

**Measurements of the Power Spectrum and Redshift  
Distortions of the Las Campanas Redshift Survey**

by

**Michael P. P. Culhane**

B. S., Stanford University, 1992

M. S., University of Colorado, 1994

A thesis submitted to the  
Faculty of the Graduate School of the  
University of Colorado in partial fulfillment  
of the requirements for the degree of  
Doctor of Philosophy  
Department of Astrophysical and Planetary Sciences

2000

This thesis entitled:  
Measurements of the Power Spectrum and Redshift Distortions of the Las Campanas Redshift  
Survey  
written by Michael P. P. Culhane  
has been approved for the Department of Astrophysical and Planetary Sciences

---

Andrew Hamilton

---

Prof. Richard McCray

Date \_\_\_\_\_

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

Culhane, Michael P. P. (Ph.D., Astrophysics)

Measurements of the Power Spectrum and Redshift Distortions of the Las Campanas Redshift Survey

Thesis directed by Professor Andrew Hamilton

I present a new method for extracting power spectra from galaxy redshift surveys called the pair weight compression method. I then apply the method to determine the redshift power spectrum of the Las Campanas Redshift Survey (LCRS). The measured redshift power spectrum has the same shape as that found by Lin et al. (1996b) but is a factor of 2 smaller. Upon expanding the method to include linear redshift distortions, the method is able to measure real space power spectra and  $\beta$ .  $\beta \approx \Omega_m^{0.6}/b$  where  $\beta$  is the linear redshift distortion parameter,  $\Omega_m$  is the cosmological density of matter and  $b$  is the linear galaxy to mass “bias factor.” Applying the method to LCRS yields a real space power spectrum that is very similar to the redshift space power spectrum. The measured value of  $\beta \approx .55_{-.30}^{+.35}$  is larger than, but consistent with, a previous measurement by Matsubara et al. (2000).

## Acknowledgements

I would like to thank my wife, family and friends for their support.

I would especially like to thank Andrew Hamilton for his help over the past several years. In particular, I thank him for his advice, patience and encouragement.

I would like to thank Dick McCray for his time and energy during the time that I worked for him. I would also like to thank him for his advice and enthusiasm over my entire graduate career.

I would like to thank Huan Lin for generously supplying me with his selection function code. I would also like to thank him for helping me to get it to work properly.

This work was supported by NASA Astrophysical Theory Grants NAG5-2797 and NAG5-7128.

## Contents

### Chapter

<b>1</b>	Introduction	1
1.1	Constraining the Universe . . . . .	1
1.1.1	Parameters . . . . .	1
1.1.2	Progress on Measuring the Parameters . . . . .	3
1.2	Pair Weight Compression Method . . . . .	6
<b>2</b>	Description of the Pair Weight Method with no Redshift Distortions	8
2.1	Pair Weight Compression . . . . .	12
2.1.1	Linear Weights . . . . .	12
2.1.2	Pair Weight Method . . . . .	14
2.2	Calculation of the Pair Weight Fisher Matrix . . . . .	17
2.2.1	Nonuniform Monte Carlo Integration of $K$ . . . . .	17
2.2.2	Step-by-Step Process of Selecting a Galaxy Foursome . . . . .	19
2.3	Comparison of Fisher Matrices . . . . .	25
2.3.1	Classical Method Versus Pair Weight Method for IRAS 1.2Jy . . . . .	25
2.3.2	Classical Method Versus Pair Weight Method for LCRS . . . . .	28
2.4	Discussion and Conclusions . . . . .	33
<b>3</b>	Application of the Pair Weight Method to the Las Campanas Redshift Survey with no Redshift Distortions	34

3.1	Preparation of Catalog . . . . .	36
3.2	Calculation of Power Spectrum . . . . .	37
3.2.1	Calculation of $W_{bkl}\delta_k\delta_l$ . . . . .	37
3.2.2	Measuring the Power Spectrum . . . . .	39
3.2.3	Correcting for Discrete FFT Errors . . . . .	41
3.3	Power Spectra . . . . .	42
3.3.1	Power Spectrum of the Full LCRS . . . . .	42
3.3.2	Power Spectrum of the Cut LCRS . . . . .	44
3.3.3	IRAS 1.2 Jy Power Spectrum . . . . .	45
3.3.4	Tests of the Pair Weight Method . . . . .	45
3.3.5	Comparison of Pair Weight Results to Previous Results . . . . .	50
3.4	Discussion and Conclusions . . . . .	50
<b>4</b>	<b>Spherical Redshift Distortions</b> . . . . .	<b>52</b>
4.1	Spherical Shape Functions . . . . .	54
4.1.1	Importance of $\alpha$ . . . . .	55
4.1.2	Shape Functions . . . . .	56
4.2	Generalized Fourier Transform Windows . . . . .	57
4.3	Pair Weight Compression with Spherical Redshift Distortions . . . . .	59
4.3.1	Shape Functions . . . . .	60
4.3.2	Window Functions . . . . .	61
4.3.3	Better Final Parameters . . . . .	61
4.3.4	Discreteness Correction . . . . .	63
4.3.5	Correlation and Decorrelation . . . . .	64
4.4	Information in LCRS . . . . .	68
4.5	Discussion . . . . .	70

<b>5</b>	Measurements of $\beta$ and the Power Spectrum from the Las Campanas Redshift Survey	72
5.1	Real Space Power Spectrum . . . . .	73
5.2	Calculated Errors . . . . .	74
5.3	Measurement of $\beta$ . . . . .	75
5.4	Discussion and Conclusions . . . . .	77
<b>6</b>	Conclusions	80
6.1	Pair Weight Method . . . . .	80
6.2	Las Campanas Redshift Survey . . . . .	80
6.2.1	Results . . . . .	81
6.3	Conclusion . . . . .	83
	<b>Bibliography</b>	84
	<b>Appendix</b>	
<b>A</b>	Calculating the Shape Functions	87
<b>B</b>	Definitions of Terms and Symbols	97
B.1	Terms . . . . .	97
B.2	Symbols . . . . .	99
B.3	Greek Symbols . . . . .	102

## Figures

### Figure

2.1	An IRAS galaxy foursome. . . . .	20
2.2	Angular selection function for LCRS . . . . .	21
2.3	Angular definitions for LCRS . . . . .	23
2.4	Measured parameters for IRAS . . . . .	27
2.5	Information in IRAS parameters. . . . .	28
2.6	Information in LCRS parameters. . . . .	30
2.7	Information in LCRS parameters: pair weight vs. Classical . . . . .	31
2.8	Measured parameters for LCRS . . . . .	32
3.1	Number of galaxies in LCRS at each distance . . . . .	36
3.2	Discretization correction . . . . .	42
3.3	Power spectrum for the full LCRS . . . . .	43
3.4	Power spectrum for the cut LCRS . . . . .	44
3.5	Power spectrum for IRAS 1.2 Jy catalog . . . . .	46
3.6	Power spectrum for a mock LCRS . . . . .	47
3.7	Power spectrum of the cut LCRS measured with a flat prior power spectrum . . . . .	48
3.8	Comparison of Power Spectra . . . . .	49
4.1	Representative rows of the fisher matrix . . . . .	65
4.2	Representative rows of the correlation matrix . . . . .	66



4.3	Representative rows of the decorrelation matrix . . . . .	69
4.4	Information in $P(k)$ , $\beta P(k)$ and $\beta^2 P(k)$ . . . . .	69
5.1	Real space power spectrum for the cut LCRS . . . . .	74
5.2	$\beta$ using correlated power spectrum . . . . .	76
5.3	Best fit amplitude for the correlated power spectrum . . . . .	77
5.4	Amplitude of power spectrum vs. $\beta$ . . . . .	78

## **Chapter 1**

### **Introduction**

#### **1.1 Constraining the Universe**

Human beings have always been fascinated by the subject of how we got to where we are and how we will get to where we are going – not just the history of human events, but also the history of the world and the universe. On the largest scales and longest time-scales this is the subject of cosmology. Many of the cosmological questions have been answered in a general way. However, there are many questions which are (seemingly) open. Specifically, the history of the universe is beginning to be generally agreed upon. It seems that the universe began, some  $13 \pm 2$  billion years ago, much smaller, denser, and hotter than today. The present data also suggest that the expansion of the universe will not only continue indefinitely but could also accelerate.

Many lines of research have come together to bring this picture into focus, including the study of galaxy redshift surveys. To understand the relevance of galaxy redshift surveys, it is instructive to look at the parameters that describe the universe and how they have been measured.

##### **1.1.1 Parameters**

Maybe the most interesting parameter to the lay-person is the ultimate fate of the universe. An expanding universe can have two ultimate conclusions. The first is a forever expanding and cooling universe. The second possibility is that the mass in the universe is sufficient

to cause the expansion to end and the universe to collapse upon itself. The parameters of chief importance in determining the fate of the universe are the omegas:  $\Omega_\Lambda$ ,  $\Omega_m$ . The omegas are ratios of the density of a particular type of matter to the critical density:

$$\Omega_a = \frac{\rho_a}{\rho_{\text{crit}}}. \quad (1.1)$$

The critical density is that density which gives an exactly flat universe. The sum of the omegas gives the shape of the universe. If the sum is greater than one, the universe is closed. If the sum, is less than one the universe is open. The universe will continue to expand indefinitely if  $\Omega_\Lambda$  is positive and  $\Omega_m$  is not significantly greater than 1 or if  $\Omega_\Lambda = 0$  and  $\Omega_m < 1$  (Carroll et al., 1992).

In addition to knowing the end result, many people are interested in how long the universe has been around. The parameter of primary importance in the question of the age of the universe is the Hubble constant ( $H_0$ ). The Hubble constant is a measure of how fast the universe is expanding. The age of the universe is given by

$$t_{\text{universe}} = \frac{1}{H_0} f(\Omega_m, \Omega_\Lambda). \quad (1.2)$$

The function  $f$  is of order unity.

Once the age of the universe and the ultimate fate of the universe are known, other cosmological parameters become interesting. For example,  $\Omega_b$ , the amount of baryons in the universe is interesting. A significant fraction of the matter in the universe is invisible, or dark. Also one would like to know what fraction of that matter is normal matter and what fraction is non-baryonic dark matter.

Studies of the universe may also yield information about the formation of galaxies. It is clear that gravity plays the largest role in galaxy formation; however, it is not clear when galaxies form and how sensitive galaxy formation is to the local environment. It is also not clear how large the “local environment” is.

## Power Spectra

The power spectrum of matter density fluctuations ( $P_k$ ) is one of the most important measurable quantities. The theory of how the power spectrum relates to the other (perhaps more interesting) parameters is quite sophisticated (Eisenstein and Hu, 1999). It allows one to measure several basic parameters (*e.g.*,  $\Omega_m$ ,  $\Omega_b$ ,  $\Omega_\Lambda$ ,  $H_0$ , etc.) from measurements of the power.

### 1.1.2 Progress on Measuring the Parameters

Before setting out to measure any parameter of the universe, it is important to note which parameters have already been constrained and by which methods.

#### Big Bang Nucleosynthesis

Big Bang nucleosynthesis is the study of which elements are produced and in what amounts from the nuclear reactions that take place in the first few minutes of the universe. The fractions of baryons in each of the light elements (H, D,  $^3\text{He}$ ,  $^4\text{He}$  and  $^7\text{Li}$ ) are sensitive to the density of baryonic matter. Current results have uncertainties in the measurement of  $\Omega_b h^2$  at the 10 per cent level (*e.g.*, Tytler et al., 2000) at  $\Omega_b h^2 = 0.019 \pm 0.0024$  ( $h$  is the Hubble constant in units of 100 km/s/Mpc). These estimates have tightened considerably in the last few years. Maybe more importantly, there are no aspects of the theory that are discrepant with observational measurements. Big Bang nucleosynthesis is one of the success stories of modern cosmology.

#### Type 1a Supernovae

In the last 2 to 3 years, the study of type 1a supernovae has yielded some of the most interesting results. Type 1a supernovae are a useful measure for cosmologists because they are believed to be extremely uniform. This means that the physics between local and more distant type 1a supernovae are the same. This allows the distance to the supernova to be measured accurately. The redshift of the object can also be measured. Having the redshift and distance to several objects allows measurements of  $H_0$  and a combination of  $\Omega_\Lambda$  and  $\Omega_m$ . Both teams trying

to measure the combination of omegas concluded that  $\Omega_\Lambda \neq 0$  (Riess et al., 1998; Perlmutter et al., 1999). In his review, Riess (2000) concludes that the supernova evidence does not yet compel us to accept an accelerating universe ( $\Omega_\Lambda > \Omega_m/2$ ); however, the accelerating universe is the most likely solution.

### **Galaxy Clusters**

Galaxy clusters can be used to measure  $\Omega_m$ . The method is to measure the mass-to-light ratio ( $M/L$ ) and the background luminosity density ( $j$ ) at the same redshift. This allows

$$\Omega_m = \frac{M}{L} \times \frac{j}{\rho_{\text{crit}}}. \quad (1.3)$$

Measurements of this kind (*e.g.*, Carlberg et al., 1998) yield  $\Omega_m \sim 0.2$  with errors at the 20 per cent level. There are two interesting points to be made about this result. First of all, it implies that a majority of the matter in the universe is non-baryonic. Second, it suggests that the density in matter is significantly lower than the critical density.

Also using galaxy clusters to probe cosmological quantities Donahue et al. (1998) observed cluster MS 1054-0321 and calculated that the probability that  $\Omega_m = 1$  is less than a few  $\times 10^{-5}$ .

### **Cosmic Microwave Background**

Measurements of the cosmic microwave background (CMB) are likely to answer most of the current cosmological questions definitively. The CMB is a great place to make a clean measurement of the power spectrum. The fluctuations in the CMB are small enough to be thoroughly treated in a linear approximation; most of the foreground contamination is small or at least removable (Tegmark et al., 2000); and the Gaussian nature of the fluctuations allows them to be completely characterized by their angular spectrum (Hu, 2000). The upcoming MAP and Planck missions should yield power spectra with sufficient resolution to measure all of the omegas,  $H_0$  and several other parameters to the accuracy of a few percent (Bond et al., 1997). A glimpse of this future was afforded to us by the BOOMERANG results. de Bernardis et al.

(2000) gives the results of the BOOMERANG experiment. The BOOMERANG experiment measured approximately 18000 pixels 14 arc-minutes in size in four frequencies. The resulting maps were then analyzed and an angular power spectrum was calculated. Their result was that  $0.88 < \Omega_0 < 1.12$  with 95 per cent confidence. Their best fit to the power spectrum had parameters  $(\Omega_b, \Omega_m, \Omega_\Lambda, n_s, h) = (0.05, 0.31, 0.75, 0.95, 0.70)$ . Already the CMB has constrained the universe to be (nearly) flat. Expect much more to come.

### Galaxy Redshift Surveys

After seeing what other methods for measuring the parameters have done, and will do soon, it is important to ask what galaxy redshift surveys can contribute. Galaxy redshift surveys are also used to measure power spectra. Unfortunately, the results are not as clean as those of the CMB. The measurement one would like is that of the matter power spectrum. Galaxy redshift surveys only yield the galaxy-galaxy power spectrum. On large scales, these power spectra differ by the square of the bias parameter ( $b$ ). On small scales, non-linearities cause more complicated departures from the matter power spectrum. Redshift surveys are also complicated by redshift distortions. Because the radial distance to an object is measured by the redshift rather than a “true” distance, local “peculiar” velocities cause errors in the distance to galaxies. At linear scales, the amplitude of redshift distortions, however, can be used as a measure of the quantity  $\beta$ .  $\beta$  is a function of the omega divided by the bias parameter:

$$\beta \approx \frac{1}{b} \left[ \Omega_m^{\frac{4}{7}} + \frac{\Omega_\Lambda}{70} \left( 1 + \frac{\Omega_m}{2} \right) \right] \quad (1.4)$$

(*e.g.*, Lahav et al., 1991).

The information extracted from the galaxy survey can be used in two ways. One way would be to make an independent measurement of the parameters. The second way is to use the parameters measured by some other technique to measure the bias parameter on all scales. The advantage that the galaxy survey has over, say, the CMB is that the power spectrum can be measured on smaller scales. However, particularly on the smallest scales, non-linearities make

parameter extraction difficult. However, using the measurements to constrain the non-linear models will yield information about how galaxies form and cluster.

## 1.2 Pair Weight Compression Method

This thesis focuses on the use of a new method, the pair weight compression method, to extract the power spectrum and measure  $\beta$  from galaxy surveys. The advantage of the pair weight compression method is that it is able to give accurate weights to each galaxy pair and accurate error bars to the results, even in the presence of redshift distortions, non-linearities and difficult data sets.

Currently favored extraction methods fall into one of two categories: the brute force methods and the classical methods. Brute force methods are the methods of choice when dealing with the largest scales (scales where the fluctuations can be treated as Gaussian). On the smallest scales, brute force techniques are too computationally expensive. Furthermore, with non-Gaussian fluctuations brute force techniques fail.

On the other hand, at small scales classical methods work the best. Classical methods rely on the selection function (expected density of galaxies contained within the catalog at a given location given the selection criteria of the survey) being slowly varying over the range of the measurement. This means that on the largest scales, classical methods are no longer ideal.

Depending on the observational procedures for a given catalog, there may be a region where both brute force and classical methods work well or there may be a region where neither technique works well at all. The pair weight method works on all size scales. Therefore, the pair weight method looks like the ideal method.

However, the pair weight method has serious disadvantages. The biggest problem with the pair weight method is that it is computationally expensive even for simple catalogs. Either of the other types of extraction method, if viable, will probably be less computationally expensive than is the pair weight method. This means that the pair weight method is likely to be the method of choice only in situations where the catalog or the prior model is complicated. In this

thesis we chose the Las Campanas Redshift Survey as an example of a catalog which has the property of being complicated enough that both brute force methods and classical methods will have a difficult time in extracting the relevant parameters (Matsubara et al., 2000).



## Chapter 2

### Description of the Pair Weight Method with no Redshift Distortions

The prospect of large, well-controlled redshift surveys such as the Two-Degree Field Survey (2dF) and the Sloan Digital Sky Survey (SDSS) and their potential for constraining cosmological parameters (Peacock 1997; Eisenstein et al. 1999) has motivated an increasing level of sophistication in the measurement of galaxy power spectra.

In recent years two approaches to measuring power spectra from galaxy surveys have come to the fore (Tegmark et al., 1998). The first is the ‘brute-force’ method, the linear maximum likelihood method pioneered by Fisher et al. (1994) and Heavens and Taylor (1995). The second is the ‘classical’ method (Feldman et al., 1994, hereafter FKP; Hamilton 2000). The two methods are complementary to each other, each performing best where the other method does the worst: the brute-force method works best at large scales, while the classical method works best at smaller scales.

Both brute-force and classical methods lay claim to being ‘optimal’ provided that certain assumptions are true. The defining assumption of the brute-force method is that density fluctuations are Gaussian. Thus the brute-force method is the method of choice at the largest, linear scales, where fluctuations may well be Gaussian. The brute force method has been applied to the *IRAS* 1.2 Jy survey by Heavens and Taylor (1995) and Ballinger et al. (1995), to the *IRAS* Point Source Catalogue redshift (PSCz) survey by Tadros et al. (1999), and to the Updated Zwicky Catalog (UZC) by Padmanabhan et al. (2000).

The defining assumption of the classical method is that the selection function  $\bar{n}(\mathbf{r})$  of a

survey, the probability of selecting a galaxy at position  $\mathbf{r}$  into the survey, is ‘slowly varying’. The classical method is good to the extent that position and wavelength of the fluctuation can be measured simultaneously. In quantum mechanics, this is the condition for a system to be classical – hence the designation. The classical method is optimal for measuring power at wavelengths much less than any characteristic scale of the survey, such as the scale over which the selection function varies.

The classical method was introduced by Feldman et al. (1994) for the case of Gaussian fluctuations. However, the method can be generalized into the nonlinear regime (Hamilton 2000). In this chapter the expression ‘classical’ refers, in general, to the nonlinear generalization of the FKP method.

While the brute-force method should be optimal at the largest scales of a survey, and the classical method should be optimal at the smallest scales, it can happen that both methods are suboptimal at intermediate scales. In the Las Campanas Redshift Survey (LCRS), for example, the width of each of the six  $1.5^\circ$  slices is  $\sim 7.5 h^{-1} \text{Mpc}$  at the median depth  $\sim 300 h^{-1} \text{Mpc}$  of the survey. Density fluctuations at this wavelength are neither fully linear nor much smaller than the scale of the survey.

The brute-force and classical estimators of the power spectrum remain unbiased (or at least asymptotically unbiased) even where they are suboptimal (Tegmark et al., 1997; Hamilton 2000). In other words, the estimators are valid estimators of the power spectrum, being biased neither high nor low, even if they are not the best estimators. However, both methods yield incorrect error bars on the power spectrum in regimes where their assumptions fail. For example, at nonlinear scales the brute-force method grossly underestimates the variance in the estimated power spectrum, by a factor  $\sim 1 + \xi$  where  $\xi$  is the correlation function. Thus, in general, it is necessary to resort to some additional procedure to compute reliable error bars. For example, error bars can be estimated from ensembles of  $N$ -body simulations (*e.g.*, Fisher et al., 1993), or empirically from the level of fluctuations observed in the data (Maddox et al., 1990; Hamilton, 1993).

The purpose of the present chapter is to explore a new method for constructing a near-optimal estimator of the power spectrum. The method goes beyond the classical approximation, but is not restricted to linear, Gaussian fluctuations. Among other things, the method is designed to yield correct error bars on the power spectrum.

Recall that in the original linear FKP approximation, the power spectrum  $P_k$  at wavenumber  $k$  is measured by weighting overdensities in pairs  $ij$  of volume elements at positions  $\mathbf{r}_i$  and  $\mathbf{r}_j$  with a pair weighting of the form

$$W_{ij} = \frac{\bar{n}_i \bar{n}_j}{(1 + \bar{n}_i P_k)(1 + \bar{n}_j P_k)} \quad (2.1)$$

where  $\bar{n}_i \equiv \bar{n}(\mathbf{r}_i)$  is the selection function at position  $\mathbf{r}_i$ . The idea of the present chapter is to admit not just a single pair weighting, as in the classical FKP method, but rather several pair weightings, judiciously chosen. In effect, the power spectrum is ‘compressed’ (Tegmark et al., 1997) into several measurements of it using different pair weightings; the several measurements of the power spectrum are then combined using their Fisher matrix. The procedure is at least somewhat familiar. For example, Fisher et al. (1993), FKP, and Sutherland et al. (1999) present measurements of the power spectrum using not one but several different pair weightings. (In the last two papers, the pair-weightings were linear FKP pair-weightings of the form of equation [2.1], with the quantity  $P_k$  in the denominator treated as an adjustable constant). The present method goes a step further by combining estimates from different pair-weightings, using their Fisher matrix.

Using the Fisher matrix to merge estimates of the power spectrum takes into account the covariance between the estimates and automatically yields the best possible combination of the estimates. In principle, an optimal measurement of the power spectrum could be obtained in the limit of many estimates of power spectra derived from a large number of different pair weightings. In practice, numerics limit the number of pair weightings that can be handled to only a few. Still, even just two sets of pair weightings is better than a single pair weighting.

For simplicity, and to facilitate comparison with other methods, the present chapter is

limited to the case of Gaussian fluctuations without redshift distortions. We include redshift distortions in later chapters and hope to include non-linearity in future work.

There is a parallel between the approach proposed in the pair weight method and the linear brute-force method for Gaussian fluctuations. In the linear brute-force method, the data in a galaxy survey, the overdensities  $\delta_i$  in many volume elements  $i$ , are compressed into a set of linearly weighted modes  $\hat{x}_a \equiv W_{ai}\delta_i$ . To retain as much information as possible, the number of modes should be as large as possible; but numerical tractability limits the number of modes to several  $\times 10^3$  or possibly several  $\times 10^4$ . The game is then to craft the modes so as to cram the largest amount of information about parameters of interest (*e.g.*, the amplitude  $P_k$  of the power spectrum at several linear wavenumbers  $k$ , and the redshift distortion parameter  $\beta \approx \Omega^{0.6}/b$ ) into the smallest number of modes (Heavens and Taylor 1995; Tegmark et al. 1997; Padmanabhan et al. 2000). Here, the data are compressed, instead, into a set of quadratically weighted modes  $\hat{X}_a = W_{aij}\delta_i\delta_j$ . The game is similar: try to choose the pair windows  $W_{aij}$  so as to cram the largest amount of information into the smallest number of pair modes.

One advantage of a quadratic weighting method is that, unlike the linear brute-force method, it is not limited to Gaussian fluctuations. Thus if a nonlinear prior is used (which, however, is not done in this work), then the quadratic weighting method will yield reliable error bars on the nonlinear power spectrum; whereas the linear weighting method will underestimate the true error bars in the nonlinear regime. A disadvantage of the quadratic weighting method is that it is numerically much more involved to compute the Fisher matrix of a set of pair weighted modes, so that in practice only a handful of pair modes can be computed. The situation is not hopeless, however, because a single pair weighting – the linear FKP pair weighting, equation (2.1) – probably already contains a large fraction of the information of interest (Heavens et al., 2000). Adding just one additional pair window is guaranteed to increase the information content of the estimate and to reduce the error bars correspondingly. The gain should be greater in surveys with more complicated selection functions, such as the LCRS.

We describe the pair weight compression method in section 2.1. In section 2.2 we de-

scribe how we actually go about calculating the matrices necessary to obtain the measurements. Section 2.3 compares the Fisher matrices for IRAS 1.2 Jy and LCRS using the classical method and the pair weight compression method. In section 2.4 we discuss the method, our results, and possible applications.

## 2.1 Pair Weight Compression

We introduce here, in subsection 2.1.2, the pair weight compression scheme for measuring the power spectrum from a galaxy survey. We begin by reviewing, in subsection 2.1.1, the linear compression scheme (Tegmark et al., 1997), which is the basis of the standard ‘brute-force’ procedure.

### 2.1.1 Linear Weights

The data contained in a galaxy survey are the values  $n_i$  of the galaxy number density in each of the infinitely many infinitesimal volume elements  $dV_i$  of the survey. To the extent that the selection function  $\bar{n}_i$  is known or can be measured with negligible uncertainty (Binggeli et al., 1988; Willmer, 1997; Tresse, 1999), the data can be taken to be the overdensities  $\delta_i \equiv (n_i - \bar{n}_i)/\bar{n}_i$ .

Suppose that these overdensities are compressed into a set of mode amplitudes  $\hat{x}_a$  (the hat signifies an estimated quantity, distinguishing it from a prior quantity) by weighting the overdensities with a linear weighting function  $W_{ai}$

$$\hat{x}_a \equiv W_{ai}\delta_i \tag{2.2}$$

(repeated Latin indices in eq. [2.2] and hereafter signify implicit integration over the volume element, so  $W_{ai}\delta_i = \int W_a(\mathbf{r})\delta(\mathbf{r})dV$ ; see section 2 of Hamilton, 2000 for an exposition). The covariance of the modes is

$$C_{ab} \equiv \langle \hat{x}_a \hat{x}_b \rangle = W_{ai} \langle \delta_i \delta_j \rangle W_{bj} \tag{2.3}$$

where  $\delta_i \delta_j$  is the expected covariance function of the survey, which is a sum of a signal term, the cosmic correlation function  $\xi(r_{ij})$ , and a noise term, commonly taken to be dominated by Poisson sampling noise

$$\langle \delta_i \delta_j \rangle = \xi_\alpha B_{\alpha ij} + \bar{n}^{-1}(\mathbf{r}_i) \delta_{ij} \quad (2.4)$$

the correlation function with respect to the linear combination of individual parameters  $\xi_\alpha$ .

Any linear combination of Gaussian fields is Gaussian, so if the overdensity field  $\delta_i$  is Gaussian – the defining assumption of the linear method – then the mode amplitudes,  $\hat{x}_a$  will also form a multivariate Gaussian. Gaussian fields have the advantage that the likelihood function can be written down explicitly:

$$\mathcal{L} = \frac{1}{|C|^{1/2}} \exp \left[ -\frac{1}{2} \hat{x}_a C_{ab}^{-1} \hat{x}_b \right]. \quad (2.5)$$

The maximum likelihood estimate  $\hat{\xi}$  of the parameters is given by the vanishing of the derivative of  $\ln \mathcal{L}$  with respect to the parameters:

$$\frac{\partial(\ln \mathcal{L})}{\partial \xi_\alpha} = \frac{1}{2} [C_{ab}^{-1} C_{bc, \alpha} C_{cd}^{-1} (\hat{x}_d \hat{x}_a - C_{da})] \quad (2.6)$$

$$\left. \frac{\partial(\ln \mathcal{L})}{\partial \xi_\alpha} \right|_{\xi=\hat{\xi}} = 0. \quad (2.7)$$

If the parameters  $\xi_a$  are the values of power spectrum at many different wavenumbers  $k_a$ , then the covariance  $C_{ab}$  depends linearly on the parameters

$$C_{ab} = \frac{\partial C_{ab}}{\partial \xi_\alpha} \xi_\alpha + N_{ab}. \quad (2.8)$$

The maximum likelihood estimator  $\hat{\xi}_\alpha$  of the power spectrum is then

$$\hat{\xi}_\alpha = F_{\alpha\beta}^{-1} C_{ab}^{-1} C_{bc, \beta} C_{cd}^{-1} \hat{x}_a \hat{x}_d \quad (2.9)$$

where  $F_{\alpha\beta}$  is the Fisher matrix of the parameters  $\xi_\alpha$

$$F_{\alpha\beta} \equiv C_{ab}^{-1} C_{bc, \alpha} C_{cd}^{-1} C_{da, \beta}. \quad (2.10)$$

Equations (2.9) and (2.10) provide the best (maximum likelihood) estimate of the power spectrum  $\xi_\alpha$  that can be deduced from the given set of modes:  $a$ . If the modes formed a complete set (which would require infinitely many modes), then the estimator would be optimal. However, computational tractability limits the number of modes that can be treated to several  $\times 10^3$  or possibly several  $\times 10^4$ . Thus, a key part of the brute-force method is to try to choose the modes to contain as much information as possible about the parameters of interest, the values of the power spectrum at large, linear scales. This can be achieved by crafting the modes wisely (Heavens and Taylor, 1995) and perhaps also by performing some kind of Karhunen-Loève (signal-to-noise) compression (Vogeley and Szalay, 1996; Tegmark et al., 1997; Tegmark et al., 1998; Padmanabhan et al., 2000).

### 2.1.2 Pair Weight Method

In the pair weight scheme, the overdensities are compressed instead into a set of pair-weighted modes

$$\hat{X}_a \equiv W_{aij} \delta_i \delta_j. \quad (2.11)$$

This seems like a natural thing to do since, after all, the power spectrum is itself an expectation value of a covariance of overdensities. Indeed, the traditional way to measure the power spectrum is based on equation (2.11) with just a single pair weighting  $W_{aij}$ , suitably normalized (that is, one estimates  $\hat{\xi}_\alpha = W_{\alpha ij} \delta_i \delta_j$ , the shot noise being removed by excluding the contribution of self-pairs of galaxies to the estimator). Commonly, some variant of the linear FKP pair-weighting, equation (2.1) is adopted.

The expectation value of the amplitude  $\hat{X}_a$  of the pair mode is

$$\langle \hat{X}_a \rangle \equiv X_a = W_{aij} C_{ij} = W_{aij} B_{\alpha ij} \xi_\alpha + W_{aij} \delta_B(r_{ij}) \bar{n}^{-1}(\mathbf{r}_i). \quad (2.12)$$

The covariance is

$$\langle \Delta \hat{X}_a \Delta \hat{X}_b \rangle = W_{aij} \langle \Delta C_{ij} \Delta C_{kl} \rangle W_{bkl} \equiv W_{aij} \langle \mathbf{c}_{ijkl} \rangle W_{bkl} \equiv K_{ab}. \quad (2.13)$$

In a sufficiently large survey, the central limit theorem implies that the estimator  $\hat{X}_a$  will be Gaussianly distributed about its expectation value. This is true irrespective of whether the underlying density field is Gaussian or not. In practice, surveys typically do contain enough information that  $\hat{X}_a$  will be near Gaussian at all but the largest scales, where the linear brute-force method is the method of choice. We assume therefore that  $\hat{X}_a$  are Gaussianly distributed, so that the likelihood function is

$$\mathcal{L} \propto \frac{1}{|K|^{1/2}} \exp \left[ -\frac{1}{2} (\hat{X}_a - W_{aij} B_{\alpha ij} \xi_\alpha) K_{ab}^{-1} (\hat{X}_b - W_{bkl} B_{\beta kl} \xi_\beta) \right]. \quad (2.14)$$

Take the derivative of  $-\ln \mathcal{L}$  with respect to  $\xi_\alpha$  and set the result to zero to find the estimate of  $\xi_\alpha$  given the set of pair weights in the model.

$$\begin{aligned} \frac{\partial(-\ln \mathcal{L})}{\partial \xi_\alpha} &= \frac{1}{2} K_{dc}^{-1} K_{dc, \alpha} - \\ &\frac{1}{2} (\hat{X}_a - \xi_\alpha B_{\alpha ij} W_{ija}) K_{ac}^{-1} K_{cd, \alpha} K_{db}^{-1} (\hat{X}_b - W_{bkl} \xi_\beta B_{\beta kl}) - \\ &W_{aij} B_{\alpha ij} K_{ab}^{-1} (\hat{X}_b - W_{bkl} \xi_\beta B_{\beta kl}). \end{aligned} \quad (2.15)$$

Take the first two terms and rewrite them as

$$\frac{1}{2} K_{ac}^{-1} K_{cd, \alpha} K_{db}^{-1} [K_{ba} - (\hat{X}_b - W_{bkl} \xi_\beta B_{\beta kl}) (\hat{X}_a - \xi_\alpha B_{\alpha ij} W_{ija})]. \quad (2.16)$$

In order to perform the full maximum likelihood calculation these terms must be calculated explicitly. Bond et al. (1998) show that it is possible to carefully include the contribution from these terms. However, Tegmark et al. (1997) shows that the estimator is asymptotically unbiased even if it is suboptimal. In the minimum variance case, the derivatives of  $K$  are set to zero and these terms vanish. In the remainder of this thesis, we take the minimum variance solution rather than the full maximum likelihood solution. This leaves:

$$-W_{aij} B_{\alpha ij} K_{ab}^{-1} (\hat{X}_b - W_{bkl} \hat{\xi}_\beta B_{\beta kl}) = 0. \quad (2.17)$$

Rearrange this equation and replace  $\hat{X}_b$  with  $W_{bkl} \delta_k \delta_l$  to get:

$$B_{\alpha ij} W_{aij} K_{ab}^{-1} W_{bkl} B_{\beta kl} \hat{\xi}_\beta = B_{\alpha ij} W_{aij} K_{ab}^{-1} W_{bkl} \delta_k \delta_l. \quad (2.18)$$



Define a new matrix (the Fisher matrix)  $M$ :

$$M_{\alpha\beta} \equiv B_{\alpha ij} W_{aij} K_{ab}^{-1} W_{bkl} B_{\beta kl}. \quad (2.19)$$

In order to get the final estimate of  $\xi_\alpha$ , multiply both sides of equation (2.18) by the inverse of  $M$ . The final result is

$$\hat{\xi}_\beta = M_{\beta\alpha}^{-1} B_{\alpha ij} W_{aij} K_{ab}^{-1} W_{bkl} \delta_k \delta_l. \quad (2.20)$$

This gives us a complete method for calculating  $\hat{\xi}_\alpha$  using the multiple pair weights  $W_{aij}$ . We can see that we no longer have to invert the  $C$  matrix directly. On the other hand, we must make sure that the computation of the matrix  $K$  is feasible. If we assume Gaussian fluctuations then

$$\langle \mathfrak{C}_{ijkl} \rangle \approx C_{ik} C_{jl} + C_{il} C_{jk}. \quad (2.21)$$

The matrix  $K$ , Equation 2.13, can then be rewritten as:

$$K_{ab} \approx 2W_{aij} C_{ik} C_{jl} W_{bkl}. \quad (2.22)$$

It is not necessary to make this particular approximation. However, we do need a prior estimate for  $\langle \mathfrak{C}_{ijkl} \rangle$ .

In principle, the pair weighting method can save computation time over a brute force technique, particularly for size scales significantly smaller than the size of the catalog. For example, doubling the resolution in the pair weight method takes four times as much computing capacity, instead of the  $\sim 2^9$  (based on the computing cost of matrix inversion which is an  $n^3$  process) times for a brute force method. (The computation of  $K$  is the slowest part, and on average each matrix element takes the same amount of time. Doubling the resolution requires four times as many matrix elements.) The down sides are that the creation of the matrix  $K$  loses information if the choice of pair weights is not complete, and the matrix  $K$  can also be quite computationally expensive. The matrix  $M$  can be used as the Fisher information matrix.

$$M_{\alpha\beta} \equiv B_{\alpha ij} W_{aij} K_{ab}^{-1} W_{bkl} B_{\beta kl} = B_{\alpha ij} W_{aij} W_{amn}^{-1} F_{mnop} W_{bop}^{-1} W_{bkl} B_{\beta kl} \quad (2.23)$$

where  $F_{mnop}$  is the Fisher matrix with respect to parameters  $\sum_{\alpha} \xi_{\alpha} B_{\alpha mn}$  and  $\sum_{\beta} \xi_{\beta} B_{\beta op}$ . To the extent that the choice of windows is a complete set

$$W_{aij} W_{amn}^{-1} \approx \delta_{im} \delta_{jn}. \quad (2.24)$$

This means that by applying the chain rule and noting that  $\frac{\partial C_{ij}}{\partial \xi_{\alpha}} = B_{\alpha ij}$

$$M_{\alpha\beta} \approx B_{\alpha ij} F_{ijkl} B_{\beta kl} = F_{\alpha\beta}. \quad (2.25)$$

This means that we can compare the pair weight method to another method by comparing the resulting Fisher matrices. In the case of the pair weight method we will insert  $M$  in place of  $F$ . If  $M$  is identical to the Fisher matrix as computed by an exact method then we know that the pair weight method loses no information. We will see in the next section that the difficulty in the pair weight method lies in the computation of  $K$ .

## 2.2 Calculation of the Pair Weight Fisher Matrix

Equation (2.25) says that the matrix  $M$  is equivalent to the Fisher matrix. Calculating  $M$  requires the matrices  $K_{ab}$  and  $B_{\alpha ij} W_{aij}$ . This section discusses the method of calculation for  $K_{ab}$  since it is by far the more difficult of the two. The calculation of  $K$  requires a 10-dimensional integral (equation 2.22), to be performed. (Each of the four volume elements contributes 3 dimensions and each window function  $W_{aij}$  and  $W_{bkl}$  involves a delta function which effectively removes a dimension.) for each pair of values  $a$  and  $b$ . There are two serious complications: the volume elements are linked by the correlation function of the separation of two pairs of two elements (this makes it impossible to reduce the 10-dimensional integral to a product of smaller integrals) and the lack of full-sky coverage. (This removes possible symmetries making analytical progress difficult). We therefore proceed by Monte Carlo techniques.

### 2.2.1 Nonuniform Monte Carlo Integration of $K$

Monte Carlo integration is a technique by which one integrates by randomly selecting points within the range of integration and computing the average value of the integrand (as

well as the average of the integrand squared) over those points. For a good introduction to the subject, see *Numerical Recipes* (Press et al., 1986). Monte Carlo integration relies on the idea that

$$\int \rho dr = \bar{\rho} \int dr \quad (2.26)$$

for the average  $\bar{\rho}$ . We can then approximate

$$\bar{\rho} \approx \frac{1}{n} \sum_{i=1}^n \rho(r_i) \quad (2.27)$$

where the values  $r_i$  are randomly chosen to be within the range of the integral. The accuracy to which the integral can be calculated depends on how well the average of the points selected represents the true average. Use the average of the integrand squared to estimate the accuracy of the calculation:

$$\Delta \bar{\rho} \approx \left[ \frac{\frac{1}{n} \sum_n \rho^2(r_i) - \bar{\rho}^2}{n} \right]^{\frac{1}{2}}. \quad (2.28)$$

There are two ways to improve the accuracy of the Monte Carlo integration. First, choose more points (increase  $n$ ). As the number of points increases, the average tends towards the true average. Second, we would like  $\rho$  to be as close to constant as possible within the range of integration (so that  $\rho^2$  is very nearly  $\bar{\rho}^2$  for all points) thereby decreasing the numerator in equation (2.28). To do this, change the variable of integration to smooth out  $\rho$ . For example, suppose that  $\rho = r^2(1 + \epsilon(r))$ , where  $|\epsilon| \lesssim 1$ , and the integral runs from  $r = 0$  to  $r = 2$ . Change the integration variable from  $r$  to  $u$  where  $u = r^3$ .

$$\int_0^2 \rho(r) dr = \frac{1}{3} \int_0^8 (1 + \epsilon(u^{\frac{1}{3}})) du. \quad (2.29)$$

By changing the variable from  $r$  to  $u$ , the integrand becomes very nearly constant. This is the equivalent to putting many points in the regions where the integrand is highest and down-weighting the contribution from each of the points in this region. In order to maintain the accuracy of the integral, lower the probability of choosing a point in a low density region and

correspondingly raise the weight given to that point. By smoothing out the contributions from each point, the integral can be calculated to the same accuracy with fewer points.

$K_{ab}$  is a whole matrix of integrals ( $a$  and  $b$  are matrix indices), so they need to be calculated quickly. To this end, choose integration variables to smooth out the entire integral. Fortunately, for the integral  $K$ , this is possible. Equation (2.28) says that the values of  $\rho$  that increase  $\Delta\bar{\rho}$  the most are the largest absolute values. Even though the integral is quite complicated, it is not hard to see which configurations of the four volume elements contribute the most to the integral. First of all, the correlation function is largest when the two volume elements are close to one another, so choose the volume elements in such a way that close pairs (within the volume foursomes) are more likely than more distant pairs. Secondly, the selection function and the chosen window functions vary with distance to the observer. Choosing the volume elements according to the window functions smoothes out the integrand. When the integrand is near zero or exactly zero (outside of the catalog), it contributes relatively little to the error. This only slows down the integral if a high fraction of foursomes has an element outside of the catalog.

### 2.2.2 Step-by-Step Process of Selecting a Galaxy Foursome

To see how the Monte Carlo integration works, it makes sense to look at the process of selecting one galaxy foursome. Equation 2.22 shows the integral that is being performed by Monte Carlo techniques. Each of the indices which is summed over can be thought of as the location of a galaxy. Because there are four summation indices, each location (in twelve-dimensional space with two separations fixed) can be thought of as a galaxy foursome. Figure 2.1 shows one such galaxy foursome for the case of IRAS.

The biggest goal in selecting one of the random variables is to make sure that the selection takes into account everything known about the integral up to that point. This allows the most effective reduction of the largest peaks. Obviously, the first step is to choose the first galaxy. At this point, use the angular mask to make sure that the galaxy will be within the catalog or at least has a good chance of being within the catalog. For a catalog like IRAS 1.2-Jy,

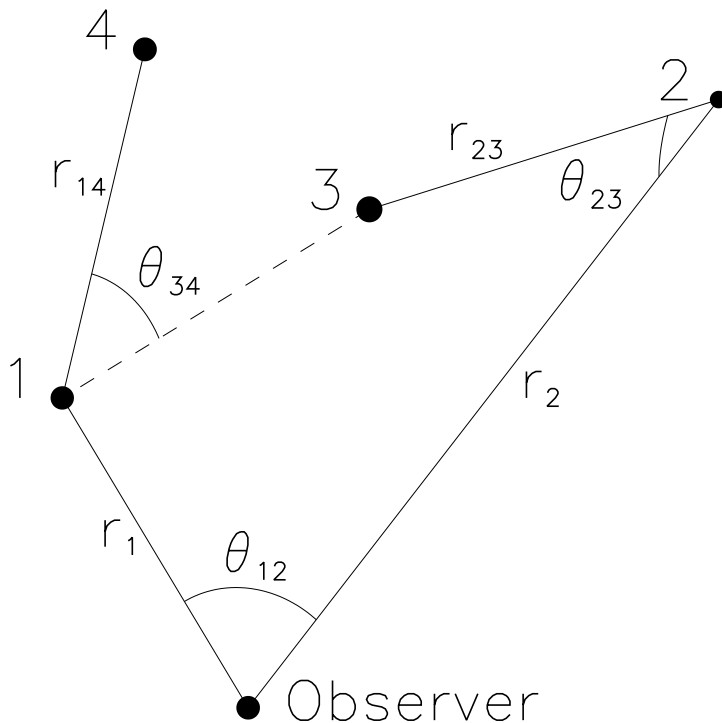


Figure 2.1: A stick figure representation of a galaxy foursome for IRAS.

the angular mask is not a significant issue since more than 90 percent of the sky is within the catalog.

On the other hand, for a catalog like LCRS (less than one steradian of sky coverage), choosing the angular position carefully greatly speeds up the calculation. In LCRS, choose the angular position to lie in one of the six slices in such a way that equal angular areas get equal weight. Figure (2.2) shows the LCRS angular mask. Each square represents one observing area. The fraction of galaxies for which redshifts were obtained is shown by the darkness of the square. (If all galaxies have redshifts then the square is black.) Cross-hatched areas are from

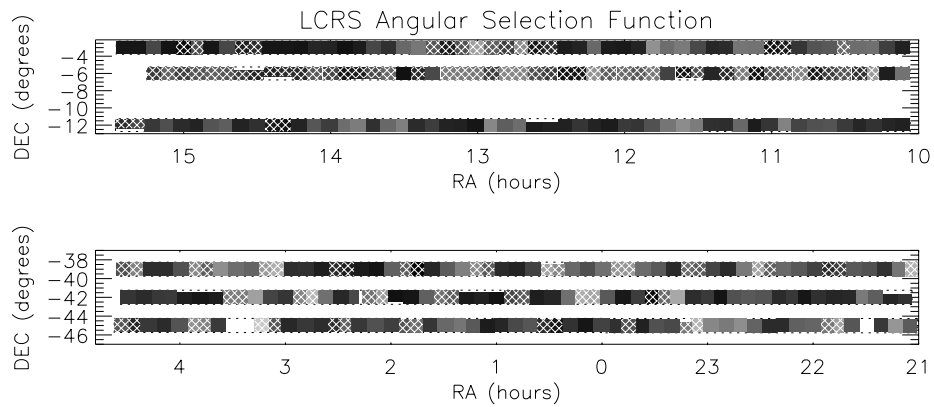


Figure 2.2: The angular selection function for LCRS. The grayscale shows the fraction of galaxies within the region for which redshifts are included. (Black is 100 per cent and white is zero.) Areas which are cross-hatched are from the 50-fiber instrument whereas solid regions are from the 112-fiber instrument.

the first stage of the catalog where 50 redshifts could be obtained simultaneously. Solid squares are from the second stage where 112 redshifts could be taken simultaneously. The dashed lines represent the outer boundaries of the slices. If the galaxy is within one of the slices, it has a very high probability (again over 90 percent) of being in the catalog. Because the slices are not complete, make sure the actual angular position is, in fact, within the catalog. If it is not, then give this galaxy foursome zero weight and move on to the next galaxy foursome. Now go about choosing the radial position of the first galaxy. Make sure that the radial position reflects the chosen window for the matrix element in question. To do this, fit the window function (times the volume at that radius) with a series of power laws. This allows rapid computation of the randomly selected radial position, while at the same time smoothing out the radial peaks. In fact, the distribution of probabilities and the corresponding weights for the radial position of the first galaxy can be precomputed.

In selecting the second galaxy, take into account not only the information about the catalog as a whole but also the position of the first galaxy. The information about the first galaxy is important because one product of the integrand is the correlation function at the separation between galaxies one and two. This means that the integrand is, in general, higher if galaxies

one and two are close together. To eliminate the peaks, increase the probability that these two galaxies are close together and decrease the probability that they are far apart. First select the radial component of the galaxy. Rather than just using the power law fits to the window function, modify the power laws such that the region near the radial component of the first galaxy is more likely to be selected for the second galaxy as well. In other words, increase the power of the power laws for regions smaller than  $r_1$  and decrease the power of the power laws for regions greater than  $r_1$ .

After choosing the radial position, the strategy for choosing the angular position is different for IRAS than it is for LCRS due to the nature of their angular masks. In the case of IRAS, calculate the correlation function for a series of points for given test angles ( $\theta_{12}$ ) between galaxy one and galaxy two. Then fit a series of power laws to the absolute values of the correlation function. (Don't allow any of these points to drop below a lower boundary for two reasons. First, the correlation function can go through zero and negative values can cause problems. Second, giving one section too little probability can cause a new peak to form when a point within this improbable region is chosen due to its large up-weighting.) Then choose the angle between galaxies one and two using these power laws as the probability. The azimuthal angle is chosen with all angles receiving equal probability.

For LCRS, choose the slice and then the location within that slice. Choose the slice with the probabilities given by the angular area of the slice times the maximum of the correlation function with the galaxy in that slice. Choose the latitude within that slice uniformly. (The slices are only  $1.5^\circ$  thick so there is not too much variation over the slice.) The longitudinal angle is found by evaluating the correlation function as if the galaxy were at a few locations along the slice. Then, apportion the probability to locations according to the corresponding values of the correlation function. Again, check that the particular angular position of the galaxy is within the valid portion of the catalog.

The third and fourth galaxies are connected to the first two galaxies by the fixed lengths  $r_{23}$  and  $r_{14}$ . Therefore, only two angles are necessary to define the location of each of these

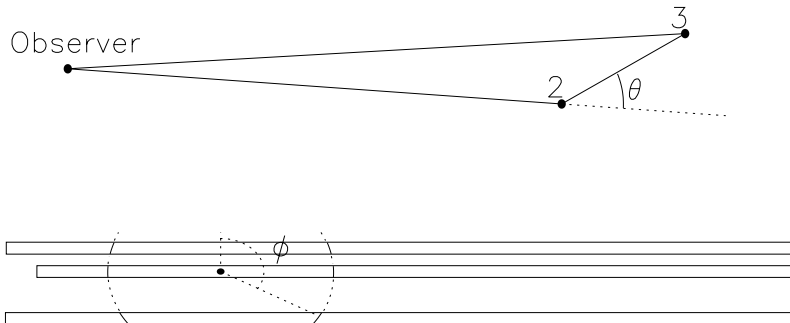


Figure 2.3: Top: definition of  $\theta$  for LCRS. Bottom: depiction of possible choices of  $\phi$  for a given value of  $\theta$ .

galaxies. Define these two angles as  $\theta_{23}$  and  $\phi_{23}$ .  $\theta_{23}$  is the latitudinal angle with its north pole pointing directly toward the center from galaxy two. For IRAS, the contribution to the integral is normally highest if  $\theta_{23}$  is low. This is because the selection function is highest closest to the center. Calculate the selection function for various values of  $\theta_{23}$  and weight the probabilities by these values. Then choose  $\phi_{23}$  with uniform probability.

The situation for LCRS is substantially different. Due to the observing strategy, the selection function is not guaranteed to be larger closer to the center. Therefore, select  $\theta_{23}$  uniformly. Figure (2.3) shows a particular value for  $\theta_{23}$  in the top panel. Use the positions of the slices to guarantee that the value of  $\phi_{23}$  will place the galaxy within one of the slices in the survey (as shown in the bottom panel of Figure 2.3). Once the angular position of galaxy three is known, check to see that it is within the survey. If it is not, (or if no value of  $\phi_{23}$  is within the survey), give this galaxy foursome zero weight and move on to the next foursome.

In the case of the IRAS survey, there are two competing goals in choosing the location of the final galaxy. Positions close to the observer have large selection functions, while at the same time, positions close to galaxy three maximize the correlation function. For simplicity, choose to do one or the other based on the relative importance of the two. If the closest possible separation for galaxies three and four is less than twenty times (empirically the best trade-off)



the smallest possible distance to galaxy four (from the observer), then try to get galaxy four close to galaxy three, as shown in Figure (2.1). Otherwise, try to get it close to the observer. In either case, set up the coordinates such that  $\theta_{34}$  is zero when it minimizes the relevant distance. Then, follow the same procedure for choosing  $\theta_{34}$  as for  $\theta_{23}$ . Once again  $\phi_{34}$  is chosen uniformly.

For LCRS, the selection function is not always decreasing with radius. Therefore, always try to get galaxies three and four close together. The mathematics are simplest if  $\theta_{14} = 0$  points directly away from the center. Again, use Figure (2.3) to see how this location is chosen. First, compare the value of the correlation function at the minimum attainable distance to the correlation functions at  $\theta_{14} = 0$  and  $\theta_{14} = \pi$ . If it is significantly larger (more than a factor of 3 or so), then calculate the location of intermediate values (approximately factors of the square root of 10) of the correlation function. Then choose  $\theta_{14}$  by the probabilities given by a power law fit to these values and locations. If the variation of the correlation is small, choose  $\theta_{14}$  uniformly. To calculate  $\phi_{14}$ , first determine which values of  $\phi_{14}$  lie in one of the slices. Then multiply the length within each slice by the value of the correlation function at the minimum separation available in that slice. This is possible because the minimum separation in the slice will be at one of the ends of the slice or at the minimum separation possible given  $\theta_{14}$ . Use this value to be the probability of choosing this slice. Once the slice is determined, compare the highest value of the correlation function within the slice to the lowest value. If the highest value is more than three times as high as the lowest, compute the positions along the slice of the predetermined values of the correlation function. (If it is not, again use a uniform distribution.) Once again, fit a power law to the determined values and locations and call this the probability distribution (this time for  $\phi_{14}$ ). Once the calculation of the position of galaxy four is complete, determine whether or not galaxy four is within the catalog.

If all of the galaxies are within the catalog, then compute the value of the integrand at that set of positions. In particular, take the product of the correlation functions of the separations ( $r_{14}$  and  $r_{23}$ ), the window functions, and the accumulated renormalizations due to the non-uniform probability functions for each variable. Then add this to the running total. Continue choosing

galaxy foursomes until the statistical error is below the error limit, at which point the value of the integral is the running total divided by the number of foursomes (including foursomes which were not in the catalog) times the 10-dimensional volume of the integral. This, of course, gives the value of the signal-signal portions of the matrix  $K$ . To find the full value of the matrix add into the matrix the components due to the delta functions (the 7-dimensional and 5-dimensional components). Then use this matrix in our calculation of the Fisher matrix.

### 2.3 Comparison of Fisher Matrices

By comparing the Fisher matrices calculated by two different methods, one can compare the relative accuracies of the two methods in measuring a particular set of parameters. In particular, in this section we compare the Fisher matrix as calculated by the pair weight method to an accepted method, the classical approximation, for the parameters  $P(k)$ . This will allow a comparison of the two methods to see whether or not the more complicated pair weight method is also the more accurate method. In particular, we would like to know which circumstances demand the pair weight method and which can be addressed by the more easily computed classical method. Note that in this comparison we are comparing the Fisher matrices of Gaussian models with no redshift distortions.

#### 2.3.1 Classical Method Versus Pair Weight Method for IRAS 1.2Jy

The IRAS 1.2 Jy redshift survey is nearly all-sky. This means that the classical approximation should work very well. This makes it a good choice for comparison between the pair weight method and the classical method. To generate the Fisher matrices, we need to make assumptions about the prior prediction of the power spectrum (or, equivalently, about the correlation function). Our power spectrum was generated by the methods described in Eisenstein and Hu (1999) with  $h = 0.65$ ,  $\Omega_\Lambda = 0.7$ ,  $\Omega_b = 0.3$ ,  $n = 1$ , and  $T = 2.73$  K. For this comparison, we use a Fisher matrix calculated as described in Hamilton (1997) for the classical method. For

the pair weight method, we use one window at each separation. The window is given by

$$W_{\alpha ij} = \frac{\bar{n}_i \bar{n}_j \delta(r_\alpha - |\mathbf{r}_i - \mathbf{r}_j|)}{(1 + 4\pi J_{3\alpha} \bar{n}_i)(1 + 4\pi J_{3\alpha} \bar{n}_j)} \quad (2.30)$$

where

$$J_{3\alpha} = \int_0^{r_\alpha} r^2 \xi(r) dr. \quad (2.31)$$

We expect this window to reflect the information content because we are counting equal volumes in regions where the distribution is well sampled and equal numbers where it is sparsely sampled.

To compare the information contained within the Fisher matrix, we would like to find statistically independent variables and find the amount of information about each of these variables contained within the data set. Hamilton and Tegmark (2000a) describes a number of ways to decorrelate the data. In this paper we will use the Fisher matrix divided by the prior value of the power spectrum at the particular  $k$  (the fractional Fisher matrix) rather than the Fisher matrix itself.

$$G_{\alpha\beta} \equiv \left[ \xi_\alpha^{-1} \langle \Delta \hat{\xi}_\alpha \Delta \hat{\xi}_\beta \rangle \xi_\beta^{-1} \right]^{-1} = \xi_\alpha F_{\alpha\beta} \xi_\beta. \quad (2.32)$$

This is so that the fractional error is the error in question rather than the absolute error. This means that the region where  $\xi$  is large will not be over-represented in regions where  $\xi$  is small. To decorrelate this matrix, measure the parameters which are the rows (normalized to unit sum) of the square-root of this matrix. This ensures that these parameters will be decorrelated with eigenvalue equal to the unnormalized sum of the row squared ( $[\sum_i F_{ij}^{1/2}]^2$ ). Compare two quantities between the two matrices. The first is addressed in Figure 2.4. This Figure shows a set of normalized rows of the square-root of the Fisher matrices. (Solid is the pair weight method and dashed is the classical method.) These are the actual parameters that we will measure. Notice two things about this Figure. First, the parameters have narrow peaks centered on the wavelength in question. This means that the decorrelated parameter does a good job of measuring the value of the power spectrum at a particular  $k$ . Second, notice that the peak of the

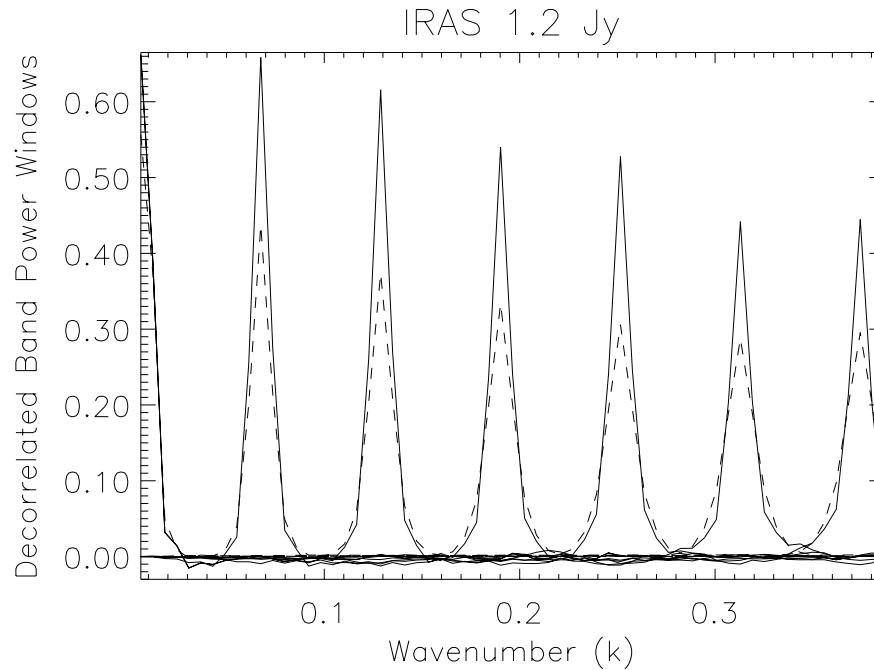


Figure 2.4: Normalized rows of the square-root of the Fisher matrix. Solid lines are for the pair weight method and dashed are for the classical method.

parameter is higher for the pair weight method than it is for the classical method. This is due to the fact that the pair weight method has values off the diagonal which are negative. These negative values are likely to be caused by errors in the calculation of the Fisher matrix rather than actual regions of negative information. Because the Fisher matrix is used in the calculation of the parameters from the data as well as to estimate the accuracy of the calculation, errors in the calculation of the Fisher matrix will reduce the accuracy with which the parameters in question can be measured.

Figure 2.5 depicts the amount of information contained within the IRAS catalog in the decorrelated band-powers. The classical method extracts as much (or more) information about the parameters on all scales as does the pair weight compression method with one window per separation. This is not too surprising due to the fact that the IRAS catalog is nearly all-sky. This means that the classical approximation should be very nearly valid. This also means that

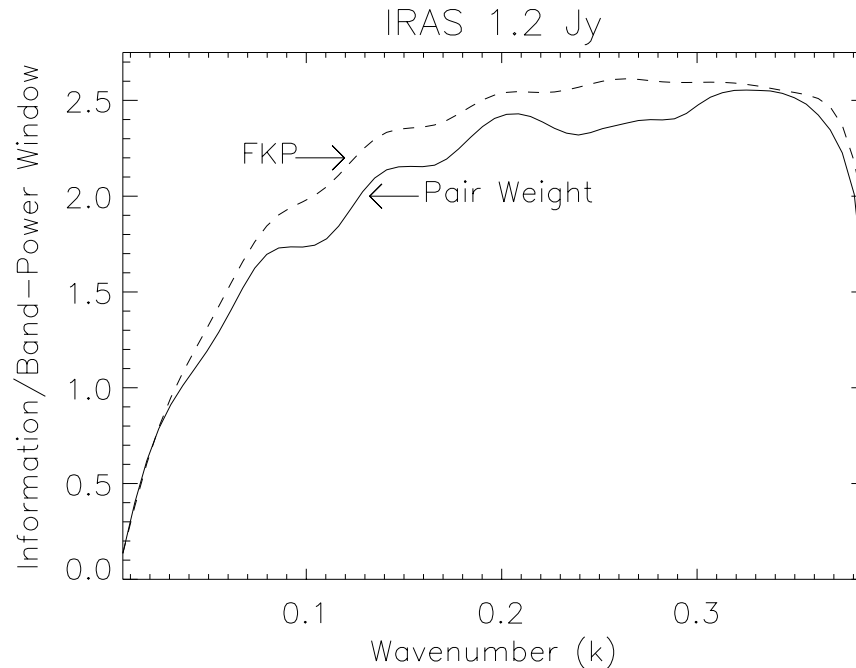


Figure 2.5: Information contained within decorrelated parameters. The  $k$  index gives the wavenumber of the location of the peak of the parameter. The solid line is for the pair weight method and the dashed line is for the classical method.

for a catalog such as IRAS, one should strongly consider using the simpler classical method on the small to moderate scales. On the largest scales, one should use one of the brute-force techniques. These techniques should eliminate the approximation used in the classical method while still being feasible.

### 2.3.2 Classical Method Versus Pair Weight Method for LCRS

The Las Campanas Redshift Survey (Shectman et al., 1996) allows us to compare the pair weight method to the classical method in a situation where we expect the classical approximation to break down. LCRS has 327 fields approximately  $1.5^\circ \times 1.5^\circ$  contained within 6 slices. In each of these fields up to 112 redshifts were taken. Due to the limitations of the observing procedure, each field has its own selection function. (The selection function depends upon the fraction of galaxies in the field that have redshifts. This fraction varies from field to field.) This

means that the largest length scale over which the catalog does not vary is  $1.5^\circ$  ( $\approx 7.5h^{-1}$  Mpc at the median depth of the survey). On scales larger than this the classical approximation should lose some of the available information. The complications of the LCRS catalog also make many of the brute force calculations difficult, though not impossible.

In addition to probing a region where the classical method should be inadequate, we wanted to investigate the benefits of using more than one window at each separation. Using more than one set of windows, should enable the extraction of more information than with only one set. We would like to know whether the additional information contained within the additional windows is enough to justify the computing cost. The windows we used for this are the same as those defined in equation 2.30 with four different choices of  $J_{3\alpha}$ . The four windows have

$$J'_{3\alpha} = 0, J_{3\alpha}/10, J_{3\alpha}, 10J_{3\alpha} \quad (2.33)$$

respectively. We chose these windows to make sure that the information contained within one window was sufficiently different from the next.

Once again, our power spectrum was generated by the methods described in Eisenstein and Hu (1999) with  $h = 0.65$ ,  $\Lambda = 0.7$ ,  $\Omega_b = 0.3$ ,  $n = 1$ , and  $T = 2.73$  K. For the classical approximation, we again use a Fisher matrix calculated as described in Hamilton (1997). The plots of the information contained within the Fisher matrices in this section have been calculated in the same way as those of the previous section. Because the differences between the pair weight method and the classical method should be largest on scales larger than  $\approx 7.5h^{-1}$  Mpc and because the difference between one window and four should be apparent on all scales, we have only calculated a  $31 \times 31$  matrix of separations.

In Figure 2.6 is the information contained within all four sets of windows (the solid line) and the information contained within each set of windows individually. It is clear that adding windows does, in fact, allow for the extraction more information. However, notice that the information contained within the best set of windows (this happens to be the third window,

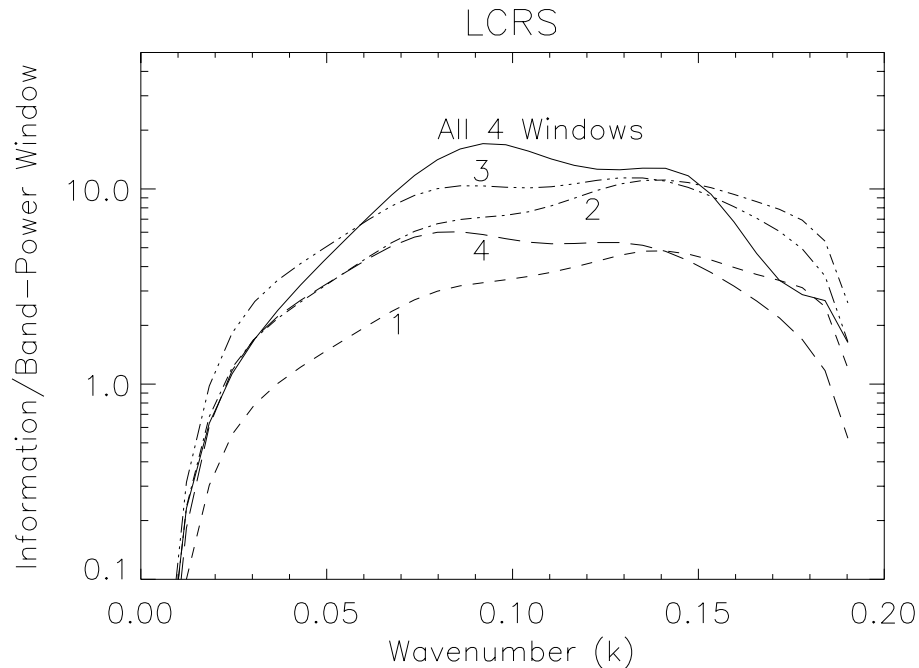


Figure 2.6: Information contained within decorrelated parameters for the pair weight method. The  $k$  index gives the wavenumber of the location of the peak of the parameter. The solid line contains information from all four sets of windows. The four other lines contain only the information from one of the sets of windows.

with  $\hat{J}_{3\alpha} = J_{3\alpha}$ ) contains no less than half of the information on all scales (as compared to the complete set of four windows). This Figure shows that this window actually contains more information on the smaller scales. This window does not actually contain more information but rather the errors in the method of calculation yield slightly inaccurate results. Also note that the fourth window ( $\hat{J}_{3\alpha} = 10J_{3\alpha}$ ) appears to contain the second most information of all the sets of windows. However, after eliminating this set of windows from the calculation, the remaining three sets of windows contain almost exactly the same amount of information as do all four sets together. The information extracted by window four has already been extracted by the first three windows (primarily window three).

Figure 2.7 shows the information contained within the classical approximation as compared to the pair weight method. This plot shows that on the very largest scales (small  $k$ ) the

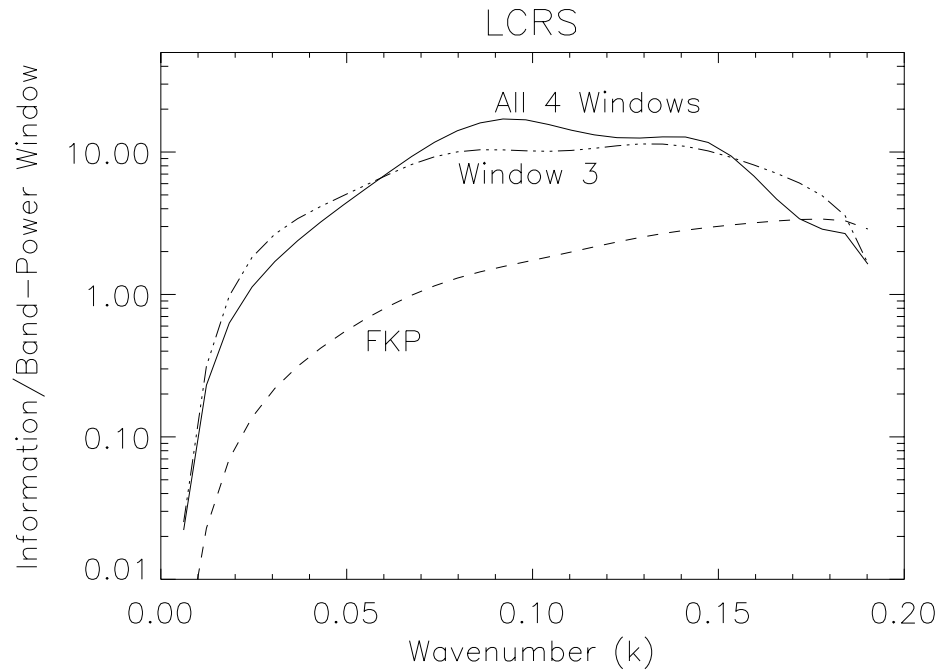


Figure 2.7: Information contained within decorrelated parameters. The  $k$  index gives the wavenumber of the location of the peak of the parameter. The solid line contains information from all four sets of windows for the pair weight method. The dot-dashed line is the information contained within the best single window for the pair weight method. The dashed line shows the information contained within the classical method.

pair weight method contains several times as much information as does the classical method. On smaller scales, the difference becomes less. On the smallest scales the classical method does retain more information than does this pair weight method. There are two causes for this. One cause is aliasing in the FFT used to go from real space to Fourier space. We found (in Chapter 4) that the latter portion of the matrices are contaminated to some degree by aliasing. Second, when the calculation is extended to include smaller scales, it does a better job of retaining the information on larger scales as well. This is particularly important for elements near the edge of the matrix. To obtain the best measurements of the power spectrum on these scales one would have to extend the Fisher matrix calculation to smaller scales.

The decorrelated band-powers, for which the information is calculated, are shown in Figure 2.8. Once again the parameters are narrow about the peak. Also, notice that the classical



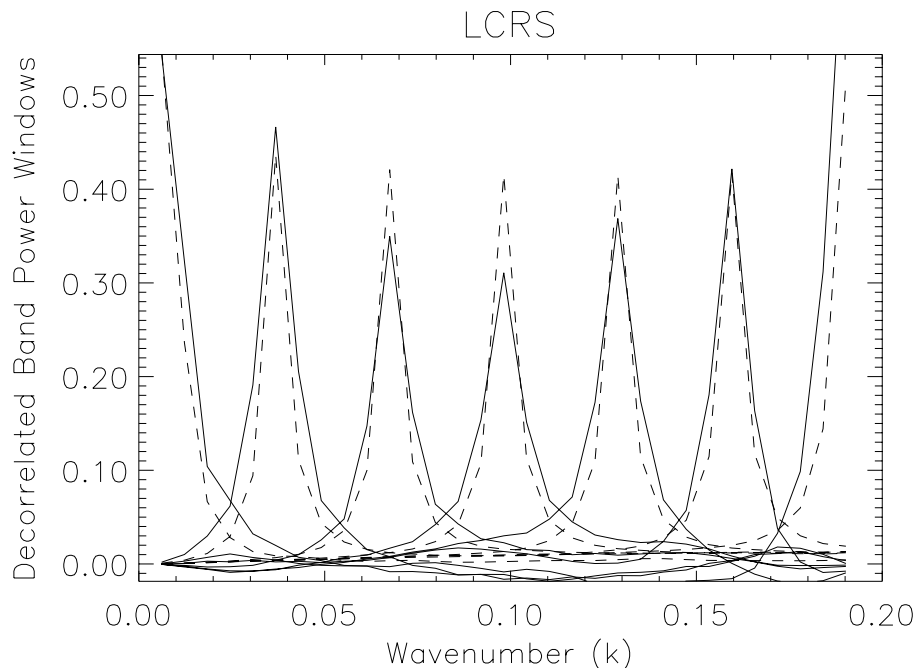


Figure 2.8: Normalized rows of the square-root of the Fisher matrix. Solid lines are for the pair weight method and dashed are for the classical method.

method parameters are very similar to the pair weight parameters. The main difference between the two is that the errors in the calculation of the pair weight Fisher matrix once again causes regions where the square root of the Fisher matrix is negative. This means that the errors in the calculation of the Fisher matrix is causing us to measure the parameters with weights which are not the best choice and to measure them with less accuracy than would be possible if the integrals of the matrix elements could be done exactly. Notice that, in the case for LCRS, the errors in the pair weight calculation mainly affect the smallest and largest scales measured while not greatly affecting the intermediate scales.

For the LCRS catalog, the pair weight method provides a more accurate measurement for the largest scales than does the classical method. Unfortunately, with only our  $31 \times 31$  matrix of data points we are unable to say for sure where the classical method becomes comparable to the pair weight method.

## 2.4 Discussion and Conclusions

We introduced a pair weight compression technique for measuring the power spectra from galaxy catalogs. This method requires the calculation of a matrix of ten-dimensional integrals. Because galaxy surveys tend to be less than all-sky, we found that the integrals cannot be performed analytically or by traditional numerical techniques. We discussed a Monte Carlo technique for computing these integrals. The advantage of the Monte Carlo technique is that the integrals can be performed even when the catalog's selection function is very complex (as is the case for LCRS). We then applied the technique to the IRAS 1.2 Jy redshift catalog and the LCRS.

We found that for the IRAS 1.2 Jy survey, the classical method would produce (slightly) smaller error bars than would the pair-weight method, if we use just one window per separation. Both methods are able to measure narrowly peaked decorrelated band-powers. Although the pair weight method works for a nearly all sky survey like IRAS, it is clear that the simplicity of the classical method makes it the preferred method of calculation.

For LCRS, we found that even using only one window per separation, the pair weight method produces error bars considerably smaller (as much as a factor of three smaller) than those of the classical method. When we used four (or the three best), windows per separation the information retained was even greater. Due to the computational expense of calculating the additional windows, it is not clear whether one should use one or more window per separation. Clearly, if computational expense is not an issue, one should use many windows.

Throughout the chapter we have assumed that the overdensities are linear and that redshift distortions are negligible. We know that both of these assumptions are not accurate in our universe. Under these simplified circumstances, we saw that the pair weight method is a viable method. For surveys with complicated selection functions, the pair weight method might well be the preferred method. It will be interesting to see whether the pair weight method remains a quality method when we relax the assumptions of linearity and small distortions.

## Chapter 3

### **Application of the Pair Weight Method to the Las Campanas Redshift Survey with no Redshift Distortions**

The power spectrum of matter in the universe can be used to measure the relevant parameters of models of the universe (*e.g.*, Eisenstein et al., 1999). Because galaxies are easily visible, it is common to use galaxies as a tracer for the underlying density field. It has been shown (Scherrer and Weinberg, 1998; Cole et al., 1988) that if galaxy formation is a local phenomenon, then the galaxy power spectrum on large scales is different from the mass power spectrum by a multiplicative constant (at least in the linear regime; Mann et al., 1998). This constant is the bias factor. Because of the expected similarities between the galaxy power spectrum and the mass power spectrum, it is not surprising that power spectra have been taken for all of the larger galaxy surveys.

Despite the wealth of measurements, the power spectrum on large scales is not yet well determined. Part of the problem lies in the fact that the redshift surveys have different selection criteria. For example, it is commonly thought (Peacock, 1997) that galaxies selected in the infrared may have a bias factor that is lower than those selected in the optical. Another problem is that the amount of data in each catalog still produces large (tens of percent) errors in the measurement of the power spectrum at any given wavenumber.

The Las Campanas Redshift Survey is one of the larger publicly available redshift surveys. The observational procedure maximized the number of galaxy redshifts given the telescope time available to the team. Although this maximization of the telescope time means

that the analysis will yield the highest possible number of galaxy redshifts, optimizing for the number of galaxy redshifts makes the analysis of the survey more difficult. Each of the 327 observing fields has a different covering factor. That is, one field may have redshifts for 85 percent of the galaxies where the next field over has redshifts for 65 percent of the galaxies. In addition to this, the catalog is 6 slices each having a width of  $1.5^\circ$ . This means that it is similar to 6 two-dimensional slices. This is a problem for analyses that rely on the measured wavelength to fit comfortably within the catalog (in any direction). These characteristics make the Las Campanas Redshift Survey difficult for traditional methods, yet an ideal catalog with which to test the pair weight compression method.

In chapter 2, we described a method for extracting the power spectrum from a redshift galaxy catalog. The pair weight compression method assigns weights to each galaxy pair within the catalog based on the selection function (the number of galaxies within the catalog divided by the true number of galaxies) of the catalog and the parameters that one wishes to measure. The method requires that the selection function of galaxies to be known for each position in the catalog. It does not require that the selection function have any particular properties (like being an angular selection function times a radial selection function). This means that the pair weight compression method should assign weights to the individual galaxy pairs to maximize the information extracted from the galaxy survey (given the limitation of the number of window functions available in the calculation). The LCRS certainly meets all requirements for the pair weight compression method.

In section 3.1 we discuss the modifications to the catalog that were made to make the analysis as accurate as possible. Section 3.2 describes the mathematics necessary to extract the power spectrum from the catalog using the pair weight method. In section 3.3 we extract the power spectrum from the LCRS. We then compare the results to power spectra from other works. We conclude in section 3.4

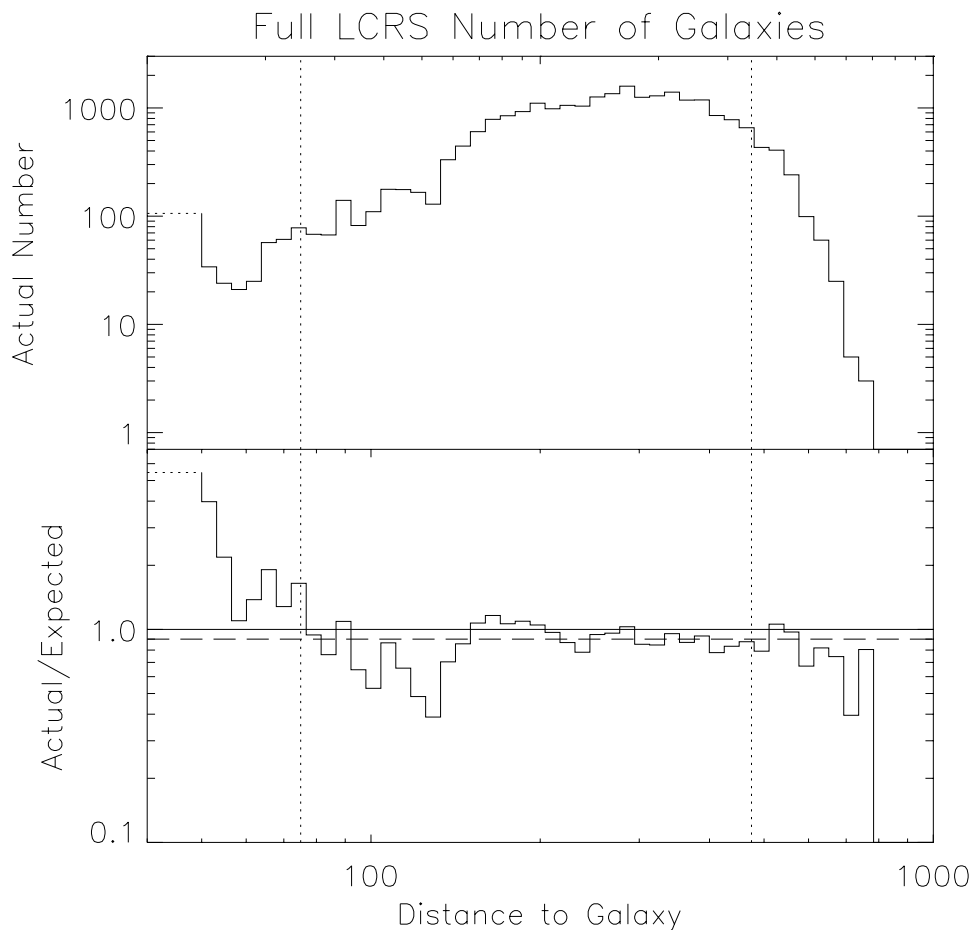


Figure 3.1: Top panel is the number of galaxies per logarithmic bin in the full Las Campanas Redshift Survey. The bottom panel shows the ratio of the number of galaxies in the full LCRS to the expected number based on the integrated selection function. The dashed line shows the ratio based on the total number of galaxies in the cut range divided by the number of galaxies found by integrating the selection function over the cut range.

### 3.1 Preparation of Catalog

The Las Campanas Redshift Survey (hereafter LCRS) of Shectman et al. (1996) contains 23720 galaxy redshifts stretching from  $\sim 10h^{-1}$  Mpc to  $\sim 800h^{-1}$  Mpc with the range of  $75h^{-1}$  Mpc to  $475h^{-1}$  Mpc being reasonably well sampled. Lin et al. (1996a) calculate the selection function of this well sampled region. The number of galaxies per logarithmic bin is shown in the top panel of Figure 3.1. In the lower panel is the ratio of the number of galaxies per

logarithmic bin to the expected number (using the selection function code generously provided by Huan Lin). This shows that the number of galaxies is well fit by the selection function in the region beyond  $75h^{-1}$  Mpc.

In order to perform the calculations, we take two catalogs from the LCRS. The first is all 23720 galaxies, while the second is the 21996 galaxies between  $75h^{-1}$  Mpc and  $475h^{-1}$  Mpc. The lower cutoff is due to the fact that the selection function does not adequately describe the catalog below  $75h^{-1}$  Mpc. The upper cut is because the catalog is rather sparse beyond this cut. This means that the statistics are not all that good. This means that the region beyond  $475h^{-1}$  Mpc should not provide too much additional information.

### 3.2 Calculation of Power Spectrum

In chapter 2, we described how to extract the power spectrum from a galaxy survey using the pair weight compression method. The final result was that the estimate of  $\xi_\beta$  where  $\beta$  is the index of the separation is given by:

$$\hat{\xi}_\beta = M_{\beta\alpha}^{-1} B_{\alpha ij} W_{aij} K_{ab}^{-1} W_{bkl} \delta_k \delta_l. \quad (3.1)$$

Chapter 2 describes the Monte Carlo Calculation of  $M$ ,  $K$ , and the product  $BW$  in real space. Notice that equation 3.1 does not indicate the space in which the calculation is performed. Because all of the data, selection functions, etc. are in real space, perform all calculations in real space until the final estimate. Then, use an FFT to obtain a Fourier space representation.

#### 3.2.1 Calculation of $W_{bkl} \delta_k \delta_l$

The only computation involved in Equation 3.1 not discussed in Chapter 2 is the calculation of  $W_{bkl} \delta_k \delta_l$ . This is the only place in which the actual galaxy catalog data is introduced into the calculation. For clarity, first look at the full integral:

$$W_{bij} \delta_i \delta_j = \int \frac{\bar{n}_i \bar{n}_j \delta(r_\alpha - |\mathbf{r}_i - \mathbf{r}_j|)}{(1 + 4\pi J_{3\alpha} \bar{n}_i)(1 + 4\pi J_{3\alpha} \bar{n}_j)} \frac{n_i - \bar{n}_i}{\bar{n}_i} \frac{n_j - \bar{n}_j}{\bar{n}_j} d^3 r_i d^3 r_j \quad (3.2)$$

where  $n_i = n(r_i)$ . For simplicity of calculation this integral splits into three separate integrals:

$$W\delta\delta = W_{ss} - (W_{ds} + W_{sd}) + W_{dd}. \quad (3.3)$$

Where the subscript  $d$  stands for data and the subscript  $s$  stands for smooth.

$$W_{ss} = \int \frac{\bar{n}_i \bar{n}_j \delta(r_\alpha - |\mathbf{r}_i - \mathbf{r}_j|)}{(1 + 4\pi J_{3\alpha} \bar{n}_i)(1 + 4\pi J_{3\alpha} \bar{n}_j)} d^3 r_i d^3 r_j. \quad (3.4)$$

This integral is very similar to the integral for  $WB$ . The only difference is that  $WB$  has an additional factor of  $\delta(r_\beta - |\mathbf{r}_i - \mathbf{r}_j|)$ . This means that the calculation for  $WB$  can be used for the calculation of  $W_{ss}$  and the only modification is in the translation from continuous representation to discretized representation. When discretized,  $WB$  is a diagonal matrix whereas  $W_{ss}$  is a vector.

Because  $i$  and  $j$  are dummy indices

$$W_{sd} + W_{ds} = 2W_{ds} = 2 \int \frac{n_i \bar{n}_j \delta(r_\alpha - |\mathbf{r}_i - \mathbf{r}_j|)}{(1 + 4\pi J_{3\alpha} \bar{n}_i)(1 + 4\pi J_{3\alpha} \bar{n}_j)} d^3 r_i d^3 r_j. \quad (3.5)$$

The value  $n_i$  is not a smoothly distributed value but rather a series of delta functions at the locations of the galaxies. This means that the integral over  $i$  becomes a sum over galaxies:

$$W_{sd} + W_{ds} = 2 \sum_{\text{galaxies}(i)} \int \frac{\bar{n}_j \delta(r_\alpha - |\mathbf{r}_i - \mathbf{r}_j|)}{(1 + 4\pi J_{3\alpha} \bar{n}_i)(1 + 4\pi J_{3\alpha} \bar{n}_j)} d^3 r_j. \quad (3.6)$$

So for each galaxy, integrate over the shell centered on the galaxy's position with radius  $r_\alpha$ . This set of 2-dimensional integrals can be performed by standard numerical techniques. It is easiest to perform this integral as the sum of integrals on the sub-shells as they intersect each of the LCRS observing areas. This is the set of integrals which takes the bulk of the computing time when calculating the term  $W\delta\delta$ .

The final term is

$$W_{dd} = \int \frac{n_i n_j \delta(r_\alpha - |\mathbf{r}_i - \mathbf{r}_j|)}{(1 + 4\pi J_{3\alpha} \bar{n}_i)(1 + 4\pi J_{3\alpha} \bar{n}_j)} d^3 r_i d^3 r_j. \quad (3.7)$$

Once again the values  $n_i$  and  $n_j$  are a series of delta functions at the locations of the galaxies.

So this integral should actually be written

$$\sum_{\text{galaxies}(i)} \sum_{\text{galaxies}(j \neq i)} \frac{\delta(r_\alpha - |\mathbf{r}_i - \mathbf{r}_j|)}{(1 + 4\pi J_{3\alpha} \bar{n}_i)(1 + 4\pi J_{3\alpha} \bar{n}_j)} \quad (3.8)$$

where the delta function is 1 if  $r_\alpha - |\mathbf{r}_i - \mathbf{r}_j| = 0$  and 0 otherwise. Because there are essentially no galaxy pairs with exactly separation  $r_\alpha$ , the calculation must be relaxed somewhat. In addition to this, the full calculation is using sums over discrete values ( $\alpha$ ) to approximate integrals. This means that the galaxy pair should contribute to the integral even if the separation is not equal to the arbitrary choice of  $\alpha$ . To accommodate this, simply apply the contributions to the two values of  $r_\alpha$  which bracketed the actual separation. The fractions ( $f_{\text{lower}}$  and  $f_{\text{upper}}$ ) apportioned to each integral are reminiscent of linear interpolation:

$$f_{\text{lower}} = \frac{r_{\text{upper}} - r_{\text{sep}}}{r_{\text{upper}} - r_{\text{lower}}}, f_{\text{upper}} = \frac{r_{\text{sep}} - r_{\text{lower}}}{r_{\text{upper}} - r_{\text{lower}}}. \quad (3.9)$$

This sum can be done exactly as written with a reasonable amount of computing time. In fact, the sums can be replaced

$$\sum_{\text{galaxies}(i)} \sum_{\text{galaxies}(j \neq i)} \rightarrow 2 \sum_{\text{galaxies}(i)} \sum_{\text{galaxies}(j > i)} \quad (3.10)$$

to save half of the computing time.

Because the terms  $W_{ss}$  and  $W_{sd} + W_{ds}$  are, in fact, calculated as true delta functions in  $\alpha$ , it is necessary to reduce the value of  $W_{dd}$  so that it will contain the same volume. The delta functions mean that the volume is actually a spherical shell (with no width) with area  $4\pi r_\alpha^2$ . So multiply by  $r_\alpha^2$  and divide by

$$\int_{r_{\alpha-1}}^{r_\alpha} \left(1 - \frac{r_\alpha - r}{r_\alpha - r_{\alpha-1}}\right) r^2 dr + \int_{r_\alpha}^{r_{\alpha+1}} \left(1 - \frac{r - r_\alpha}{r_{\alpha+1} - r_\alpha}\right) r^2 dr. \quad (3.11)$$

This places all of the terms in  $W\delta\delta$  on the same footing, so they can be added together. With the calculations from chapter 2, there is now a complete description of the calculation of  $\hat{\xi}_\beta$  in real space.

### 3.2.2 Measuring the Power Spectrum

The next step is to put it all together to obtain a measurement of the power spectrum with error bars. Equation 3.1 gives the measurement of the power spectrum at wavelength  $k_\alpha$



marginalized over all other possible parameters. Since  $M$  is just the inverse of the covariance matrix for the parameters  $\xi_\alpha$ , the error is given by  $\sqrt{M_{\alpha\alpha}^{-1}}$ . Although this gives the proper measurement at a particular wavelength (a delta function in width) this is not usually the quantity of interest. In fact, this quantity is anti-correlated. This means that neighboring points have covariances that are negative. Because neighboring points are expected to have similar values it makes sense to use a correlated measurement rather than an anti-correlated measurement. Smoothing over an appropriate window produces smaller error bars at the expense of measuring a region around  $k_\alpha$  rather than measuring only  $k_\alpha$  itself.

The first smoothing of interest produces what we will call the correlated spectrum. This is obtained by multiplying equation 3.1 by  $M_{\gamma\beta} / \sum_\beta M_{\gamma\beta}$ :

$$\hat{\xi}_\alpha^{\text{correlated}} = \frac{1}{\sum_\beta M_{\gamma\beta}} B_{\gamma ij} W_{aij} K_{ab}^{-1} W_{bkl} \delta_k \delta_l. \quad (3.12)$$

We must divide by the sum of the row of the Fisher matrix to make sure that the measurement is unbiased. Notice that this smoothes the measurement using the Fisher matrix as the smoothing function. This will correlate the data over a reasonably small width. This is a nice representation because it is the measurement that emerges most naturally from the data. Also, this measurement is nice and smooth. This is because smoothing reduces the error bars because, instead of using the measurement at just the point of interest, it averages over the surrounding points as well. In this case, the error bars are given by  $\sqrt{(M_{\gamma\gamma})} / \sum_\beta M_{\gamma\beta}$ . If the Fisher matrix is reasonably sharply peaked, then this representation will yield a valid, unbiased estimate of the power spectrum. In fact, most analyses use a similar weighting of the data to achieve their result. It means, however, that the measurements are not independent of one another.

The second smoothing function that produces an interesting result is obtained by using the square-root of the Fisher matrix as the smoothing function. There are many possible versions of the square-root. In this paper we diagonalize the Fisher matrix and take the square root of the Eigenvalues. This measurement is interesting because

$$\langle \Delta M^{\frac{1}{2}} \xi \Delta M^{\frac{1}{2}} \xi \rangle = 1 \quad (3.13)$$

That is to say that, smoothing with the square root of the Fisher matrix results in measurements that are uncorrelated. For this reason, we call the result decorrelated.

$$\hat{\xi}_\gamma^{\text{decorrelated}} = \frac{1}{\sum_\beta M_{\gamma\beta}^{1/2}} M_{\gamma\alpha}^{-1/2} B_{\alpha ij} W_{a ij} K_{ab}^{-1} W_{bkl} \delta_k \delta_l. \quad (3.14)$$

The errors on the measurements  $\hat{\xi}_\gamma^{\text{decorrelated}}$  are just  $1/\sum_\beta M_{\gamma\beta}^{1/2}$ .

### 3.2.3 Correcting for Discrete FFT Errors

The process of using discrete matrices rather than the true continuous values leads to one other problem. The calculation of the estimates of the power spectrum is actually a calculation of the Fourier transform of the estimate of the correlation function. Taking the Fourier transform of the discretized version of the prior value of the correlation function should yield the discretized version of the prior power spectrum. However, performing this calculation actually yields a power spectrum which is lower (and at the large  $k$  end much lower) than the prior value of the power spectrum. This is because the discrete FFT of the prior power spectrum is necessarily bumpy. On the other hand, the prior correlation function (obtained using a continuous Fourier transform) is smooth. The power contained within the extra bumps are lost in the calculation. As the binning becomes finer (as more points are added) the discrepancy becomes smaller.

Fortunately, the prior value of the correlation function tells us exactly how to correct for this error. If the measured correlation function is exactly the prior correlation function, then the measured power spectrum should be exactly the prior power spectrum. This means that in order to obtain the corrected estimate of the power spectrum, take the uncorrected value of the power spectrum (before doing any smoothing) and multiply by the ratio of the prior power spectrum to the FFT of the prior discretized correlation function.

Figure 3.2 shows the results of the correction in both the 31 by 31 matrix and the 63 by 63 matrix. The dashed lines show the estimate for the power spectrum without the correction. The solid lines show the same results with the appropriate corrections applied. In both cases

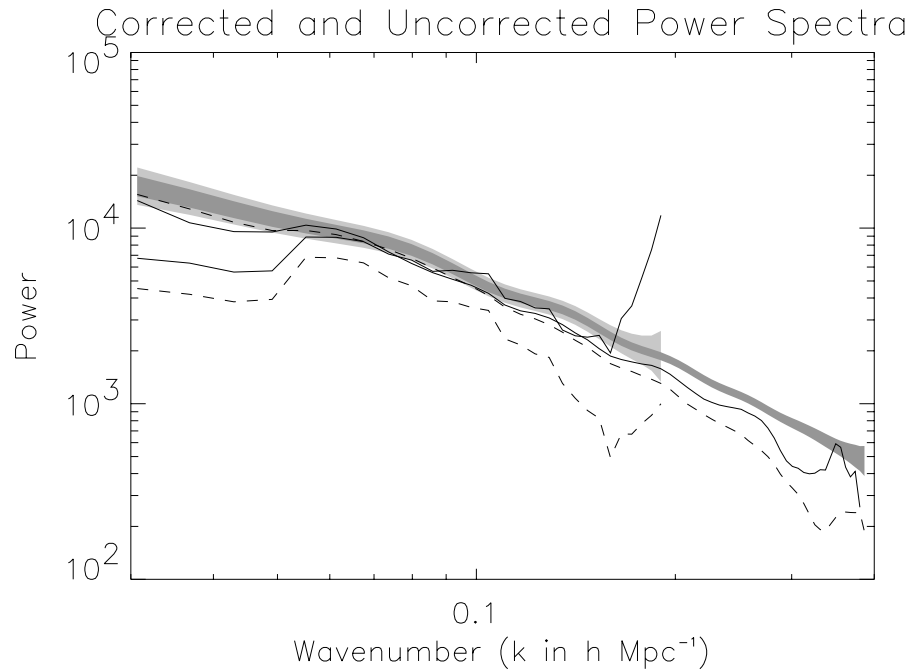


Figure 3.2: Solid lines show the correlated versions of the corrected power spectra. Dotted lines show the same versions for the uncorrected power spectra. Shaded regions show the expected ( $1-\sigma$ ) errors.

the correction is a factor of a few at the largest  $k$  probed. Notice that the agreement between the measured power spectra becomes much better in the corrected version than it was in the uncorrected version.

### 3.3 Power Spectra

#### 3.3.1 Power Spectrum of the Full LCRS

Figure 3.3 shows the correlated and decorrelated estimates of the redshift space power spectrum for the full LCRS catalog. On small scales (large  $k$ ) the measured power is lower than that of the prior. On large scales, the measured power is larger than that of the prior. The correlated version (the solid line) appears to have variations that are within the expected errors. However, the decorrelated measurement has errors which are larger than expected (i.e.,

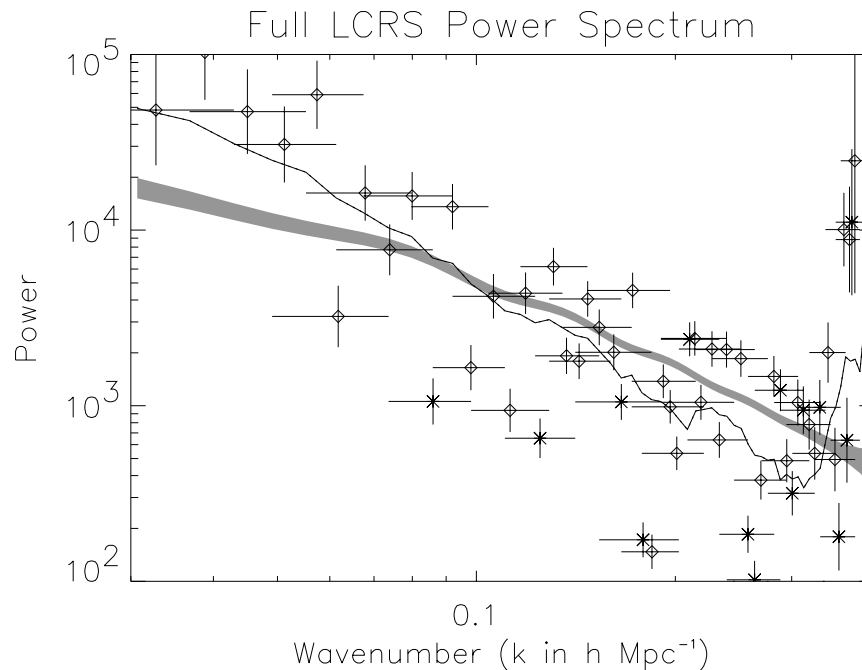


Figure 3.3: Power spectrum of all 23720 LCRS galaxies. The shaded region shows the expected deviations from the prior power spectrum for the correlated measurements. Diamonds show the measured decorrelated power for a bin with horizontal and vertical error bars. The horizontal error bars show the full width at 1/4-maximum. The position of the points is given by the center of the full width at 1/4-maximum region. Stars also show the measured power for a bin, however the measured power is negative. The solid line is the positive portion of the correlated power spectrum measurement. Notice that the scatter between points is much larger than the expected errors. This is due to the galaxies where the selection function is very small.

neighboring points are separated by much more than the error bars). The scatter is due to the large contributions from the few galaxies in sparsely sampled locations. Also, the power on the largest scales is likely to have been enhanced by these low selection function galaxies.

At locations where  $k > 0.3h/\text{Mpc}^{-1}$  the estimates take a large upturn. Hamilton (2000) found that when taking the fast Fourier transform, the outer portions of a matrix are not always faithfully represented. It appears that (Figure 3.2) the last few points of the measurement are systematically too high. This means that when looking at the results, one should take the outer portions of the matrix less seriously than the middle.

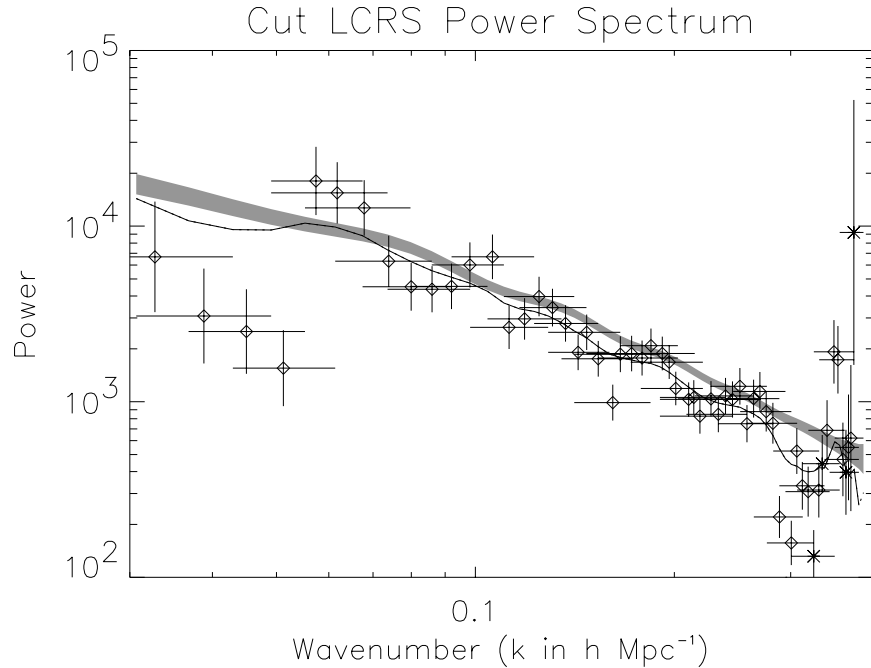


Figure 3.4: Power spectrum of the cut catalog's 21996 LCRS galaxies. The lines and symbols are as described in Figure 3.3. Notice that the error bars are now consistent with the scatter of the points.

### 3.3.2 Power Spectrum of the Cut LCRS

By cutting the LCRS at  $75h^{-1}$  Mpc and  $475h^{-1}$  Mpc, only the region which is well fit (this eliminates the small redshift galaxies) by the Lin et al. (1996a) selection function and which is well sampled (this eliminates the most distant region) is included in the measurement. By using only the best region of the catalog, systematic errors (from a mismeasured selection function where the selection function is too low to be included in the fit to the selection function) in the measurements will be reduced. The results of this analysis are shown in Figure 3.4.

The correlated power spectrum is positive at all wavelengths. The errors in the measurement of the power spectrum using the cut catalog appear to agree with the expected level of errors. It is interesting to note that the measured power spectrum agrees rather well with the prior power spectrum (particularly in shape). The region with  $k \gtrsim .3 h \text{ Mpc}^{-1}$  is likely to be

inaccurate. This means that both of the negative values measured in the decorrelated power spectrum occur where it is likely to be affected by aliasing.

The prior value of the power spectrum cannot be completely ruled out by our measurement. First of all, on the largest scales (small  $k$ ), redshift distortions will tend to add power to the measurement. On the very smallest scales (large  $k$ ), redshift distortions will tend to reduce the measured power. Although the prior cannot be completely ruled out, it does appear that the measured power spectrum is steeper than the prior.

There is some evidence that the power spectrum is turning over at about  $.06 \text{ h Mpc}^{-1}$ . The region with smaller values of  $k$  is likely to be contaminated by the FFT aliasing power from larger  $k$  into the bins with smaller  $k$ . In fact, the first few measured points are not plotted due to severe aliasing problems. Unfortunately, the measurement is too insensitive to make any definitive statement about the turn-over of the power spectrum.

### 3.3.3 IRAS 1.2 Jy Power Spectrum

Figure 3.5 shows the measured power spectrum for the IRAS 1.2 Jy survey. Once again, the line is the pair weight measurement. The line is solid where positive. The diamonds show the canonical measurement by Fisher et al. (1993) with error bars. The stars show the power spectrum measured by the combined 1.2 Jy and QDOT surveys as measured by Tadros and Efstathiou (1995). The pair weight measurement appears to be noisier than expected. This may be due to errors in the calculation of the Fisher matrix elements. It is clear, however, that the pair weight measurement is similar to the measurements by Fisher et al. (1993) and Tadros and Efstathiou (1995).

### 3.3.4 Tests of the Pair Weight Method

Now that we have a power spectrum measured by the pair weight method, we would like to know the limitations of the method. The pair weight method can be tested in two additional ways. The first is to work with a data set with a known power spectrum. The second is to

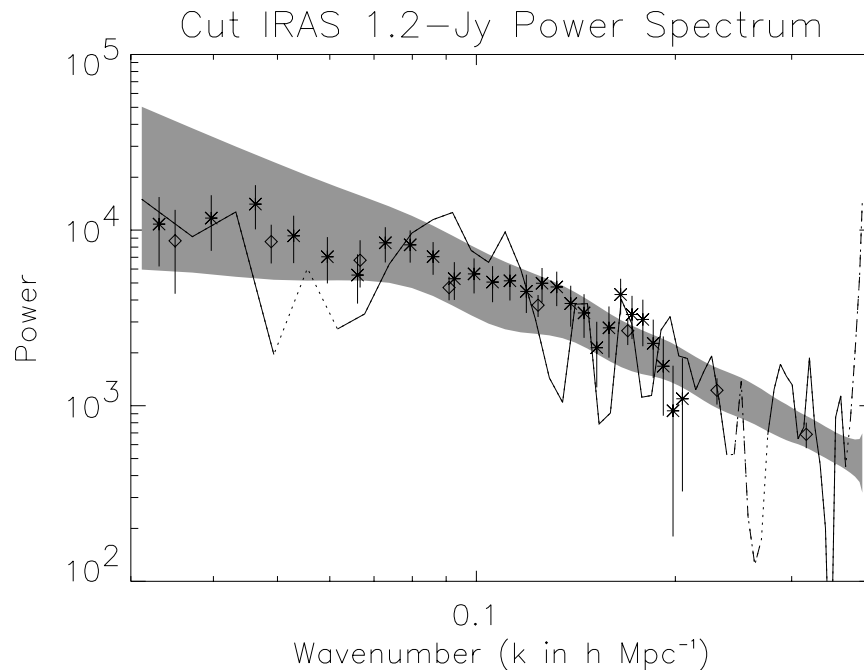


Figure 3.5: Power spectrum for IRAS 1.2 Jy catalog. The line is the pair weight measurement. The line is solid where positive, dotted where switching signs and dot-dashed where negative. Also shown are measurements by Fisher et al. (1993) (diamonds) and Tadros and Efstathiou (1995) (stars). Tadros and Efstathiou (1995) results report the power spectrum measured from the combined 1.2 Jy and QDOT surveys.

work with a prior power spectrum which is not expected to be correct to see if the selected prior affects the outcome of the measurement.

### Mock Catalog with Zero Power

In order to test whether the pair weight method is able to extract the correct power spectrum from data with a known power spectrum, we tested it against a mock galaxy catalog. The mock galaxy catalog was created by randomly selecting galaxies based on the Huan Lin selection function. This means that the power spectrum (for nonzero values of  $k$ ) should be zero. The results of the analysis using this mock catalog are shown in Figure 3.6. Clearly, the power measured is much smaller than the prior power spectrum. In fact the measured value in the correlated power spectrum is approximately equal to the error bar supplied on the prior. This is

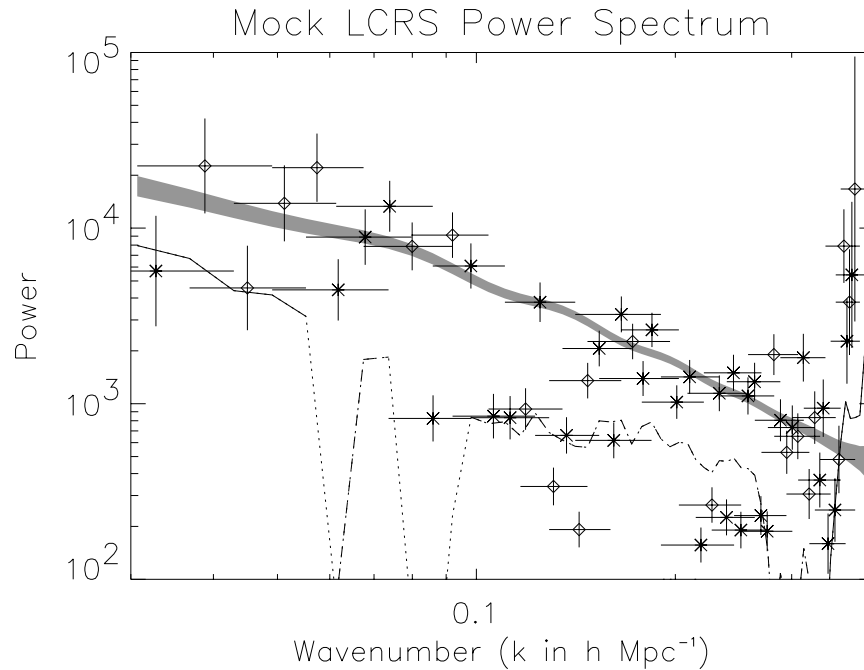


Figure 3.6: Power spectrum as measured from a mock catalog. The mock catalog has the correct selection function for the entire range of the LCRS. The input power for the catalog is zero. Here the error bars are constant in the logarithm assuming that the measured power is approximately equal to the prior power.

what one would expect. However, the measured power is systematically negative. This means that when the power in a data set is much lower than the prior value, there may be some sort of imprinting of the prior onto the data. This also says that the pair weight method, due to the errors associated with the calculation, may not measure zero all that well.

### Cut LCRS with Flat Prior

When working with a prior value of the power spectrum it is important to test whether the prior value biases the calculation in some way. In other words, it is important to know that the method is measuring the data rather than measuring the prior. To test this, we used a second prior with a very different shape than that of Eisenstein and Hu (1999). We used a prior that was uniformly  $10^4 (h^{-1} \text{Mpc})^3$ . This prior leads to the simplest analysis. This is because the



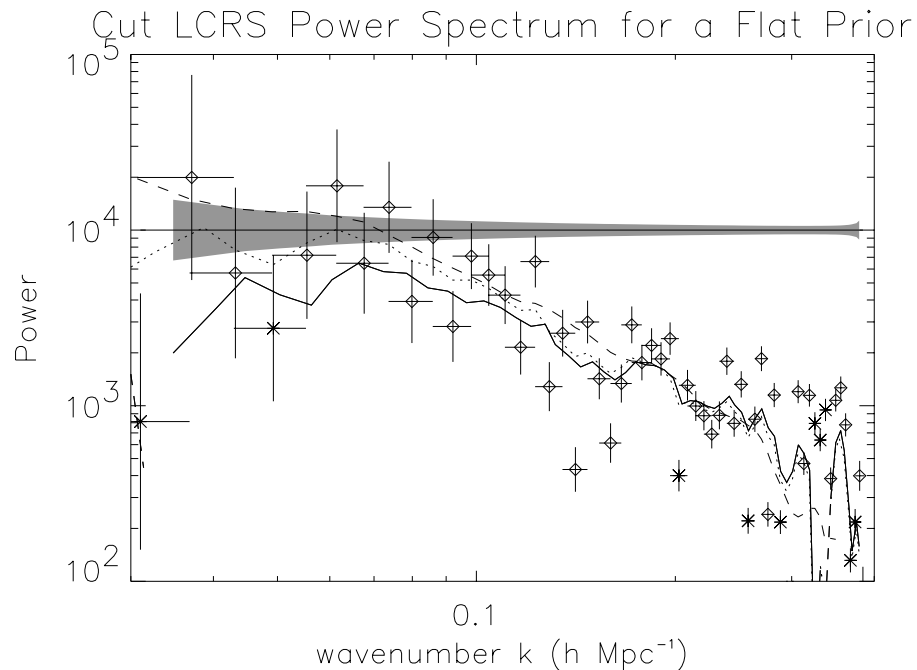


Figure 3.7: Power spectrum measured with the 21996 galaxies in the cut LCRS catalog. Again, the shaded region is the expected errors assuming the flat prior is correct. The data points (diamonds for positive and stars for negative) show the decorrelated results. The solid line is the correlated version smoothed with the typical smoothing of  $\xi_{\text{prior}} M \xi_{\text{prior}}$ . The dotted line is the correlated version using just the Fisher matrix  $M$  as the smoothing function. The dashed line is the result from the more realistic prior power spectrum.

Fourier transform of a flat power spectrum is a  $\delta$  function. This means that our matrices become diagonal. The other benefit of the flat prior is that even when discretized the Fourier transform of the real space correlation function is still a flat power spectrum. This means that there is no discretization correction. The results from the analysis of the cut LCRS using the flat prior are shown in Figure 3.7. The measurement is positive and is consistent with the expected errors. In addition to this, the measured power is similar to that measured with the more realistic prior. On the largest scales, the flat prior measures a smaller value for the power spectrum. On the smallest scales, the flat prior measures a slightly larger value for the power spectrum. This is probably due to power shifting from locations where the power divided by the prior is high to locations where the power divided by the prior is low. This is because the smoothing function

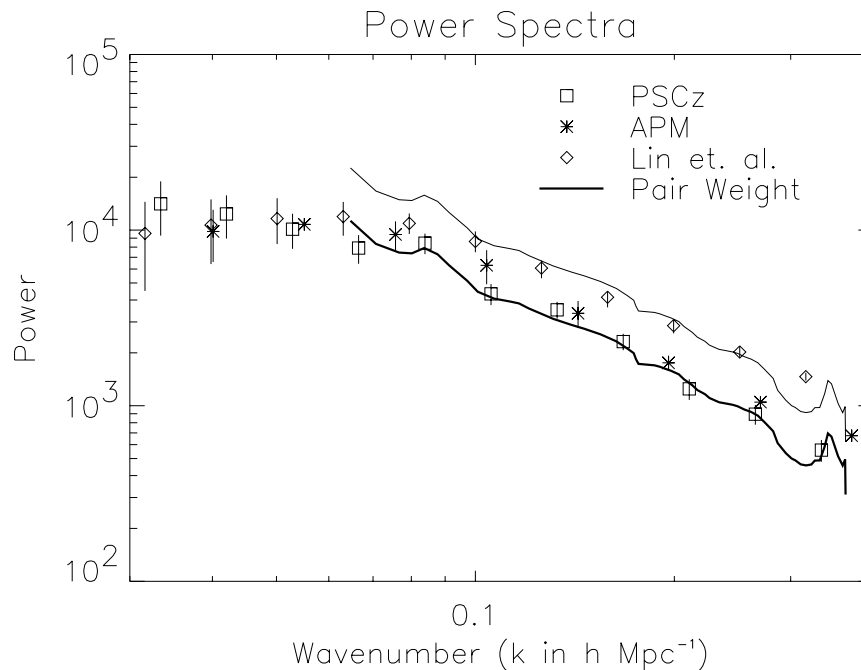


Figure 3.8: Power spectra for PSCz, APM and LCRS using different methods. The solid shows the power spectrum of LCRS measured with the pair weight method using the Eisenstein and Hu (1999) power spectrum as the prior. The diamonds and the associated error bars are the power in LCRS measured by Lin et al. (1996b). The squares and associated error bars are the power measured by Sutherland et al. (1999) from the PSCz. The stars and associated error bars are the power spectrum measured by Gaztañaga and Baugh (1998) for the APM.

( $\xi_{\text{prior}} M \xi_{\text{prior}}$  where  $\xi_{\text{prior}}$  is the discretized version of the prior power spectrum, and thus not constant) assumes that the measured value of the power spectrum is close to the prior value. This is clearly false with this outrageous prior power spectrum. To correct this the dotted line shows the correlated power spectrum using  $M$  as the smoothing function. This reduces the problem. In fact, the flat prior power spectrum produces a measured power spectrum that is very similar to that measured by the more realistic power spectrum. This is a good indication that the prior power spectrum is not biasing the results.

### 3.3.5 Comparison of Pair Weight Results to Previous Results

Figure 3.8 shows the comparison of the Lin et al. (1996b) and the correlated measurement using the pair weight method with the Eisenstein and Hu (1999) prior. Like Lin et al. (1996b) the pair weight method shows some evidence that the power spectrum turns over at small  $k$ . Notice that the pair weight method measures smaller power than does Lin et al. (1996b), on almost all scales. However, by multiplying the result from the pair-weight method by 2, we obtain the same result as Lin et al. (1996b) over our well-measured region. It is not at all clear why two analyses should obtain results that are exactly off by a factor of 2.

Also plotted is the measurement of the power in PSCz by Sutherland et al. (1999) (squares) and the results from the APM by Gaztañaga and Baugh (1998). The pair weight method measures very nearly the same power in LCRS as Sutherland et al. (1999) does for PSCz. This result is surprising due to the fact that it is commonly believed that IRAS selected galaxies will have a smaller bias than optically selected galaxies. However, the APM results are also from an optically selected survey, and they also appear to agree with the PSCz results. This may be an indication that the bias factor of optically selected galaxies is not all that different from that of IRAS selected galaxies.

## 3.4 Discussion and Conclusions

In this chapter we described how to apply the pair weight method to an actual data set. We then applied the pair weight method to the Las Campanas Redshift Survey using a selection function provided by Huan Lin and a  $\Lambda$ CDM prior power spectrum from Eisenstein and Hu (1999). The measured power spectrum of the entire LCRS was noisier than expected. Inclusion of galaxies where the selection function is small (several galaxies are in locations where the actual selection function is zero) caused errors in the power spectrum.

We then cut the LCRS to include only the region where the selection function fit the data well. After calculating the power spectrum, we found that the discrete nature of our measure-

ment introduces errors that can be corrected. The corrected power spectrum calculated with the cut LCRS appears to be good. The fluctuations in the measurement are consistent with the expected errors. This measured power spectrum is quite similar to the prior power spectrum. The measured power spectrum was slightly larger on large scales and slightly smaller on small scales. It is possible that these differences are due to redshift distortions.

Next, we performed the pair weight calculation with a flat prior power spectrum. The measured power for the flat prior case is consistent on scales where the calculation should be accurate with the measurement using the more realistic prior. On the large scales, however, the flat prior yields lower power. This is due to power shifting from large  $k$  to small  $k$ . When smoothed over a more appropriate window the flat prior result agrees very well with the result from the more reasonable prior.

Finally the measured power spectrum was compared to other works. The power spectrum measured from LCRS using the pair weight method was lower than that measured by Lin et al. (1996b) by a factor of 2. It is not clear what causes this discrepancy in the measurement. However, our results look nearly identical to the results of Sutherland et al. (1999) from their analysis of the PSCz survey. This is surprising due to the fact that the PSCz is an IRAS selected survey which is commonly thought to be less biased than an optical survey. However, the APM, another optical catalog, has very nearly the same power spectrum (Gaztañaga and Baugh, 1998) as measured in LCRS by the pair weight method and in the PSCz survey. It appears that optically selected galaxies and IRAS galaxies may be more similar than is commonly believed.

## Chapter 4

### Spherical Redshift Distortions

Redshift surveys give a 3-dimensional picture of the galaxy distribution in which the redshift of a galaxy is taken as a measure of its distance, through Hubble's law. In addition to redshift, due to cosmological expansion, galaxies have "peculiar" velocities caused by gravitational (and perhaps other) interactions. This causes distortions in the catalog's redshift patterns. Kaiser (1987) pointed out that the distortions have a simple form in Fourier space. He showed that a wave with amplitude  $\delta(\mathbf{k})$  has an amplification in redshift space of  $1 + \beta\mu^2$ . So

$$\delta^s(\mathbf{k}) = (1 + \beta\mu^2)\delta^r(k). \quad (4.1)$$

where the superscripts,  $s$  and  $r$ , refer to redshift and real space respectively (as they will throughout this chapter). Here  $\mu$  is the cosine of the angle between the line of sight and the wavevector  $\mathbf{k}$ .  $\beta = f/b$  is the ratio of the dimensionless linear growth rate ( $f$ ) to the linear light-to-mass bias ( $b$ ). In the linear regime, peculiar velocities ( $v$ ) are related to overdensities ( $\delta$ ) by the continuity equation  $\nabla \cdot \mathbf{v} + \beta\delta = 0$  (where  $H = 1$ ). In standard pressure-less Friedmann cosmologies, the growth rate is just a function of  $\Omega$ ,

$$f \approx \Omega_m^{\frac{4}{7}} + \frac{\Omega_\Lambda}{70} \left(1 + \frac{\Omega_m}{2}\right) \quad (4.2)$$

(Lahav et al., 1991) From equation 4.1 Kaiser concluded that since the Fourier transform of the correlation function, the power spectrum or  $P(k) = \langle \delta(k'), \delta^*(k + k') \rangle$  is also amplified. In the plane-parallel limit this becomes:

$$P^s(\mathbf{k}) = (1 + \beta\mu^2)^2 P^r(k). \quad (4.3)$$

Simply stated, the plane-parallel approximation is the assumption that the angular separation of the galaxies is small enough such that their line-of-sight displacements are effectively parallel. This assumption was applied by many subsequent papers (Lilje and Efstathiou 1989; McGill 1990; Loveday et al. 1992; Hamilton 1992, 1993; Gramann et al. 1994; Bromley 1994; Fry and Gaztañaga 1994; Fisher et al. 1994; Cole et al. 1994, 1995). In order for the plane parallel approximation to be close to accurate one must only take into consideration pairs of galaxies which are separated by a small angle on the sky. Unfortunately, in order to obtain good statistics (to include the largest number of galaxy pairs), one would like to have this angle be as large as possible. Also, linear theory of clustering is most accurate in regions with small overdensities: the largest regions possible. Therefore, methods involving the plane parallel approximation must strike a compromise between angles which are too large and separations which are not large enough. For example, Hamilton (1993) and Cole et al. (1994) both took this angle to be  $\leq 50$  deg. Cole et al. (1994, fig. 8) show from simulations that the plane parallel approximation causes errors of about 5 per cent if the analysis is cut off at  $50^\circ$ .

In reality the distortions are not parallel but rather radially directed towards (or away from) the observer. A proper treatment of the spherical distortions should be superior to the plane parallel treatment in several ways. First, the same catalog can provide better statistics for large wavelengths. Second, larger wavelengths (regions where the linear clustering is more valid) can be included in the analysis. Finally, the systematic errors introduced by the plane parallel approximation can be eliminated completely.

Progress on spherical distortions began with Fisher et al. (1994). They expanded the density field of the 1.2-Jy survey into spherical harmonics, windowing the density in the radial direction with Gaussian windows at several depths. Heavens and Taylor (1995) expanded the radial direction in a complete set of spherical waves. Both of these methods required a prior shape of the power spectrum. Ballinger et al. (1995) allowed the power spectrum to vary in six bins thereby eliminating the prior assumption of the shape for the power spectrum. Hamilton and Tegmark (2000b) introduced several improvements, allowing them to measure a full, high

resolution, power spectrum and also measure  $\beta$  as a function of  $k$ . These brute force techniques have been used to analyze most of the current available catalogs (*e.g.*, Tadros et al. 1999, Hamilton and Tegmark 2000). The drawback to brute force techniques are that they require the assumption of Gaussian fluctuations (not true in the nonlinear regime) and computing costs constrain the brute force techniques to work on only the largest scales.

Hamilton and Culhane (1996) began working on a method for measuring the power spectrum and  $\beta$  using a fully spherically symmetric analysis that had the flexibility to be used in the nonlinear regime. In particular we derived a spherical distortion operator which when applied to the undistorted correlation function gives the distorted correlation function:

$$\xi^s(r_{12}, r_1, r_2) = D_{\text{dist}} \xi^r(r_{12}) \quad (4.4)$$

where

$$D_{\text{dist}} = \left[ 1 + \beta \left( \frac{\partial^2}{\partial r_1^2} + \frac{\alpha(r_1)}{r_1} \frac{\partial}{\partial r_1} \right) \nabla_1^{-2} \right] \left[ 1 + \beta \left( \frac{\partial^2}{\partial r_2^2} + \frac{\alpha(r_2)}{r_2} \frac{\partial}{\partial r_2} \right) \nabla_2^{-2} \right] \quad (4.5)$$

Here,

$$\alpha(r) \equiv \frac{\partial \ln r^2 \bar{n}^s(r)}{\partial \ln r} \quad (4.6)$$

where  $\bar{n}^s(r)$  is the galaxy selection function in redshift space for the catalog in question. Looking at the distortion operator in equation 4.5 notice that the distortions depend only on  $\beta$  and the selection function of the catalog. One would like to use the properties of the distortion operator to find a way to measure  $\beta$  for each catalog. Hamilton and Culhane (1996) outlined a method for doing exactly that. However, in applying this method we found that some changes must be made in order to account for the properties of the individual catalog.

## 4.1 Spherical Shape Functions

In order to measure the real space power spectrum and  $\beta$  from a redshift survey it is necessary to have a model for the redshift distortions. In the Appendix, it is shown that, in

linear theory, the redshift space correlation function can be written as

$$\xi^s(r_{12}) = \sum_i g_i(\beta) B_i(r_{12}, r_1, r_2) \Xi_i(\xi(r_{12}), \check{\xi}(r_{12}), \bar{\xi}(r_{12}), \bar{\bar{\xi}}(r_{12})) \quad (4.7)$$

with

$$\begin{aligned} \check{\xi} &\equiv \frac{2}{r_{12}^2} \int_0^{r_{12}} \xi(r) r dr, \\ \bar{\xi} &\equiv \frac{3}{r_{12}^3} \int_0^{r_{12}} \xi(r) r^2 dr, \\ \bar{\bar{\xi}} &\equiv \frac{5}{r_{12}^5} \int_0^{r_{12}} \xi(r) r^4 dr. \end{aligned} \quad (4.8)$$

Each of the  $B$ 's is a function only of the shape of the triangle formed by  $r_1$ ,  $r_2$ , and  $r_{12}$ . Each of the measurable quantities

$$\check{\xi}_i(r_{12}) \equiv g_i(\beta) \Xi_i(r_{12}) \quad (4.9)$$

can be separated into a function of the separation (which can be transformed into  $P(k)$ ) times a function only of  $\beta$ . The choice of  $B$ 's and  $\check{\xi}$ 's is not unique, however the sum in Equation 4.7 must be correct.

#### 4.1.1 Importance of $\alpha$

Hamilton and Culhane (1996) assumed that  $\alpha$  could be treated as a constant, in particular, that

$$\int \alpha^i B_j(r_{12}, r_1, r_2) B_k(r_{12}, r_1, r_2) d^3 r_1 d^3 r_2 = \alpha^i M_{j,k} \quad (4.10)$$

for all  $i, j, k$ . In practice, the true  $\alpha$  does not satisfy this equation. So then the question becomes: what is the meaning of  $\alpha$  in the first place? We can see from the definition that  $\alpha$  (Equation 4.6) measures the distortions caused by selection function of the catalog. Since each catalog has a different  $\bar{n}(r)$  each catalog has a different  $\alpha(r)$ . Ideally, the distortions can be calculated exactly. In the above description of the method  $\alpha$  is not specifically mentioned. However, in order to treat the effect of  $\alpha$  properly it must be allowed to vary over  $r_1$  or  $r_2$ . In the method outlined



above the only functions allowed to vary over  $r_1$  and  $r_2$  are the shape functions ( $B_i(r_{12}, r_1, r_2)$ ). So in order to treat  $\alpha$  properly the shape functions must include  $\alpha$ . Notice that the inclusion of  $\alpha$  forces us to use a selection function with no radial jumps (except for when the selection function goes to zero).

#### 4.1.2 Shape Functions

For our calculations we have chosen the following shape functions:

$$B_0 \equiv 1,$$

$$B_1 \equiv \left[ \frac{\alpha(r_1)}{r_1} r_{12} \mu_1 + \frac{\alpha(r_2)}{r_2} r_{12} \mu_2 \right],$$

$$B_2 \equiv \frac{1}{2} [P_2(\mu_1) + P_2(\mu_2)],$$

$$B_3 = -\frac{2}{3} \mu_{12}^2 + \frac{2}{3} + \frac{1}{6} \frac{\alpha(r_1) \alpha(r_2) \mu_{12} r_{12}^2}{r_1 r_2} + \frac{1}{6} \frac{\alpha(r_1) \alpha(r_2) \mu_1 \mu_2 r_{12}^2}{r_1 r_2} - \frac{1}{2} \left[ \frac{\alpha(r_1) r_{12}}{r_1} \mu_1 \mu_2^2 + \frac{\alpha(r_2) r_{12}}{r_2} \mu_1^2 \mu_2 \right] - \frac{1}{3} \left[ \frac{\alpha(r_1) r_{12}}{r_1} \mu_{12} \mu_2 + \frac{\alpha(r_2) r_{12}}{r_2} \mu_1 \mu_{12} \right],$$

$$B_4 \equiv \frac{1}{8} (35 \mu_1^2 \mu_2^2 - 15 \mu_1^2 - 15 \mu_2^2 + 3),$$

$$B_5 = \frac{4}{5} \mu_{12}^2 - \frac{4}{5} - \frac{1}{30} \frac{\alpha(r_1) \alpha(r_2) \mu_{12} r_{12}^2}{r_1 r_2} - \frac{1}{10} \frac{\alpha(r_1) \alpha(r_2) \mu_1 \mu_2 r_{12}^2}{r_1 r_2} + \frac{1}{2} \left[ \frac{\alpha(r_1) r_{12}}{r_1} \mu_1 \mu_2^2 + \frac{\alpha(r_2) r_{12}}{r_2} \mu_1^2 \mu_2 \right] + \frac{1}{5} \left[ \frac{\alpha(r_1) r_{12}}{r_1} \mu_{12} \mu_2 + \frac{\alpha(r_2) r_{12}}{r_2} \mu_1 \mu_{12} \right] - \frac{1}{10} \left[ \frac{\alpha(r_1) r_{12}}{r_1} \mu_1 + \frac{\alpha(r_2) r_{12}}{r_2} \mu_2 \right],$$

$$B_6 = -\frac{\alpha(r_1) \alpha(r_2) \mu_{12} r_{12}^2}{r_1 r_2}. \quad (4.11)$$

This choice of shape functions leads to the following expressions for  $\check{\xi}_i$ :

$$\check{\xi}_0 \equiv \left(1 + \frac{2}{3}\beta + \frac{1}{5}\beta^2\right)\xi(r_{12}),$$

$$\check{\xi}_1 \equiv \left[\frac{1}{3}\beta + \frac{1}{6}\beta^2\right]\bar{\xi}(r_{12}), \quad (4.12)$$

$$\check{\xi}_2 \equiv \left(\frac{4}{3}\beta + \frac{4}{7}\beta^2\right) [\xi(r_{12}) - \bar{\xi}(r_{12})], \quad (4.13)$$

$$\check{\xi}_3 \equiv \beta^2\bar{\xi}(r_{12}), \quad (4.14)$$

$$\check{\xi}_4 \equiv \frac{8}{35}\beta^2 \left[ \xi(r_{12}) + \frac{5}{2}\bar{\xi}(r_{12}) - \frac{7}{2}\bar{\bar{\xi}}(r_{12}) \right], \quad (4.15)$$

$$\check{\xi}_5 \equiv \beta^2\bar{\bar{\xi}}(r_{12}), \quad (4.16)$$

$$\check{\xi}_6 \equiv \frac{\beta^2}{6}\check{\xi}(r_{12}). \quad (4.17)$$

## 4.2 Generalized Fourier Transform Windows

In Cole et al. (1995), Hamilton (1992) and Hamilton and Culhane (1996) it was stated that the functions  $\check{\xi}_l^s(r_{12})$  (where  $l = 0, 2, 4$ ) can be transformed into a function of  $\beta$  times  $P(k)$  using Bessel functions:

$$P(k) = 4\pi \int \Xi_{0,2,4}^s(r_{12}) j_{0,2,4}(kr_{12}) r_{12}^2 dr_{12}. \quad (4.18)$$

Unfortunately, in the spherically symmetric case a single Bessel function is not sufficient for some of the  $\Xi_l^s(r_{12})$  (where  $l \neq 0, 2, 4$ ). However, there are functions which will serve the same purpose.

The first thing to notice is that the window over  $k$  will not effect  $\beta$ . This means that the function for  $\check{\xi}^s$  needs to be a function of  $\beta$  times the function of the averages of  $\xi^r(r)$ . This

is the reason for setting  $\check{\xi}_i = g_i(\beta)\Xi_i$ . In other words, take the function  $g_i(\beta)$  outside of the windowing integral. Assume that

$$\begin{aligned} & 4\pi \int \Xi_i^s(r_{12})\check{j}_l(kr_{12})r_{12}^2 dr_{12} = \\ & 4\pi \int f_l(\xi(r_{12}), \check{\xi}(r_{12}), \bar{\xi}(r_{12}), \bar{\bar{\xi}}(r_{12}))\check{j}_l(kr_{12})r_{12}^2 dr_{12} = P(k) \end{aligned} \quad (4.19)$$

in order to find  $\check{j}_l(kr_{12})$ . In this equation  $\check{j}_l(kr_{12})$  is the generalized window. In the cases of ( $l = 0, 2, 4$ )  $\check{j}_l(kr_{12}) = 4\pi i^l j_l(kr_{12})$  where  $j_l(kr_{12})$  is the usual spherical Bessel function. Since this is not true for all  $i$  we want to find  $\check{j}_l(kr_{12})$  such that:

$$\begin{aligned} & 4\pi \int f_l(\xi(r_{12}), \check{\xi}(r_{12}), \bar{\xi}(r_{12}), \bar{\bar{\xi}}(r_{12}))\check{j}_l(kr_{12})r_{12}^2 dr_{12} = \\ & 4\pi \int \xi(r_{12})j_0(kr_{12})r_{12}^2 dr_{12}. \end{aligned} \quad (4.20)$$

To find the appropriate windows solve this equation for each value of  $l$ . First, choose functions  $f_l$  in such a way to make the calculations of the windows as simple as possible. The above separation leaves  $f_l, l \neq 0, 2, 4$  to be a function of only one of the averages of  $\xi^r(r_{12})$ .

$$\int_0^\infty \left[ \frac{n}{r_{12}^n} \int_0^{r_{12}} r^{n-1} \xi^r(r) dr \right] \check{j}_l(kr_{12})r_{12}^2 dr_{12} = \int_0^\infty \xi^r(r_{12})j_0(kr_{12})r_{12}^2 dr_{12}. \quad (4.21)$$

Now switch the order of integration of the left hand side of equation 4.21:

$$\int_0^\infty n r^{n-1} \left[ \int_r^\infty r_{12}^{2-n} \check{j}_l(kr_{12}) dr_{12} \right] \xi^r(r) dr = 4\pi \int_0^\infty \xi^r(r_{12})j_0(kr_{12})r_{12}^2 dr_{12}. \quad (4.22)$$

Rename the integration variables so that the semi-infinite integrals are over the same variable:

$$\int_0^\infty n s^{n-1} \left[ \int_s^\infty r_{12}^{2-n} \check{j}_l(kr_{12}) dr_{12} \right] \xi^r(s) ds = \int_0^\infty \xi^r(s)j_0(ks)s^2 ds. \quad (4.23)$$

One solution, perhaps the most restrictive solution, to this equation is found by equating the integrands of the two sides of the equation:

$$n s^{n-1} \left[ \int_s^\infty r_{12}^{2-n} \check{j}_l(kr_{12}) dr_{12} \right] \xi^r(s) = \xi^r(s)j_0(ks)s^2. \quad (4.24)$$

This can be further simplified to

$$\int_s^\infty r_{12}^{2-n} \check{j}_l(kr_{12}) dr_{12} = \frac{s^{3-n}}{n} j_0(ks). \quad (4.25)$$

To find the new window differentiate equation 4.25 with respect to  $s$ . This gives:

$$-s^{2-n}\check{j}_l(ks) = (-ksj_1(ks) + (3-n)j_0(ks))\frac{s^{2-n}}{n}. \quad (4.26)$$

The final form of the windows for functions of just one of the averages of  $\xi^r$  (we have employed recursion relations for spherical Bessel functions so that only even functions remain) is:

$$\check{j}_l(ks) = \frac{1}{n} \left( \frac{k^2 s^2}{3} j_0(ks) + \frac{k^2 s^2}{3} j_2(ks) + (n-3)j_0(ks) \right). \quad (4.27)$$

So estimates of  $\xi_\alpha(r)$  can be transformed into estimates of  $f_i(\beta)\xi(k)$ .

It turns out that the inverse relation is equally (if not more) important. Fortunately, the inverse window functions are even easier to calculate. The equation

$$\frac{4\pi}{(2\pi)^3} \int_0^\infty \tilde{j}_n(kr_{12})P(k)k^2 dk = \frac{n}{r_{12}^n} \int_0^{r_{12}} r^{n-1}\xi(r)dr \quad (4.28)$$

needs to be solved for  $\tilde{j}_n(kr_{12})$ . However, because

$$\frac{4\pi}{(2\pi)^3} \int_0^\infty j_0(kr)P(k)k^2 dk = \xi(r) \quad (4.29)$$

equation 4.28 can be rewritten

$$\frac{4\pi}{(2\pi)^3} \int_0^\infty \tilde{j}_n(kr_{12})P(k)k^2 dk = \frac{n}{r_{12}^n} \int_0^{r_{12}} r^{n-1} \left[ \frac{4\pi}{(2\pi)^3} \int_0^\infty j_0(kr)P(k)k^2 dk \right] dr \quad (4.30)$$

or

$$\tilde{j}_n(kr_{12}) = \frac{n}{r_{12}^n} \int_0^{r_{12}} r^{n-1} j_0(kr) dr. \quad (4.31)$$

Now the  $\Xi$ 's can be transformed to and from  $P(k)$ .

### 4.3 Pair Weight Compression with Spherical Redshift Distortions

In chapter 2 we showed how to extract power spectra from redshift surveys disregarding redshift distortions. However, the pair weight compression method becomes more powerful when redshift distortions are explicitly included. Not only can the real space power spectrum be extracted, but it can also yield an estimate of  $\beta$ .

First of all, look at the resulting estimate of  $\hat{\xi}_\gamma$  (Equation 2.20):

$$\hat{\xi}_\gamma = M_{\gamma\alpha}^{-1} B_{\alpha ij} W_{aij} K_{ab}^{-1} W_{bkl} \delta_k \delta_l \quad (4.32)$$

where

$$M_{\alpha\gamma} \equiv B_{\alpha ij} W_{aij} K_{ab}^{-1} W_{bkl} B_{\gamma kl} \quad (4.33)$$

and

$$K_{ab} \equiv W_{aij} \langle \mathbf{e}_{ijkl} \rangle W_{bkl} \approx 2 W_{aij} C_{ik} C_{jl} W_{bkl}. \quad (4.34)$$

No major modifications of these formulae are necessary to explicitly include redshift distortions. However, the meanings of some of these symbols change.

### 4.3.1 Shape Functions

The most necessary change comes from the definition of  $B$ . Recall from chapter 2 that  $B$  comes from taking the derivative of  $\langle \delta_i \delta_j \rangle$  with respect to the parameters we wish to measure. In the case where redshift distortions are ignored, the only parameters of interest were the value of the power spectrum at various values of  $k$ . Now that redshift distortions are included

$$\langle \delta_i \delta_j \rangle = \xi_{ij} + \bar{n}^{-1}(\mathbf{r}_i) \delta_{ij} = \sum_{\alpha} \check{\xi}_{\alpha} B_{\alpha ij} + \bar{n}^{-1}(\mathbf{r}_i) \delta_{ij}. \quad (4.35)$$

Now each value of  $\check{\xi}_{\alpha}$  (after transforming into Fourier space) gives a measurement of the power spectrum at a particular wavenumber times some function of  $\beta$ . This means that the information about  $P_k$  and  $\beta_{(k)}$  is contained within the parameters  $\check{\xi}_{\alpha}$ . Of course, one would like to go straight to measuring the actual values of  $P_k$  and  $\beta$  with no intermediate steps. However, because the data lie in real space, the direct measurement of  $P_k$  and  $\beta$  is infeasible. One can, however, measure the parameters  $\check{\xi}_{\alpha}$ , transform these parameters into Fourier space, and then do a maximum likelihood calculation using these parameters (and the associated error bars) to find the values of  $P_k$  and  $\beta$ .

So differentiating Equation 4.35 with respect to  $\check{\xi}_\alpha$  yields just the shape functions  $B_{\alpha ij}$ . Actually, the differentiation occurs at a particular separation so the shape functions also include a delta function:

$$\frac{\partial \langle \delta_i \delta_j \rangle}{\partial \check{\xi}_\alpha} = \delta(r_\alpha - |\mathbf{r}_i - \mathbf{r}_j|) B_{\alpha ij}. \quad (4.36)$$

Notice that the subscript  $\alpha$  stands for the subscript of the particular shape function at the particular separation.

### 4.3.2 Window Functions

At each separation, instead of measuring one parameter, there are now seven parameters to measure. Because there are seven parameters to measure it makes sense to use at least seven window functions at each separation. In the case of the plane-parallel approximation the window functions to use were obvious. The three different  $\Xi$ 's were orthogonal to one another. Each window was tuned to measure exactly one of the parameters of interest. This meant that the data needed to be weighted by the Legendre polynomials of order 0, 2 and 4. In the case of spherical distortions the parameters are not orthogonal to one another. However, we can still make an analogous choice of window functions. In other words, use window functions like those described in chapter 2 weighted by the shape functions. This should be a good choice because each window function will be picking out a majority of the information about a particular parameter. Adding additional windows will again raise the amount of information available to the parameter estimation. However, because the computation is so lengthy as it is, we want to use the smallest possible number of parameters.

### 4.3.3 Better Final Parameters

The seven  $\Xi$ 's are the natural parameters to measure. However, the physical meaning of some of the  $\Xi$ 's is not so clear. The Fisher matrix representation allows us to take a linear combination of the measured parameters and form a more physically motivated set of parameters.

To do this notice that

$$F_{\theta_1, \theta_2} = \frac{\partial \phi_1}{\partial \theta_1} F_{\phi_1, \phi_2} \frac{\partial \phi_2}{\partial \theta_2} \quad (4.37)$$

and

$$\Theta_i = \frac{\partial \theta_i}{\partial \phi_j} \Phi_j. \quad (4.38)$$

However, it is common to have  $\Phi$  as a function of  $\Theta$  rather than the other way around. In this case, the quantity  $\frac{\partial \theta_i}{\partial \phi_j}$  is unknown. In fact, if there are fewer values of  $\Theta$  than  $\Phi$  then  $\Theta$  may be over-determined. This means that the best choice is the one where the values of  $\Theta$ , as determined by  $\Phi$ , make the best use of the data.

The best use of the data can be obtained by looking at the log-likelihood function

$$-\ln \mathcal{L} = \frac{1}{2} (\Phi_i - \hat{\Phi}_i) F_{\phi_i \phi_j} (\Phi_j - \hat{\Phi}_j) \quad (4.39)$$

which can be rewritten in terms of  $\theta$ :

$$-\ln \mathcal{L} = \frac{1}{2} (\Phi_i - \frac{\partial \phi_i}{\partial \theta_k} \hat{\Theta}_k) F_{\phi_i \phi_j} (\Phi_j - \frac{\partial \phi_j}{\partial \theta_l} \hat{\Theta}_l). \quad (4.40)$$

Setting the first derivative of  $-\ln \mathcal{L}$  with respect to  $\hat{\Theta}$  equal to zero gives the best estimate of  $\hat{\Theta}$ :

$$\frac{\partial(-\ln \mathcal{L})}{\partial(\hat{\Theta}_k)} = -\frac{\partial \phi_i}{\partial \theta_k} F_{\phi_i \phi_j} (\Phi_j - \frac{\partial \phi_j}{\partial \theta_l} \hat{\Theta}_l) = 0. \quad (4.41)$$

Rearranging this equation yields:

$$\frac{\partial \phi_i}{\partial \theta_k} F_{\phi_i \phi_j} \Phi_j = \frac{\partial \phi_i}{\partial \theta_k} F_{\phi_i \phi_j} \frac{\partial \phi_j}{\partial \theta_l} \hat{\Theta}_l. \quad (4.42)$$

Using equation 4.37 gives:

$$\frac{\partial \phi_i}{\partial \theta_k} F_{\phi_i \phi_j} \Phi_j = F_{\theta_k \theta_l} \hat{\Theta}_l. \quad (4.43)$$

This means that

$$\Theta_i = F_{\theta_i \theta_j}^{-1} \frac{\partial \phi_k}{\partial \theta_j} F_{\phi_k \phi_l} \Phi_l \quad (4.44)$$

is the best representation of  $\Theta$  given  $\Phi$ . Notice that if there are as many values of  $\Theta$  as there are of  $\Phi$  then equation 4.44 collapses to the expected:

$$\Theta_i = \left(\frac{\partial\phi_j}{\partial\theta_i}\right)^{-1}\Phi_j. \quad (4.45)$$

Using this formalism, each of the spherical distortion parameters  $\Xi$  can be expressed as a function of  $P(k)$ ,  $\beta P(k)$  and  $\beta^2 P(k)$ . These are parameters that make more physical sense. The first is just the galaxy-galaxy power spectrum. The second is the galaxy-velocity power spectrum. The third is the velocity-velocity power spectrum Hamilton and Tegmark (2000b). Not only are these more physically motivated, but since we are ultimately trying to measure the power spectrum and  $\beta$ , this gives us a small number of parameters with which to do a non-linear  $\chi^2$  fit.

#### 4.3.4 Discreteness Correction

In Chapter 3, we discussed the fact that a discrete representation of the correlation function can cause errors in the measurement of the power spectrum. This is true for the case with redshift distortions as well. Upon transforming the prior value of the  $\Xi$ 's (using the proper generalized FFT window) the result should be the prior power spectrum. Unfortunately, for any realistic prior power spectrum, a discrete FFT of the  $\Xi$ 's does not yield exactly the prior power spectrum. To correct for this, it is necessary to multiply the result of each FFT by the discreteness correction ( $\zeta_i$ ):

$$\zeta_i(k) = \frac{P_{\text{prior}}(k)}{\sum_i \Xi_{i\text{prior}}(r) \tilde{j}_i(r)}. \quad (4.46)$$

This means that, after the correction, if the measurement of  $\Xi$  corresponds exactly to the prior value of  $\Xi$  then the measurement of  $P(k)$  will be the prior value of the power spectrum, as expected.



### 4.3.5 Correlation and Decorrelation

As in Chapter 3, the result from equation 4.32 is a measurement of the parameter in question with no width in  $k$ . This result actually anti-correlates neighboring points. That is to say, the cross-covariance of two neighboring points will be negative. This is not an ideal way to present the data, particularly because we expect neighboring points to have similar values. The two ways that are sensible are the correlated and decorrelated versions. The decorrelated version has the advantage of having each estimate being uncorrelated with any other estimate. This means that the decorrelated points can be used in a least squares fit as a simple sum of the deviation from each estimate. For obtaining the best least-squares result, the decorrelated estimates are clearly the best choice.

The correlated estimates give results which are easy to visually compare one set of results with another. That is to say, that when using correlated points, the scatter is small enough that the most likely value of the parameter is easily determined from a graph.

The method for obtaining the correlated and decorrelated estimates were given in Chapter 3. There is one major additional concern when using the redshift distortion analysis. This is that the decorrelation (correlation) matrix naturally combines some of the  $P(k)$  estimate in with pieces of the  $\beta P(k)$  and the  $\beta^2 P(k)$  terms. A representative row of the Fisher matrix is shown in Figure 4.1. In going from the anti-correlated version of the estimates, it would be nice to avoid entangling one type of estimate with another.

To disentangle the three estimates from one another first define  $\hat{q}$  (here the hat denotes a measurement):

$$\langle \hat{q}_\alpha \rangle = \mathbf{F}_{\alpha\gamma} \xi_\gamma. \quad (4.47)$$

where  $\mathbf{F}$  is just the Fisher matrix and  $\xi_\gamma$  is one of the new parameters  $P(k)$ ,  $\beta P(k)$ , or  $\beta^2 P(k)$ . Notice that the Fisher matrix smoothes over all three parameters over all wavenumbers. To

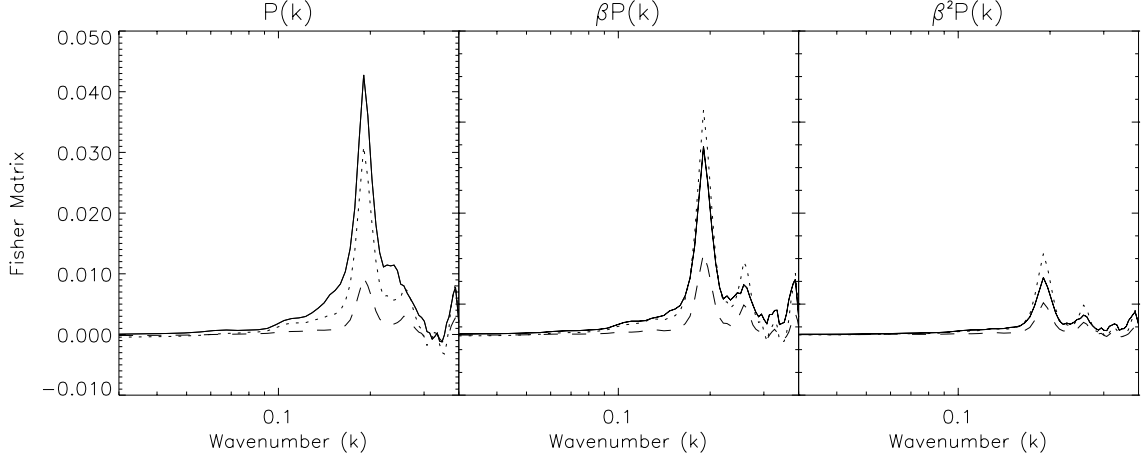


Figure 4.1: The rows of the Fisher matrix are one of the smoothing windows. Plotted here are rows corresponding to the  $P(k)$  term (solid) the  $\beta P(k)$  term (dotted) and the  $\beta^2 P(k)$  term (dashed). Each row is peaked about the nominal wavelength and contains contributions from each of the three terms.

explicitly disentangle the three parameters rewrite equation 4.47:

$$\langle \hat{q}_a(k) \rangle = \sum_b \left[ \frac{\langle \hat{\xi}_b(k) \rangle}{\xi_b(k)} \right] \sum_k \mathbf{F}_{ab}(k, k') \xi_b(k'). \quad (4.48)$$

Notice that the Greek indices have been expanded into Latin indices over the type of parameter and  $k$  the wavenumber. Also the summations are now explicitly included because some of the following equations are not summed. The term in square brackets is added to form the definition of  $\mathcal{F}_{ab}$ :

$$\mathcal{F}_{ab}(k) \equiv \frac{1}{\xi_b(k)} \sum_k \mathbf{F}_{ab}(k, k') \xi_b(k') \quad (4.49)$$

with no summation over  $b$ . This allows 4.48 to be rewritten

$$\langle \hat{q}_a(k) \rangle = \sum_b \mathcal{F}_{ab}(k) \langle \hat{\xi}_b(k) \rangle \quad (4.50)$$

with no summation over  $k$ . Now,  $\mathcal{F}_{ab}(k)$  is a  $3 \times 3$  matrix at each value of  $k$ . To obtain a measurement of each of the  $\xi$ 's multiply both sides of equation 4.50 by the inverse of  $\mathcal{F}_{ab}(k)$ :

$$\xi_b(k) = \sum_a \mathcal{F}_{ba}^{-1}(k) \langle \hat{q}_a(k) \rangle \quad (4.51)$$

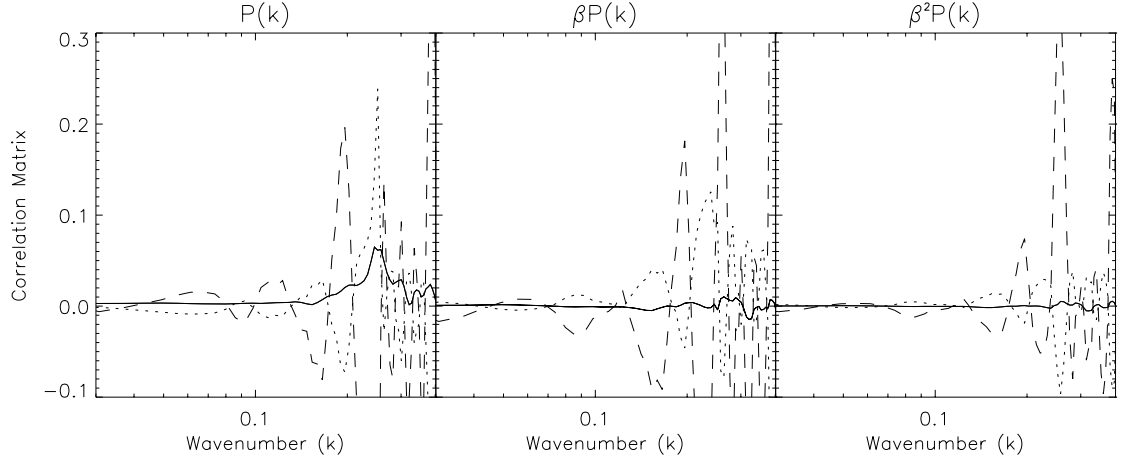


Figure 4.2: Representative rows of the correlation matrix. Each line shows the window (one of the rows of  $\mathcal{C}$ ) over which one of the parameters has been smoothed to obtain a correlated estimate. These are the same rows that are plotted in Figure 4.1. The  $P(k)$ ,  $\beta P(k)$  and  $\beta^2 P(k)$  rows are shown by the solid, dotted and dashed lines respectively.

or

$$\hat{\xi}_b(k) = \sum_a \mathcal{F}_{ba}^{-1}(k) \hat{q}_a. \quad (4.52)$$

Equation 4.52 gives the estimate of  $\xi_b(k)$  which has an expectation value of the prior power spectrum at the given wavelength. Furthermore, each of the parameters has no contribution from the others (provided that the true power spectrum is similar in shape to the prior power spectrum). When starting with the anti-correlated estimate of  $\hat{\xi}_b$  (i.e.,  $\mathbf{F}_{ab}^{-1} \hat{q}_b$ ), to obtain the untangled and correlated measurement of  $\xi$  use the band-power window matrix  $\mathcal{C}_{ab}$ :

$$\hat{\xi}_a^{\text{corr}}(k) = \sum_b \mathcal{F}_{ab}^{-1}(k) \sum_{c,k'} \mathbf{F}_{bc}(k, k') \hat{\xi}_c(k') \equiv \sum_{c,k'} \mathcal{C}_{ac}(k, k') \hat{\xi}_c(k') \quad (4.53)$$

where

$$\mathcal{C}_{ac}(k, k') \equiv \sum_b \mathcal{F}_{ab}^{-1}(k) \mathbf{F}_{bc}(k, k'). \quad (4.54)$$

Figure 4.2 shows rows of the correlation matrix ( $\mathcal{C}$ ). Notice that each of the parameters is nicely peaked about the nominal wavenumber within the same type of parameter. Similarly,

the  $P(k)$  term is nicely zero in the other parameters. However, the contributions to the terms  $\beta P(k)$  and  $\beta^2 P(k)$  from the  $P(k)$  term can be quite large if the power spectrum is not similar in shape to the prior power spectrum.

To obtain the covariance matrix of the correlated estimates take

$$\langle \Delta \hat{\xi}_a^{\text{corr}}(k) \Delta \hat{\xi}_b^{\text{corr}}(k') \rangle = C_{ac}(k, k'') \langle \Delta \hat{\xi}_c(k'') \Delta \hat{\xi}_d(k''') \rangle C_{bd}(k', k'''). \quad (4.55)$$

Next recognize that by the definition of the Fisher matrix

$$\langle \Delta \hat{\xi}_c(k'') \Delta \hat{\xi}_d(k''') \rangle = \mathbf{F}_{cd}^{-1}(k'', k'''), \quad (4.56)$$

and from the definition of  $\mathcal{C}$  (Equation 4.54):

$$\langle \Delta \hat{\xi}_a^{\text{corr}}(k) \Delta \hat{\xi}_b^{\text{corr}}(k') \rangle = \mathcal{F}_{ae}^{-1}(k) \mathbf{F}_{fe}(k, k') \mathcal{F}_{bf}^{-1}(k') \quad (4.57)$$

with implicit summation over each pair of Latin indices and no summation over  $k$  or  $k'$ .

Notice that if the smoothing matrix is instead defined as (Hamilton and Tegmark, 2000b)

$$\mathbf{M} = \mathbf{F}^{1/2} \quad (4.58)$$

then the covariance matrix is then the unit matrix. Here the square-root of a matrix is any matrix that satisfies:

$$\mathbf{F} = \mathbf{M}^\top \mathbf{M}. \quad (4.59)$$

For example, let

$$\hat{z}_\alpha \equiv \mathbf{M}_{\alpha\beta} \hat{\xi}_\beta \quad (4.60)$$

then

$$\langle \Delta \hat{z}_\alpha \Delta \hat{z}_\beta \rangle = \mathbf{1}_{\alpha\beta} \quad (4.61)$$

where  $\mathbf{1}$  is the unit matrix. This allows the estimates to be decorrelated from one another.

Once again, however, it would be better if the estimate of a particular parameter were disentangled from the other parameters. To do this, rewrite

$$\langle \hat{z}_a(k) \rangle = \sum_b \left[ \frac{\langle \hat{\xi}_b(k) \rangle}{\xi_b(k)} \right] \sum_k \mathbf{M}_{ab}(k, k') \xi_b(k'). \quad (4.62)$$

So defining  $\mathcal{M}$ :

$$\mathcal{M}_{ab}(k) \equiv \frac{1}{\xi_b(k)} \sum_k \mathbf{M}_{ab}(k, k') \xi_b(k') \quad (4.63)$$

again with no summation over  $b$ . This allows 4.62 to be rewritten

$$\langle \hat{z}_a(k) \rangle = \sum_b \mathcal{M}_{ab}(k) \langle \hat{\xi}_b(k) \rangle. \quad (4.64)$$

Now to disentangle the decorrelated parameters multiply each side by the inverse of  $\mathcal{M}$ :

$$\hat{\xi}_\beta^{\text{decorr}}(k) = \mathcal{M}_{\beta\alpha}^{-1}(k) \hat{z}_\alpha(k). \quad (4.65)$$

Notice that by disentangling the different parameters at each wavenumber the three parameters at each wavelength are no longer decorrelated from one another. However, all parameters at each wavenumber remain decorrelated from all parameters at every other wavenumber. Specifically,

$$\langle \Delta \hat{\xi}_a^{\text{decorr}}(k) \Delta \hat{\xi}_b^{\text{decorr}}(k) \rangle = \mathcal{M}_{ac}^{-1} \mathcal{M}_{bc}^{-1} \quad (4.66)$$

with summation over  $c$  only.

Figure 4.3 shows the representative decorrelated smoothing windows. Notice that in each case the smoothing window of the relevant parameter is sharply peaked about the nominal wavenumber. Also the contribution to  $P(k)$  from the  $\beta P(k)$  and  $\beta^2 P(k)$  terms is minimal as is the contribution to the  $\beta P(k)$  term from the  $\beta^2 P(k)$  term.

#### 4.4 Information in LCRS

Figure 4.4 shows the information contained within each of the decorrelated parameters. LCRS has the highest amount of information for the redshift space power spectrum. It is clear

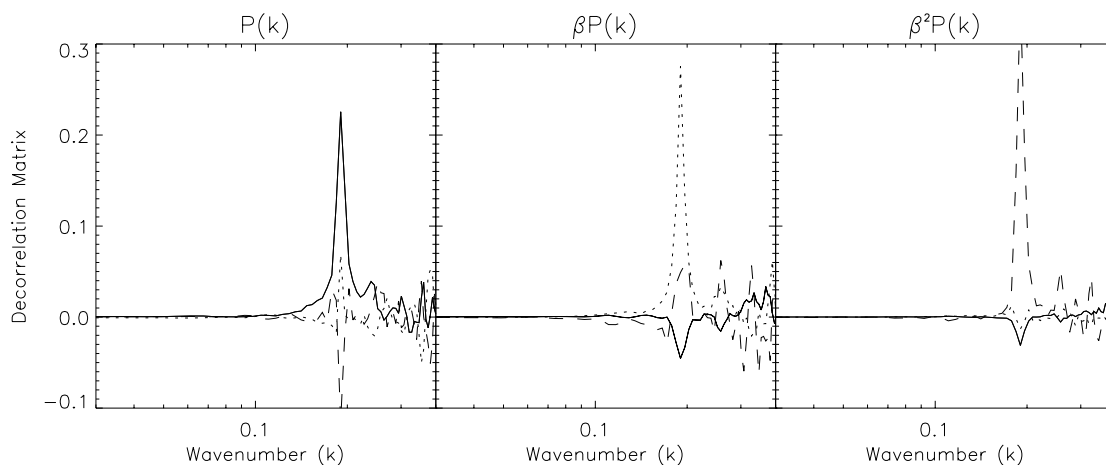


Figure 4.3: Representative rows of the decorrelation matrix. Each line shows the window (one of the rows of  $\mathcal{D}$ ) over which one of the parameters has been smoothed to obtain a decorrelated estimate. These are the same rows that are plotted in Figure 4.1 and Figure 4.2. The  $P(k)$ ,  $\beta P(k)$  and  $\beta^2 P(k)$  rows are shown by the solid, dotted and dashed lines respectively.

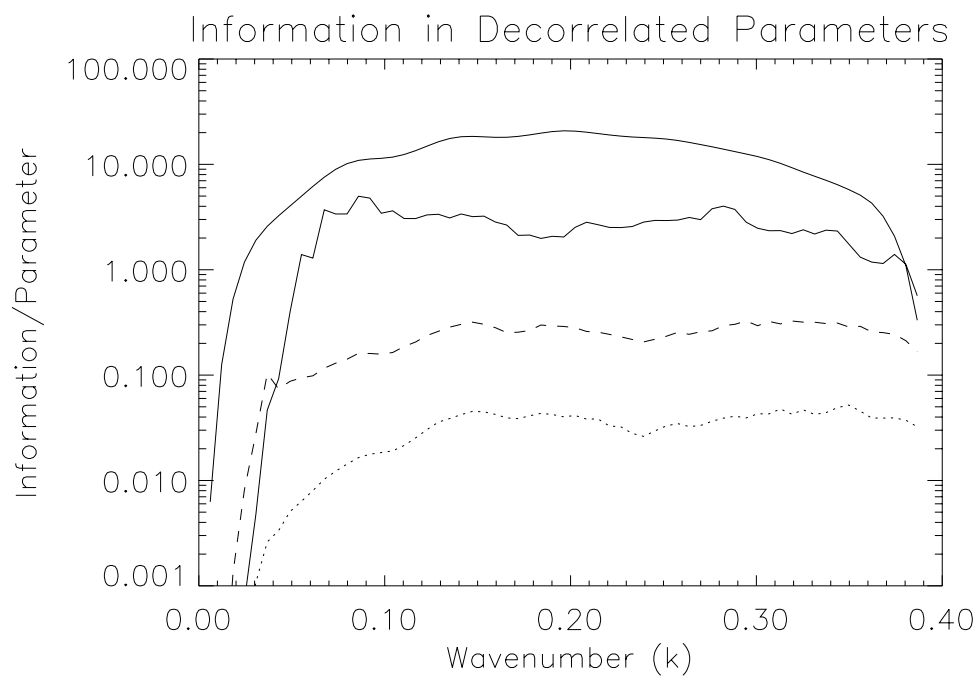


Figure 4.4: Information in the decorrelated parameters. The upper solid line is for the redshift space power spectrum. The lower solid line is for the real space power spectrum. The dashed line is for  $\beta P(k)$  and the dotted line is for  $\beta^2 P(k)$ .

that this should always be the case. Upon adding additional measured parameters, the information per parameter should go down.

In this case the total amount of information appears to be reduced by at least a factor of two. Clearly, the information has not been apportioned out to the various terms. It appears that the terms containing  $\beta$  do not contain a large fraction of the information. There has been an actual leak in the analysis. This is largely due to the inability to measure the covariance between parameters exactly. The off diagonal elements of the covariance matrix are typically calculated to about 5 per cent expected errors. This means that if one were able to measure the off-diagonal elements of the covariance matrix to better precision, the information retained would go up by as much as a factor of a few.

This Figure also shows that, to get an appreciable amount of signal for the  $\beta P(k)$  term or the  $\beta^2 P(k)$  term, one must use coarse gridding.

## 4.5 Discussion

In this chapter, we discuss the extension of the pair weight method to the case of spherical redshift distortions. In linear theory, the spherical redshift distortions can be modeled as the sum of seven shape functions times functions of the correlation function (the  $\Xi$ 's) and  $\beta$  (the  $g$ 's). Using the generalized Fourier transform functions, each of the  $\Xi$ 's can be transformed into a measurement of the real space power spectrum. Using the properties of the Fisher matrix the seven estimates can be combined into three physically meaningful quantities:  $P(k)$ ,  $\beta P(k)$  and  $\beta^2 P(k)$ .

Using the Las Campanas Redshift Survey, we showed that the estimates of  $P(k)$  can be disentangled from the estimates of  $\beta P(k)$  and  $\beta^2 P(k)$ . The Fisher matrix, the disentangled correlated band-powers and the disentangled decorrelated band-powers are all sharply peaked about the wavenumber of interest. Thus, any smoothing will leave an estimate that is representative of the nominal wavenumber.

It is clear that the measurement of the real space galaxy-galaxy power spectrum is a

more difficult measurement than that of the redshift space galaxy-galaxy power spectrum. For LCRS, using the pair weight compression method, the expected errors for the real galaxy-galaxy power spectrum are only about a factor of two as large as those for the redshift space power spectrum.



## Chapter 5

### Measurements of $\beta$ and the Power Spectrum from the Las Campanas Redshift Survey

The Las Campanas Redshift Survey is one of the largest publicly available 3-dimensional galaxy catalogs. The selection criteria allowed the team to extract the largest number of galaxy redshifts given their telescope time. Unfortunately, this method of selecting targets makes the resulting survey more difficult to analyze. Although the catalog has been publicly available for nearly four years, there are no published real space power spectra. This is remarkable particularly since LCRS is still the publicly available catalog with the largest number of redshifts. By contrast, the IRAS PSCz catalog has been publicly available for less than a year and there are two available real-space power spectra (Tadros et al., 1999; Hamilton et al., 2000).

Matsubara et al. (2000) have made the most progress, by publishing an estimate for  $\beta$  (the amplitude of the redshift distortions),  $\sigma_8$  (the amplitude of the power spectrum) and  $\Gamma$  (a shape of the power spectrum). This parameterization allows for a partial characterization of the power spectrum. The parameterization does not reveal how well the data fit the simple model and may hide interesting features. As the measurements of cosmological parameters (from galaxy catalogs as well as other sources other sources) become more precise, these features become more interesting. Particularly interesting is the size scale ( $k$ ) at which the redshift distortions become nonlinear.

In chapter 4, I generalized the pair weight compression method to allow the analysis of redshift distortions and real space power spectra. The generalization calls for measuring (up to)

seven parameters at each separation

$$\xi^s(r) = \sum_{i=1}^7 \check{\xi}_i B_i(r_1, r_2, r_{12}) \quad (5.1)$$

where  $\xi^s(r)$  is the redshift space correlation function, each of the  $\check{\xi}_i$ 's is a measurable parameter and each of the  $B_i$ 's is function of the shape of the triangle formed by a galaxy pair and the observer. Each  $\check{\xi}_i$  can be transformed into a measure of  $P(k)$  and  $\beta$ . By properly combining the transformed  $\check{\xi}_i$ 's one is able to obtain measurements of, not only the galaxy-galaxy power spectrum ( $P(k)$ ), but also the galaxy-velocity ( $\beta P(k)$ ) and the velocity-velocity ( $\beta^2 P(k)$ ) power spectra. Unfortunately, chapter 4, also said that the measurements of  $\beta P(k)$  and  $\beta^2 P(k)$  are too imprecise to be measured without seriously reducing the number of independent estimates. Here I apply the pair weight method to the Las Campanas Redshift Survey to obtain a real space power spectrum. Then using this power spectrum as the shape of the true power spectrum, I perform a least-squares fit to the amplitude of the power spectrum and  $\beta$ .

## 5.1 Real Space Power Spectrum

Figure 5.1 shows the correlated real space power spectrum as derived by the pair weight method. The error bars on the real space power spectrum are larger than those for the redshift space power spectrum. This is to be expected since there are now more parameters to fit with the same information. The real space power spectrum appears to be similar to the redshift space counterpart. Due to aliasing at the large  $k$  end (see Figure 3.2), and inaccuracy at the low  $k$  end, the inner half of the points (from  $k = .1 - .3$ ) of both curves should be considered to be the most reliable. The similarity in amplitude of the real space power spectrum and the redshift space power spectrum would imply a low value of  $\beta$ .

The calculation of this estimate only uses shape functions 0, 2 and 4 as defined in Appendix A. The inclusion of additional shape functions reduced the quality of the estimate. That is to say that, not only was the estimate less believable, but the error bars were enlarged. When calculating the Fisher matrix with Monte Carlo techniques, there are inevitably errors in some

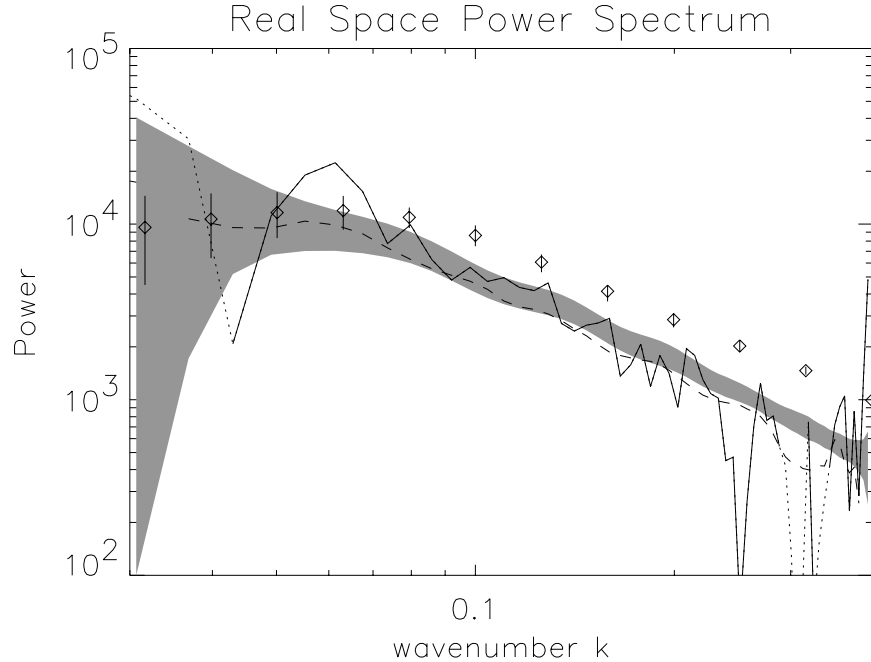


Figure 5.1: Power spectrum of the cut catalog’s 21996 LCRS galaxies. The solid line shows where the measured real space power spectrum is positive. The line is dotted where either end of the segment is negative. The shaded region is the expected error about the prior power spectrum. The dashed line is the redshift space power spectrum. The diamonds with error bars is the redshift space power spectrum measured by Lin et al. (1996b).

of the matrix elements. Unfortunately, when the matrices are inverted these errors can dominate the calculation. In particular, errors in the off diagonal terms involving  $\Xi_{0, 2 \text{ or } 4}$  and  $\Xi_{1, 3, 5 \text{ or } 6}$  can contaminate the calculation of the estimates of  $\Xi_{0, 2 \text{ or } 4}$ . It is also true that the scatter in the measurement using only shape function 0, 2, and 4 is larger than the calculated error. This can be seen in Figure 5.1. The fluctuations in the correlated power spectrum are larger than the predicted errors (shaded region) especially at the large  $k$  end. This is addressed in the following section.

## 5.2 Calculated Errors

The fluctuations in the measurements of neighboring points is empirically larger than the error bars calculated from the Fisher matrix. It is clear that part of the problem is that high

covariances between the measurements of two parameters gives a lower measurement of the error than is reasonable. In addition to this, miscalculation of Fisher matrix elements increases the problem. We have not yet been able to correct the error bars to reflect the true level of the errors. In the remainder of this Chapter, we use an empirically calculated, multiplicative factor to increase the level of the errors. That is to say, we use the relative scatter in the nominal “unit variance” calculation as the correction factor.

### 5.3 Measurement of $\beta$

Figure 4.4 shows that the amount of information in any one estimate of  $\beta P(k)$  or  $\beta^2 P(k)$  is very low. For this reason, to find the best estimate for  $\beta$  it is important to use all the available information. On the smallest scales, however, nonlinearities may become important. The most likely result of nonlinear contamination is that the estimate of  $\beta$  would be reduced. This is because the linear model for redshift distortions squashes structures along the line of sight. However, on the smallest scales the “fingers of God” clearly stretch structures along the line of sight. At the same time, however, these bins at large  $k$  will also be contaminated by aliasing. For these reasons, the calculation for the most likely value of  $\beta$  is performed at each wavenumber by including information cumulatively up to that wavenumber. If nonlinearities become important then  $\beta$  should drop perceptibly at large wavenumbers.

Figure 5.2 shows the best fit value of  $\beta$  for each wavenumber. In each case, the power spectrum is assumed to have the shape of the measured correlated power spectrum. The first six bins of the power spectrum have (unphysical) negative values. The power spectrum in these bins is not accurately measured. For this reasons, these bins have been ignored for all of the calculations. In each calculation, the only measured parameters are the amplitude of the power spectrum and  $\beta$ . As expected, as more data is added to the calculation the variation in the best measurement is reduced. Figure 5.2 shows no conclusive evidence for entering the nonlinear regime. This is only slightly surprising since the smallest length scale probed is  $8h^{-1}\text{Mpc}$  and the location where one would expect to see nonlinearities is also affected by aliasing.

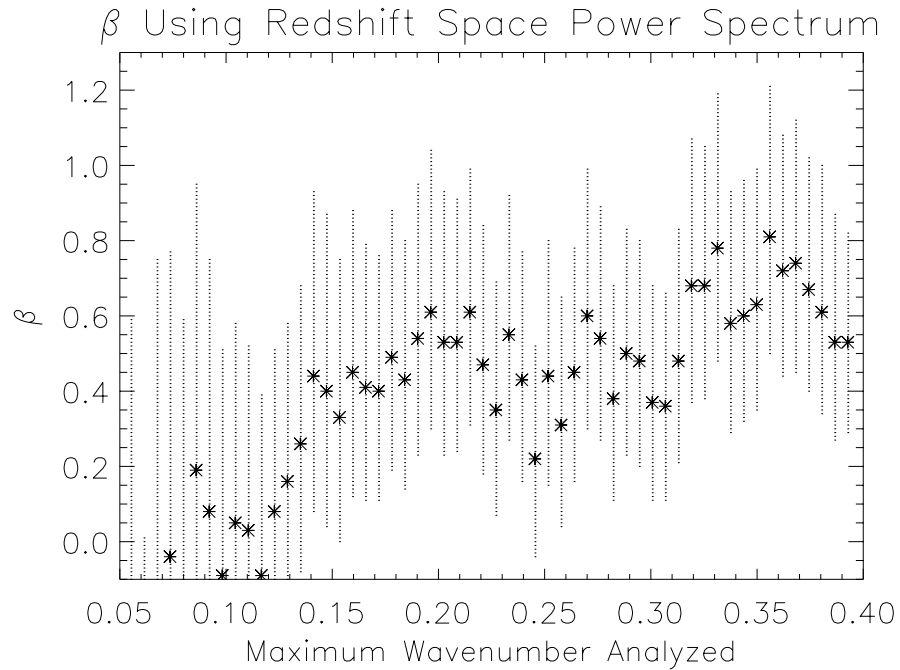


Figure 5.2: This Figure shows the  $1 - \sigma$  errors and best fit for  $\beta$  using the amplitude of the correlated power spectrum as the only other free parameter. At each plotted wavenumber, the fit uses information at smaller wavenumbers only. The stars show the best fit value for  $\beta$  at each wavenumber.

Figure 5.3 shows the best fit value for the measurement of the amplitude of the correlated power spectrum. As expected, as more information is added to the calculation the amplitude of the power spectrum becomes less uncertain and the value tends toward unity. This means that the best measured amplitude for the decorrelated power spectrum is the same as for the correlated power spectrum.

At each wavenumber the amplitude and the value of  $\beta$  are highly anti-correlated. Figure 5.4 shows the typical error region in the amplitude- $\beta$  plane. The marginalized  $1 - \sigma$  error in each parameter is the most extreme value along the  $1 - \sigma$  contour. This marginalized error is the error bar plotted in Figures 5.2 and 5.3.

The final result is that the amplitude of the correlated, real space power spectrum should be  $1. \pm .25$  times the curve shown in Figure 5.1 and  $\beta \approx 0.55^{+.35}_{-.30}$ . This value is obtained from

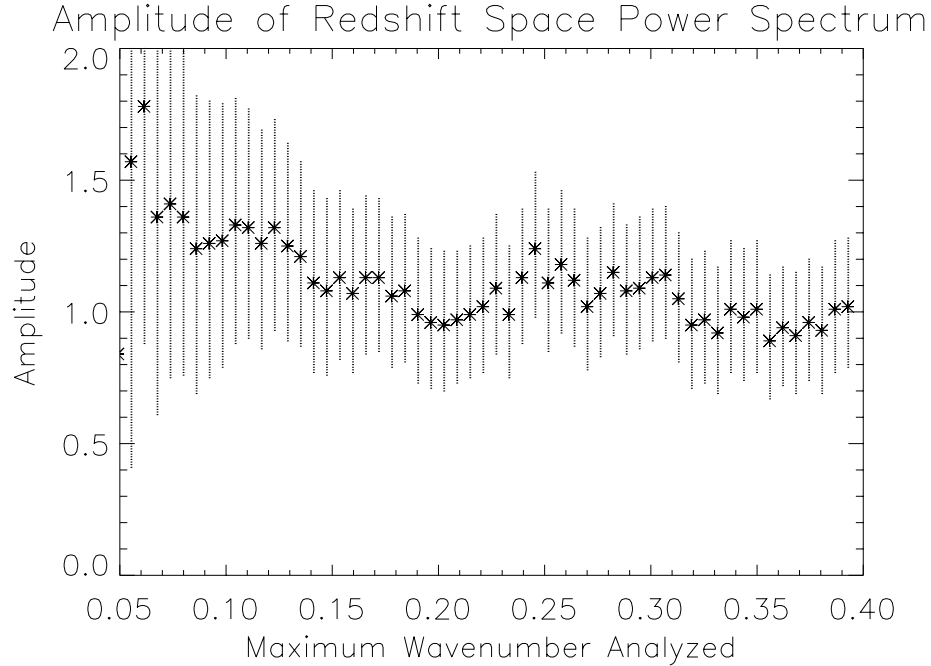


Figure 5.3: This figure shows the best fit amplitude and the  $1 - \sigma$  error for the amplitude of the correlated power spectrum. At each wavenumber, the fit uses only information from smaller wavenumbers. The only other free parameter is  $\beta$ .

the best overall fit at  $k = .209h\text{Mpc}^{-1}$ . This is a good choice because it uses as much of the linear information as possible while not being contaminated by aliasing.

#### 5.4 Discussion and Conclusions

The pair weight compression method is able to extract a real space power spectrum and an estimate of  $\beta$  from LCRS. Unfortunately, only three of the seven shape functions yielded useful results. In fact, including the additional shape functions provided an estimate that was clearly incorrect. In order to include information from all seven shape functions it is necessary to calculate all of the matrix elements in the Fisher matrix to adequate precision. For this calculation, the monopole-monopole elements were calculated to expected errors of 2.5 percent. The off-diagonal elements typically had errors of a few percent but routinely there were elements with errors of tens of percent. Reducing the calculation to only those shape functions which

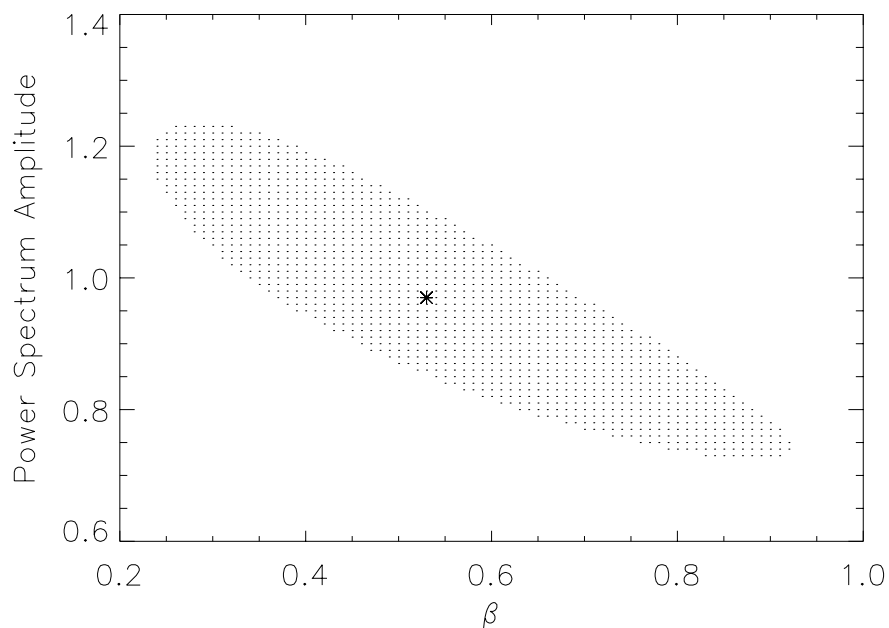


Figure 5.4: The  $1-\sigma$  error region for the calculation including all wavenumbers up to  $k = 0.209 \text{ h Mpc}^{-1}$ . The star is the best fit value.

are non-zero at small separations, reduced the errors caused by the off-diagonal terms. However, the scatter in the calculated estimates were still larger than the calculated error bars. As a result, the real space power spectrum is noisier than the redshift space power spectrum. This would be true even if the Fisher matrix were calculated exactly because the same information is now being shared amongst three estimates. The resulting real space power spectrum has similar shape to the redshift space power spectrum. This is not at all surprising since a majority of the analysis is on wavenumbers which should not be greatly affected by nonlinearities.

The estimate of  $\beta \approx .55_{-.30}^{+.35}$  is consistent with a flat low matter density universe. In particular, it is consistent with the estimates from rich clusters (Carlberg et al., 1998) and cosmic microwave background measurements (de Bernardis et al., 2000). A flat matter dominated universe is once again ruled out unless the bias factor is rather large. Although the estimate of  $\beta$  is larger than that of Matsubara et al. (2000) ( $\beta = .30 \pm .39$ ) the estimates are consistent within

the errors.



## Chapter 6

### Conclusions

#### 6.1 Pair Weight Method

We presented a new method for analyzing galaxy redshift surveys called the pair weight compression method. This method is able to assign weights to each galaxy pair within the catalog. It is also able to determine the error bars on each measured parameter. The pair weight method is completely general. That is to say it works for models that include redshift distortions and nonlinearities (not included in this work). The pair weight method is able to accommodate any selection function that has no discrete jumps in the radial direction. The flexibility allows nearly any catalog to be analyzed using nearly any model.

However, the pair weight compression method is computationally expensive. If a classical method or brute force method will obtain the same results, then it is a good bet that the other method will be less computationally expensive. This means that for scales smaller than the smallest physical dimension in the catalog, classical methods will be the methods of choice. If a brute force analysis is possible, then on the largest scales the brute force technique will be the method of choice. There are situations, however, where neither brute force techniques nor classical methods will work.

#### 6.2 Las Campanas Redshift Survey

The Las Campanas Redshift Survey is the largest (in terms of number of galaxies) publicly available redshift catalog. It is divided into 327 observing regions. In each observing

window, redshifts could be taken for up to 112 galaxies (at the same time). This means that the observing windows were not necessarily complete to the same level. This means that the selection function varies from observing window to observing window. This causes problems for both the classical method and for brute force techniques. The smallest size scale for the catalog is then approximately  $7.5 h^{-1}$  Mpc. This means that classical methods will only be effective for the very smallest (and probably nonlinear) regions. The fact that the selection function is different from direction to direction makes the brute force route more tricky. This means that the number of “modes” that a brute force analysis can use is limited. This limits brute force techniques to only the largest scales as was done by Matsubara et al. (2000). LCRS, however, can be analyzed using the pair weight method. In fact, LCRS has exactly the kinds of problems that would force one to use the pair weight method.

### 6.2.1 Results

#### Redshift Space Power Spectrum

In Chapter 3, we presented the redshift power spectrum for LCRS. The power spectrum measured by the pair weight method was nearly a factor of two smaller than that measured by Lin et al. (1996b). In fact, the power spectrum agreed closely with the measurement of power from the PSCz (Sutherland et al., 1999) and the AMP measured by Gaztañaga and Baugh (1998). This came as quite a surprise since it is commonly thought that IRAS selected galaxies (PSCz) should be less biased, and thus have a lower galaxy-galaxy power spectrum, than optically selected galaxies (LCRS).

#### Real Space Power Spectrum

The real space power spectrum (Figure 5.1) is similar in shape and amplitude to the redshift space power spectrum. This implies that over the range of  $k \approx .1 - .3$  nonlinearities and redshift distortions are small. This would lead to a relatively low estimate of  $\beta$ . Unfortunately,

the power spectrum is noisier than one would like. This is due to the fact that the information is now shared amongst more parameters and that additional error is caused by miscalculation of off-diagonal Fisher matrix elements.

### **Measurement of $\beta$**

Using the correlated measurement of the real space power spectrum as the shape of the power spectrum, we did a least-squares analysis to find the best fit amplitude and value of  $\beta$ . The calculation yielded an amplitude that was essentially one. This means that the best fit amplitude is the measured amplitude, as expected. The measurement of  $\beta$  was  $\beta \approx .55^{+.35}_{-.30}$ . This result is consistent with, but slightly higher than, the calculation done by Matsubara et al. (2000). The hope was that evidence of nonlinearities would arise in this calculation. However, the measurement of the smallest scales was too contaminated by aliasing of the Fourier transform to see anything conclusive about nonlinearities.

**Possible Improvements** With the analysis of the LCRS being the first analysis using the pair weight method, we made a number of decisions that turned out not to be ideal.

In the analysis, we used linearly spaced bins. This is not the best way to go. Logarithmically spaced bins have two clear advantages. The first is that by using the same number of bins one could cover a much larger portion of the spectrum. The second is that the calculations of the covariances involving large separations take much longer than calculations which only involve small separations. Using logarithmicly spaced bins places a higher fraction of the bins at small scales and thereby speeds the calculations.

The second place for improvement is in the calculation of the selection function. The selection function routine for this analysis fit the selection function to incomplete gamma functions. It is clear that incomplete gamma functions are not the fastest way to go. This causes the longest calculations to take several times as long to complete.

### 6.3 Conclusion

The pair weight compression method is a viable method of extracting power spectra and  $\beta$  from galaxy catalogs. The flexibility in the pair weight method allows for even a catalog with the selection properties of LCRS to be analyzed. The value of  $\beta \approx .55^{+.35}_{-.30}$  extracted from LCRS is consistent with the analysis done by Matsubara et al. (2000). Unlike the analysis of Matsubara et al. (2000) the pair weight method was able to extract a real space power spectrum.

The pair weight method is successful in the analysis of the most complicated of catalogs. It is clear, however, that when the catalog is simpler then the computational overhead of the pair weight method makes other methods more attractive. For large angle catalogs with uniform depth, classical methods and brute force methods will be the methods of choice.

## Bibliography

- Ballinger, W. E., Heavens, A. F., and Taylor, A. N. 1995, MNRAS **276**, L59
- Binggeli, B., Sandage, A., and Tammann, G. A. 1988, ARA&A **26**, 509
- Bond, J. R., Efstathiou, G., and Tegmark, M. 1997, MNRAS **291**, L33, (astro-ph/9702100)
- Bond, J. R., Jaffe, A. H., and Knox, L. 1998, Phys. Rev. D **57**, 2117
- Bromley, B. C. 1994, ApJ **437**, 541
- Carlberg, R. G., Yee, H. K. C., Lin, H., Sawicki, M., Shepherd, C. W., Ellingson, E., Morris, S. L., Schade, D., Hesser, J. E., Hutchings, J. B., Oke, J. B., Patton, D., Wirth, G., Balogh, M., Hartwick, F. D. A., Pritchet, C. J., Abraham, R., and Smecker-Hane, T. 1998, poco.work p. 119
- Carroll, S. M., Press, W. H., and Turner, E. L. 1992, AARA **30**
- Cole, S., Fisher, K. B., and Weinberg, D. H. 1994, MNRAS **267**, 785
- Cole, S., Fisher, K. B., and Weinberg, D. H. 1995, MNRAS **275**, 515
- Cole, S., Hatton, S., Weinberg, D. H., and Frenk, C. S. 1988, MNRAS **300**, 945, (astro-ph/9801250)
- de Bernardis, P., Ade, P. A. R., Bock, J. J., Bond, J. R., Borrill, J., Boscaleri, A., Coble, K., Crill, B. P., Gasperis, G. D., Farese, P. C., Ferreira, P. G., Ganga, K., Glacometti, M., Hivon, E., Hristov, V. V., Iacoangeli, A., Jaffe, A. H., Lange, A. E., Martinis, L., Masi, S., Mason, P. V., Mauskopf, P. D., Melchiorri, A., Miglio, L., Montroy, T., Netterfield, C. B., Pascale, E., Piacentini, F., Pogosyan, D., Prunet, S., Rau, S., Romeo, G., Ruhl, J. E., Scaramuzzi, F., Sforna, D., and Vittorio, N. 2000, Science **404**, 955
- Donahue, M., Voit, G. M., Gioia, I., Lupino, G., Hughes, J. P., and Stocke, J. T. 1998, ApJ **502**, 550
- Eisenstein, D. J. and Hu, W. 1999, ApJ **511**, 5
- Eisenstein, D. J., Hu, W., and Tegmark, M. 1999, ApJ **518**, 2, (astro-ph/9807130)
- Feldman, H. A., Kaiser, N., and Peacock, J. A. 1994, ApJ **426**, 23
- Fisher, K. B., Davis, M., Strauss, M. A., Yahil, A., and Huchra, J. P. 1993, ApJ **402**, 42

- Fisher, K. B., Scharf, C. A., and Lahov, O. 1994, MNRAS **266**, 219
- Fry, J. N. and Gaztañaga, E. 1994, ApJ **425**, 1
- Gaztañaga, E. and Baugh, C. M. 1998, MNRAS **294**, 229
- Gramann, M., Cen, R., and III, J. R. G. 1994, ApJ **425**, 382
- Hamilton, A. J. S. 1992, ApJ **385**, L5
- Hamilton, A. J. S. 1993, ApJ **417**, 19
- Hamilton, A. J. S. 1997, MNRAS **289**, 285
- Hamilton, A. J. S. 2000, MNRAS **312**, 257, (astro-ph/9907102)
- Hamilton, A. J. S. and Culhane, M. 1996, MNRAS **278**, 73
- Hamilton, A. J. S. and Tegmark, M. 2000a, ApJ **312**, 285
- Hamilton, A. J. S. and Tegmark, M. 2000b, MNRAS **312**, 285, (astro-ph/9905192)
- Hamilton, A. J. S., Tegmark, M., and Padmanabhan, N. 2000, MNRAS, (astro-ph/0004334)
- Heavens, A. F., Jimenez, R., and Lahov, O. 2000, MNRAS, (astro-ph/99111102)
- Heavens, A. F. and Taylor, A. N. 1995, MNRAS **275**, 483
- Hu, W. 2000, in Birth and Evolution of the Universe, RESCEU 1999, (astro-ph/0002520)
- Kaiser, N. 1987, MNRAS **227**, 1
- Lahav, O., Rees, M. J., Lilje, P. B., and Primack, J. R. 1991, MNRAS **251**, 128
- Lilje, P. B. and Efstathiou, G. 1989, MNRAS **236**, 851
- Lin, H., Kirshner, R. P., Sheckman, S. A., Landy, S. D., Oemler, A., Tucker, D. L., and Schechter, P. L. 1996a, ApJ **464**, 60, (astro-ph/9602064)
- Lin, H., Kirshner, R. P., Sheckman, S. A., Landy, S. D., Oemler, A., Tucker, D. L., and Schechter, P. L. 1996b, ApJ **471**, 617, (astro-ph/9606055)
- Loveday, J., Efstathiou, G., Peterson, B. A., and Maddox, S. J. 1992, ApJ **400**, 43
- Maddox, S. J., Efstathiou, G., Sutherland, W. J., and Loveday, J. 1990, MNRAS **242**, 43P
- Mann, R. G., Peacock, J. A., and Heavens, A. F. 1998, MNRAS **293**, 209, (astro-ph/9708031)
- Matsubara, T., Szalay, A. S., and Landy, S. D. 2000, ApJ **535**, 1L, (astro-ph/9911151)
- McGill, C. 1990, MNRAS **242**, 428
- Padmanabhan, N., Tegmark, M., and Hamilton, A. J. S. 2000, ApJ, submitted (astro-ph/9911421)
- Peacock, J. A. 1997, MNRAS **284**, 885

- Perlmutter, S., Aldering, G., Goldhaber, G., Knop, R. A., Nugent, P., Castro, P. G., Deustua, S., Fabbro, S., Goobar, A., Groom, D. E., Hook, I. M., Kim, A. G., Kim, M. Y., Lee, J. C., Nunes, N. J., Pain, R., Pennypacker, C. R., Quimby, R., Lidman, C., Ellis, R. S., Irwin, M., McMahon, R. G., Ruiz-Lapuente, P., Walton, N., Schaefer, B., Boyle, B. J., Filippenko, A. V., Matheson, T., Fruchter, A. S., Panagia, N., Newbert, H. J. M., and Couch, W. J. 1999, ApJ **517**, 565
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. 1986, Numerical Recipes in Fortran 77, Cambridge University Press
- Riess, A. G. 2000, PASP, To Appear as an Invited Review, (astro-ph/0005229)
- Riess, A. G., Filippenko, A. V., Challis, P., Clocchiatti, A., Diercks, A., Garnavich, P. M., Gilliland, R. L., Hogan, C. J., Jha, S., Kirshner, R. P., Leibundgut, B., Phillips, M. M., Reiss, D., Schmidt, B. P., Schommer, R. A., Smith, R. C., Spyromilio, J., Stubbs, C., Suntzeff, N. B., and Tonry, J. 1998, AJ **116**, 1009
- Scherrer, R. J. and Weinberg, D. H. 1998, ApJ **504**, 607, (astro-ph/9712192)
- Shectman, S. A., Landy, S. D., Oemler, A., Lin, H., Kirshner, R. P., and Schechter, P. L. 1996, ApJS **470**, 172, (astro-ph/9604167)
- Sutherland, W., Tadros, H., Efstathiou, G., Frenk, C. S., Keeble, O., Maddox, S., McMahon, T. G., Oliver, S., Rowan-Robinson, M., Saunders, W., and White, S. D. M. 1999, MNRAS **308**, 289
- Tadros, H., Ballinger, W. E., Taylor, A. N., Heavens, A. F., Efstathiou, G., Saunders, W., Frenk, C. S., Keeble, O., McMahon, R., Maddox, S. J., Oliver, S., Rowan-Robinson, M., Sutherland, W. J., and White, S. D. M. 1999, MNRAS **305**, 527, (astro-ph/9901351)
- Tadros, H. and Efstathiou, G. 1995, MNRAS **276**, 45, (astro-ph/9507050)
- Tegmark, M., Eisenstein, D. J., Hu, W., and de Oliveira-Costa, A. 2000, ApJ **530**, 133, (astro-ph/9905257)
- Tegmark, M., Hamilton, A. J. S., Strauss, M., Vogeley, M., and Szalay, A. 1998, ApJ **499**, 555
- Tegmark, M., Taylor, A., and Heavens, A. F. 1997, ApJ **480**, 22
- Tresse, L. 1999, in O. L. Fèvre and S. Charlot (eds.), Formation and Evolution of Galaxies, Les Houches School Series, Springer-Verlag, (astro-ph/9902209)
- Tytler, D., O'Meara, J. M., Suzuki, N., and Lubin, D. 2000, Phys. Scr., to appear, (astro-ph/0001318)
- Vogeley, M. S. and Szalay, A. S. 1996, ApJ **465**, 34
- Willmer, C. N. A. 1997, AJ **114**, 898

## Appendix A

### Calculating the Shape Functions

Equation 4.5 shows the relationship between the redshift space correlation function ( $\xi^s$ ) and its real space counterpart ( $\xi^r$ ). For current galaxy catalogs (and those in preparation)  $\xi^r$  is not directly available. Therefore we must find a way to extract  $\beta$  from the properties of  $\xi^s$ . Hamilton (1993) did this by weighting the calculations of  $\xi$  by the Legendre polynomials of  $\mu$ . To find the equivalent functions for the spherically symmetric case, expand  $\xi^s$  and  $\xi^r$  into sets of orthogonal components. This will allow the general functions  $\xi$  to be replaced by specific functions. Then the distortion operator will operate on a set of specific functions. The results from the distortion operation can then be added together to yield a final result for the general  $\xi$ .

One can expand in any complete set of functions but the best choice is to find a set which makes the calculations simplest. The operator  $\partial/\partial \ln r_{12} |_{\Delta}$  commutes with all of the individual operators contained within the distortion operator ( $\alpha$  is the only component of the distortion operator which does not necessarily commute with  $\partial/\partial \ln r_{12} |_{\Delta}$ ). Here  $\Delta$  is the shape of the triangle connecting the observer to the two galaxies in question. Therefore the logical choice for the expansion is the set of eigenfunctions of  $\partial/\partial \ln r_{12} |_{\Delta}$ :

$$r_{12}^{-\gamma - i\omega_{12}}. \tag{A.1}$$

Therefore, define the quantities  $\xi^r(\omega_{12})$  and  $\xi^s(\omega_{12})$ :

$$\begin{aligned} \xi^s(\omega_{12}, \Delta) &= \frac{1}{2\pi} \int_0^\infty \xi^s(r_{12}, \Delta) r_{12}^{\gamma + i\omega_{12}} dr_{12}/r_{12} |_{\Delta}, \\ \xi^r(\omega_{12}) &= \frac{1}{2\pi} \int_0^\infty \xi^r(r_{12}) r_{12}^{\gamma + i\omega_{12}} dr_{12}/r_{12} |_{\Delta} \end{aligned} \tag{A.2}$$



with the inverse transforms:

$$\begin{aligned}\xi^s(r_{12}, \Delta) &= \int_{-\infty}^{\infty} \xi^s(\omega_{12}, \Delta) r_{12}^{-\gamma-i\omega_{12}} d\omega_{12}, \\ \xi^r(r_{12}) &= \int_{-\infty}^{\infty} \xi^r(\omega_{12}) r_{12}^{-\gamma-i\omega_{12}} d\omega_{12}.\end{aligned}\quad (\text{A.3})$$

Now applying the distortion operator on  $\xi^r(r_{12})$  actually applies it to the  $r_{12}^{-\gamma-i\omega}$  inside the integral. This gives:

$$\begin{aligned}\xi^s(r_{12}) &= \int_{-\infty}^{\infty} \xi^r(\omega_{12}) D_{\text{dist}} \left[ r_{12}^{-\gamma-i\omega_{12}} \right] d\omega_{12} = \\ &= \int_{-\infty}^{\infty} \xi^r(\omega_{12}) \left[ 1 + A_1(\eta_{12}, \Delta)\beta + A_2(\eta_{12}, \Delta)\beta^2 \right] r_{12}^{-\gamma-i\omega_{12}} d\omega_{12}.\end{aligned}\quad (\text{A.4})$$

Let  $\eta = \gamma + i\omega$ . Now calculate the actual values of each  $A$ .

To calculate the values of  $A$  operate with the distortion operator on  $r^{-\eta}$ . First of all, note that the operators with subscripts 1 commute with those with subscripts 2. So

$$A_1(\eta_{12}, \Delta)r^{-\eta} = \left[ \left( \frac{\partial^2}{\partial r_1^2} + \frac{\alpha(r_1)}{r_1} \frac{\partial}{\partial r_1} \right) \nabla_1^{-2} + \left( \frac{\partial^2}{\partial r_2^2} + \frac{\alpha(r_2)}{r_2} \frac{\partial}{\partial r_2} \right) \nabla_2^{-2} \right] r_{12}^{-\eta} \quad (\text{A.5})$$

and

$$A_2(\eta_{12}, \Delta)r^{-\eta} = \left[ \left( \frac{\partial^2}{\partial r_1^2} + \frac{\alpha(r_1)}{r_1} \frac{\partial}{\partial r_1} \right) \left( \frac{\partial^2}{\partial r_2^2} + \frac{\alpha(r_2)}{r_2} \frac{\partial}{\partial r_2} \right) \nabla_1^{-2} \nabla_2^{-2} \right] r_{12}^{-\eta}.\quad (\text{A.6})$$

The first job here is to operate with  $\nabla_i^{-2}$  on  $r_{12}^{-\eta}$ . Because  $r_{12}$  is a spherically symmetric quantity:

$$\nabla_1^{-2} r_{12}^{-\eta} = \nabla_2^{-2} r_{12}^{-\eta} = \left[ \frac{1}{r_{12}^2} \frac{\partial}{\partial r_{12}} r_{12}^2 \frac{\partial}{\partial r_{12}} \right]^{-1} r_{12}^{-\eta}.\quad (\text{A.7})$$

Set this equal to a new variable  $\phi$  and solve for  $\phi$ :

$$\left[ \frac{1}{r_{12}^2} \frac{\partial}{\partial r_{12}} r_{12}^2 \frac{\partial}{\partial r_{12}} \right] \phi = r_{12}^{-\eta} \quad (\text{A.8})$$

becomes

$$\left[ \partial \left( r_{12}^2 \frac{\partial}{\partial r_{12}} \right) \right] \phi = r_{12}^{2-\eta} \partial r_{12}.\quad (\text{A.9})$$

After performing one integration

$$\left[ r_{12}^2 \frac{\partial}{\partial r_{12}} \right] \phi = \frac{r_{12}^{3-\eta}}{3-\eta}. \quad (\text{A.10})$$

or

$$\partial \phi = \frac{r_{12}^{1-\eta}}{3-\eta} \partial r_{12}. \quad (\text{A.11})$$

Performing this integration yields

$$\phi = \frac{r_{12}^{2-\eta}}{(3-\eta)(2-\eta)}. \quad (\text{A.12})$$

Note that this equation serves for any  $\eta$  provided that  $\eta$  does not equal 2 or 3. This means that

$$\nabla_1^{-2} r_{12}^{-\eta} = \nabla_2^{-2} r_{12}^{-\eta} = \frac{r_{12}^{2-\eta}}{(3-\eta)(2-\eta)} \quad (\text{A.13})$$

and by considering  $-\eta' = 2 - \eta$ :

$$\nabla_1^{-2} \nabla_2^{-2} r_{12}^{-\eta} = \frac{r_{12}^{4-\eta}}{(5-\eta)(4-\eta)(3-\eta)(2-\eta)}. \quad (\text{A.14})$$

Now the values for  $A$  become

$$A_1(\eta_{12}, \Delta) r^{-\eta} = \left[ \left( \frac{\partial^2}{\partial r_1^2} + \frac{\alpha(r_1)}{r_1} \frac{\partial}{\partial r_1} \right) + \left( \frac{\partial^2}{\partial r_2^2} + \frac{\alpha(r_2)}{r_2} \frac{\partial}{\partial r_2} \right) \right] \frac{r_{12}^{2-\eta}}{(3-\eta)(2-\eta)} \quad (\text{A.15})$$

and

$$A_2(\eta_{12}, \Delta) r^{-\eta} = \left[ \left( \frac{\partial^2}{\partial r_1^2} + \frac{\alpha(r_1)}{r_1} \frac{\partial}{\partial r_1} \right) \left( \frac{\partial^2}{\partial r_2^2} + \frac{\alpha(r_2)}{r_2} \frac{\partial}{\partial r_2} \right) \right] \frac{r_{12}^{4-\eta}}{(5-\eta)(4-\eta)(3-\eta)(2-\eta)}. \quad (\text{A.16})$$

To calculate the results for  $A_1$  and  $A_2$  start by doing some preliminary calculations. First define the cosines of the angles in the triangle:

$$\mu_1 = \frac{r_{12}^2 + r_1^2 - r_2^2}{2r_{12}r_1}, \mu_2 = \frac{r_{12}^2 + r_2^2 - r_1^2}{2r_{12}r_2}, \mu_{12} = \frac{r_1^2 + r_2^2 - r_{12}^2}{2r_1r_2}. \quad (\text{A.17})$$

This means that

$$r_{12} = [r_1^2 + r_2^2 - 2r_1r_2\mu_{12}]^{\frac{1}{2}}. \quad (\text{A.18})$$

So,

$$\frac{\partial}{\partial r_1} r_{12}^n = n r_{12}^{n-2} (r_1 - r_2 \mu_{12}) \quad (\text{A.19})$$

and

$$\frac{\partial}{\partial r_2} r_{12}^n = n r_{12}^{n-2} (r_2 - r_1 \mu_{12}). \quad (\text{A.20})$$

Simplify the terms in parentheses by noting

$$r_1 - r_2 \mu_{12} = \frac{2r_1^2 r_2 - r_2 (r_1^2 + r_2^2 - r_{12}^2)}{2r_1 r_2} = \frac{r_1^2 - r_2^2 + r_{12}^2}{2r_1} = \mu_1 r_{12}, \quad (\text{A.21})$$

and similarly

$$r_2 - r_1 \mu_{12} = \mu_2 r_{12}. \quad (\text{A.22})$$

Another useful equation is

$$2\mu_1 \mu_2 \mu_{12} + \mu_1^2 + \mu_2^2 + \mu_{12}^2 = 1. \quad (\text{A.23})$$

Now divide  $A_1$  into two pieces. The first piece has  $\alpha$  terms only:

$$\left[ \frac{\alpha(r_1)}{r_1} \frac{\partial}{\partial r_1} + \frac{\alpha(r_2)}{r_2} \frac{\partial}{\partial r_2} \right] \frac{r_{12}^{2-\eta}}{(3-\eta)(2-\eta)}. \quad (\text{A.24})$$

Differentiating gives:

$$\left[ \frac{\alpha(r_1)}{r_1} r_{12} \mu_1 + \frac{\alpha(r_2)}{r_2} r_{12} \mu_2 \right] \frac{r_{12}^{-\eta}}{3-\eta}. \quad (\text{A.25})$$

The second term has no terms with  $\alpha$ :

$$\left[ \frac{\partial^2}{\partial r_1^2} + \frac{\partial^2}{\partial r_2^2} \right] \frac{r_{12}^{2-\eta}}{(3-\eta)(2-\eta)} = [1 - \eta \mu_1^2 + 1 - \eta \mu_2^2] \frac{r_{12}^{-\eta}}{3-\eta}. \quad (\text{A.26})$$

Later in the calculation, all factors of  $\eta$  will need to be in the denominator. So by noticing that

$$\frac{-\eta}{3-\eta} = 1 - \frac{3}{3-\eta} \quad (\text{A.27})$$

$A_1$  can be written:

$$A_1(\eta, \Delta)r_{12}^{-\eta} = \left[ \frac{\frac{\alpha(r_1)}{r_1}r_{12}\mu_1 + \frac{\alpha(r_2)}{r_2}r_{12}\mu_2 + 2 - 3\mu_1^2 - 3\mu_2^2}{3 - \eta} + \mu_1^2 + \mu_2^2 \right] r_{12}^{-\eta}. \quad (\text{A.28})$$

To attack the differentiation necessary for  $A_2$  first notice that it can be rewritten:

$$A_2(\eta, \Delta)r_{12}^{-\eta} = \left[ \left( \frac{\partial}{\partial r_1} + \frac{\alpha(r_1)}{r_1} \right) \left( \frac{\partial}{\partial r_2} + \frac{\alpha(r_2)}{r_2} \right) \frac{\partial}{\partial r_1} \frac{\partial}{\partial r_2} \right] \frac{r_{12}^{4-\eta}}{(5-\eta)(4-\eta)(3-\eta)(2-\eta)}. \quad (\text{A.29})$$

First differentiating with respect to  $r_2$  yields:

$$A_2(\eta, \Delta)r_{12}^{-\eta} = \left[ \left( \frac{\partial}{\partial r_1} + \frac{\alpha(r_1)}{r_1} \right) \left( \frac{\partial}{\partial r_2} + \frac{\alpha(r_2)}{r_2} \right) \frac{\partial}{\partial r_1} \right] \frac{(r_2 - r_1\mu_{12})r_{12}^{2-\eta}}{(5-\eta)(3-\eta)(2-\eta)}. \quad (\text{A.30})$$

After the second differentiation:

$$A_2(\eta, \Delta)r_{12}^{-\eta} = \left[ \left( \frac{\partial}{\partial r_1} + \frac{\alpha(r_1)}{r_1} \right) \left( \frac{\partial}{\partial r_2} + \frac{\alpha(r_2)}{r_2} \right) \right] \left[ \frac{-\mu_{12}r_{12}^{2-\eta}}{(5-\eta)(3-\eta)(2-\eta)} + \frac{(r_1 - r_2\mu_{12})(r_2 - r_1\mu_{12})r_{12}^{-\eta}}{(5-\eta)(3-\eta)} \right]. \quad (\text{A.31})$$

Once again separate out the different terms and calculate them separately to obtain the simplest looking results. The first term is the term with two factors of  $\alpha$ :

$$\frac{\alpha(r_1)}{r_1} \frac{\alpha(r_2)}{r_2} \left[ \frac{-\mu_{12}r_{12}^{2-\eta}}{(5-\eta)(3-\eta)(2-\eta)} + \frac{\mu_1\mu_2r_{12}^{2-\eta}}{(5-\eta)(3-\eta)} \right]. \quad (\text{A.32})$$

The second term has one factor of  $\alpha$ :

$$\left[ \frac{\alpha(r_1)}{r_1} \frac{-2\mu_{12}\mu_2r_{12} + \mu_1r_{12} - \eta\mu_1\mu_2^2r_{12}}{(5-\eta)(3-\eta)} + \frac{\alpha(r_2)}{r_2} \frac{-2\mu_{12}\mu_1r_{12} + \mu_2r_{12} - \eta\mu_1^2\mu_2r_{12}}{(5-\eta)(3-\eta)} \right] r_{12}^{-\eta}. \quad (\text{A.33})$$

The final term contains no factors of  $\alpha$ :

$$\begin{aligned} \frac{\partial}{\partial r_1} \left[ \frac{-2\mu_{12}(r_2 - r_1\mu_{12})r_{12}^{-\eta}}{(5-\eta)(3-\eta)} + \frac{(r_1 - r_2\mu_{12})r_{12}^{-\eta}}{(5-\eta)(3-\eta)} \right. \\ \left. - \frac{\eta(r_1 - r_2\mu_{12})(r_2 - r_1\mu_{12})^2r_{12}^{-2-\eta}}{(5-\eta)(3-\eta)} \right] = \\ \frac{2\mu_{12}^2r_{12}^{-\eta}}{(5-\eta)(3-\eta)} + \frac{4\eta\mu_{12}\mu_1\mu_2r_{12}^{-\eta}}{(5-\eta)(3-\eta)} + \frac{r_{12}^{-\eta}}{(5-\eta)(3-\eta)} + \\ \frac{-\eta\mu_1^2r_{12}^{-\eta}}{(5-\eta)(3-\eta)} + \frac{-\eta\mu_2^2r_{12}^{-\eta}}{(5-\eta)(3-\eta)} + \frac{\eta(2+\eta)\mu_1^2\mu_2^2r_{12}^{-\eta}}{(5-\eta)(3-\eta)}. \quad (\text{A.34}) \end{aligned}$$

Using equation A.23 rewrite this term:

$$\begin{aligned}
& \frac{2\mu_{12}^2 r_{12}^{-\eta}}{(5-\eta)(3-\eta)} + \frac{-2\eta(\mu_{12}^2 + \mu_1^2 + \mu_2^2 - 1)r_{12}^{-\eta}}{(5-\eta)(3-\eta)} + \frac{r_{12}^{-\eta}}{(5-\eta)(3-\eta)} + \\
& \frac{-\eta\mu_1^2 r_{12}^{-\eta}}{(5-\eta)(3-\eta)} + \frac{-\eta\mu_2^2 r_{12}^{-\eta}}{(5-\eta)(3-\eta)} + \frac{\eta(2+\eta)\mu_1^2 \mu_2^2 r_{12}^{-\eta}}{(5-\eta)(3-\eta)} = \\
& \frac{2\mu_{12}^2(1-\eta)r_{12}^{-\eta}}{(5-\eta)(3-\eta)} + \frac{(1+2\eta)r_{12}^{-\eta}}{(5-\eta)(3-\eta)} + \\
& \frac{-3\eta\mu_1^2 r_{12}^{-\eta}}{(5-\eta)(3-\eta)} + \frac{-3\eta\mu_2^2 r_{12}^{-\eta}}{(5-\eta)(3-\eta)} + \frac{\eta(2+\eta)\mu_1^2 \mu_2^2 r_{12}^{-\eta}}{(5-\eta)(3-\eta)}. \quad (\text{A.35})
\end{aligned}$$

Each  $A_i(\eta_{12}, \Delta)$  can be expressed in functions of  $r_{12}, r_1$  and  $r_2$  times functions of  $\frac{1}{n-\eta_{12}}$  where  $n = 2, 3, 5$ . The factors of  $\beta$  and the  $r$ 's in equation (A.4) can be brought out of integral over  $\omega$  to become:

$$\xi^s(r_{12}) = \sum_j g_j(\beta) B_j(r_{12}, r_1, r_2) \int_{-\infty}^{\infty} \frac{\xi^r(\omega_{12}) r_{12}^{-\eta_{12}}}{n_j - \eta_{12}} d\omega_{12}. \quad (\text{A.36})$$

Rearranging the integral in this equation:

$$\int_{-\infty}^{\infty} \frac{\xi^r(\omega_{12}) r_{12}^{-\eta_{12}}}{n - \eta_{12}} d\omega_{12} = \int_{-\infty}^{\infty} \xi^r(\omega_{12}) \left[ \frac{1}{r_{12}^n} \int_0^{r_{12}} r^{n-\eta_{12}-1} dr \right] d\omega_{12}, \quad (\text{A.37})$$

and switching the order of integration:

$$\frac{1}{r_{12}^n} \int_0^{r_{12}} r^{n-1} \left[ \int_{-\infty}^{\infty} \xi^r(\omega) r^{-\eta_{12}} \right] dr = \frac{1}{r_{12}^n} \int_0^{r_{12}} r^{n-1} \xi^r(r) dr \quad (\text{A.38})$$

So the integral can be expressed in an average over  $\xi$ . Let:

$$\xi^s(r_{12}) = \sum_j g_j(\beta) B_j(r_{12}, r_1, r_2) \Xi_j(\xi(r_{12}), \check{\xi}(r_{12}), \bar{\xi}(r_{12}), \bar{\bar{\xi}}(r_{12})) \quad (\text{A.39})$$

with

$$\begin{aligned}
\check{\xi} &\equiv \frac{2}{r_{12}^2} \int_0^{r_{12}} \xi(r) r dr, \\
\bar{\xi} &\equiv \frac{3}{r_{12}^3} \int_0^{r_{12}} \xi(r) r^2 dr, \\
\bar{\bar{\xi}} &\equiv \frac{5}{r_{12}^5} \int_0^{r_{12}} \xi(r) r^4 dr.
\end{aligned} \quad (\text{A.40})$$

Note that in equation A.39 the functions  $B_j(r_i, r_j, r_{ij})$ ,  $g_j(\beta)$  and  $f_j$  are not unique. Choose any set of functions which satisfy equation A.39. Therefore, choose the functions to make the rest of the analysis as simple and accurate as possible.

In order to obtain results that tend towards the distant observer approximation, choose three of the shape functions as:

$$B_0 \equiv 1, \quad (\text{A.41})$$

$$B_2 \equiv \frac{1}{2} [P_2(\mu_1) + P_2(\mu_2)], \quad (\text{A.42})$$

and

$$B_4 \equiv \frac{1}{8} (35\mu_1^2\mu_2^2 - 15\mu_1^2 - 15\mu_2^2 + 3) \quad (\text{A.43})$$

where

$$P_2 = \frac{3}{2}\mu^2 - \frac{1}{2}. \quad (\text{A.44})$$

Define  $\check{\xi}_i$  as the function  $g_i(\beta)\Xi_i(\xi(r_{12}), \check{\xi}(r_{12}), \bar{\xi}(r_{12}), \bar{\bar{\xi}}(r_{12}))$  which corresponds to shape function  $B_i$ . This means that the term with no factors of  $\beta$  contributes  $\xi(r_{12})$  to  $\check{\xi}_0$ . The terms given by  $A_1$  contribute  $\frac{4}{3}\beta(\xi(r_{12}) - \bar{\xi}(r_{12}))$  to  $\check{\xi}_2$  and  $\frac{2}{3}\beta\xi(r_{12})$  to  $\check{\xi}_0$  leaving

$$\frac{\beta}{3}\bar{\xi}(r_{12}) \left[ \frac{\alpha(r_1)}{r_1} r_{12}\mu_1 + \frac{\alpha(r_2)}{r_2} r_{12}\mu_2 \right]. \quad (\text{A.45})$$

So, define

$$B_1 \equiv \left[ \frac{\alpha(r_1)}{r_1} r_{12}\mu_1 + \frac{\alpha(r_2)}{r_2} r_{12}\mu_2 \right]. \quad (\text{A.46})$$

This means that the terms of  $A_1$  also contributes  $\frac{\beta}{3}\bar{\xi}(r_{12})$  to  $\check{\xi}_1$ .

The terms in equation A.35 will contribute to  $\check{\xi}_4$ ,  $\check{\xi}_2$ , and  $\check{\xi}_0$ . To find the exact contributions, replace  $\mu_1^2\mu_2^2$

$$\mu_1^2\mu_2^2 = \frac{8}{35}B_4 + \frac{3}{7}(\mu_1^2 + \mu_2^2) - \frac{3}{35} \quad (\text{A.47})$$

to get

$$\begin{aligned} \frac{r_{12}^{-\eta}}{(5-\eta)(3-\eta)} [2\mu_{12}^2(1-\eta) + (1+2\eta) - 3\eta(\mu_1^2 + \mu_2^2) + \\ \eta(2+\eta)\left(\frac{8}{35}B_4 + \frac{3}{7}(\mu_1^2 + \mu_2^2) - \frac{3}{35}\right)] = \\ \frac{r_{12}^{-\eta}}{(5-\eta)(3-\eta)} \left[ 2\mu_{12}^2(1-\eta) + \left(1 + \frac{64}{35}\eta - \frac{3}{35}\eta^2\right) + \right. \\ \left. (\mu_1^2 + \mu_2^2)\left(\frac{3}{7}\eta^2 - \frac{15}{7}\eta\right) + \eta(2+\eta)\frac{8}{35}B_4 \right]. \quad (\text{A.48}) \end{aligned}$$

Now replace  $\mu_1^2 + \mu_2^2$  using

$$\mu_1^2 + \mu_2^2 = \frac{4}{3}B_2 + \frac{2}{3} \quad (\text{A.49})$$

to get

$$\begin{aligned} \frac{r_{12}^{-\eta}}{(5-\eta)(3-\eta)} \left[ 2\mu_{12}^2(1-\eta) + \left(1 + \frac{64}{35}\eta - \frac{3}{35}\eta^2\right) + \left(\frac{4}{3}B_2 + \frac{2}{3}\right)\left(\frac{3}{7}\eta^2 - \frac{15}{7}\eta\right) + \right. \\ \left. \eta(2+\eta)\frac{8}{35}B_4 \right] = \\ \frac{r_{12}^{-\eta}}{(5-\eta)(3-\eta)} \left[ 2\mu_{12}^2(1-\eta) + \left(1 + \frac{2}{5}\eta + \frac{1}{5}\eta^2\right) + \right. \\ \left. \left(\frac{3}{7}\eta^2 - \frac{15}{7}\eta\right)\frac{4}{3}B_2 + \eta(2+\eta)\frac{8}{35}B_4 \right]. \quad (\text{A.50}) \end{aligned}$$

Before continuing, note a few relations:

$$\frac{\eta(2+\eta)}{(5-\eta)(3-\eta)} = 1 + \frac{15}{2} \frac{1}{3-\eta} - \frac{35}{2} \frac{1}{5-\eta} \implies \xi(r_{12}) + \frac{5}{2}\bar{\xi}(r_{12}) - \frac{7}{2}\bar{\bar{\xi}}(r_{12}), \quad (\text{A.51})$$

$$\frac{\frac{3}{7}\eta^2 - \frac{15}{7}\eta}{(5-\eta)(3-\eta)} = \frac{3}{7} - \frac{9}{7} \frac{1}{3-\eta} \implies \frac{3}{7}\xi(r_{12}) - \frac{3}{7}\bar{\xi}(r_{12}), \quad (\text{A.52})$$

$$\frac{1 + \frac{2}{5}\eta + \frac{1}{5}\eta^2}{(5-\eta)(3-\eta)} = \frac{1}{5} + \frac{2}{3-\eta} - \frac{4}{5-\eta} \implies \frac{1}{5}\xi(r_{12}) + \frac{2}{3}\bar{\xi}(r_{12}) - \frac{4}{5}\bar{\bar{\xi}}(r_{12}), \quad (\text{A.53})$$

and

$$\frac{1-\eta}{(5-\eta)(3-\eta)} = -\frac{1}{3-\eta} + \frac{2}{5-\eta} \implies -\frac{1}{3}\bar{\xi}(r_{12}) + \frac{2}{5}\bar{\bar{\xi}}(r_{12}). \quad (\text{A.54})$$

The terms of  $\beta^2$  which contain factors of  $\alpha$  cannot contribute to the values of  $\check{\xi}_0$ ,  $\check{\xi}_2$ , or  $\check{\xi}_4$  because the corresponding shape functions contain no factors of  $\alpha$ . This means that

$$\check{\xi}_0 = \left[ 1 + \frac{2}{3}\beta + \frac{1}{5}\beta^2 \right] \xi(r_{12}), \quad (\text{A.55})$$

$$\check{\xi}_2 = \left[ \frac{4}{3}\beta + \frac{4}{7}\beta^2 \right] [\xi(r_{12}) - \bar{\xi}(r_{12})], \quad (\text{A.56})$$

and

$$\check{\xi}_4 = \frac{8}{35}\beta^2 \left[ \xi(r_{12}) + \frac{5}{2}\bar{\xi}(r_{12}) - \frac{7}{2}\bar{\bar{\xi}}(r_{12}) \right]. \quad (\text{A.57})$$

This leaves the terms

$$2\beta^2 \left[ -\frac{1}{3}\bar{\xi}(r_{12}) + \frac{2}{5}\bar{\bar{\xi}}(r_{12}) \right] \mu_{12}^2 + \beta^2 \left[ \frac{2}{3}\bar{\xi}(r_{12}) - \frac{4}{5}\bar{\bar{\xi}}(r_{12}) \right] \quad (\text{A.58})$$

and all of the terms with factors of  $\alpha$  for some other shape function. Reduce the terms containing factors of  $\alpha$  with

$$\frac{1}{(5-\eta)(3-\eta)(2-\eta)} = \frac{1}{3} \frac{1}{2-\eta} - \frac{1}{2} \frac{1}{3-\eta} + \frac{1}{6} \frac{1}{5-\eta} \implies \frac{1}{6}\check{\xi}(r_{12}) - \frac{1}{6}\bar{\xi}(r_{12}) + \frac{1}{30}\bar{\bar{\xi}}(r_{12}), \quad (\text{A.59})$$

$$\frac{1}{(5-\eta)(3-\eta)} = \frac{1}{2} \frac{1}{3-\eta} - \frac{1}{2} \frac{1}{5-\eta} \implies \frac{1}{6}\bar{\xi}(r_{12}) - \frac{1}{10}\bar{\bar{\xi}}(r_{12}), \quad (\text{A.60})$$

and

$$\frac{\eta}{(5-\eta)(3-\eta)} = \frac{3}{2} \frac{1}{3-\eta} - \frac{5}{2} \frac{1}{5-\eta} \implies \frac{1}{2}\bar{\xi}(r_{12}) - \frac{1}{2}\bar{\bar{\xi}}(r_{12}). \quad (\text{A.61})$$

Notice that the remaining terms contain no factors of  $\xi(r_{12})$  and that all of the remaining terms have a factor of  $\beta^2$ . The easiest solution to capturing all of the remaining terms is to define

$$\check{\xi}_3 = \beta^2 \bar{\bar{\xi}}(r_{12}), \quad (\text{A.62})$$



$$\check{\xi}_5 = \beta^2 \bar{\xi}(r_{12}), \quad (\text{A.63})$$

and

$$\check{\xi}_6 = \frac{\beta^2}{6} \check{\xi}(r_{12}). \quad (\text{A.64})$$

These then have the correspondingly complicated shape functions:

$$B_3 = -\frac{2}{3}\mu_{12}^2 + \frac{2}{3} + \frac{1}{6} \frac{\alpha(r_1)\alpha(r_2)\mu_{12}r_{12}^2}{r_1r_2} + \frac{1}{6} \frac{\alpha(r_1)\alpha(r_2)\mu_1\mu_2r_{12}^2}{r_1r_2} - \frac{1}{2} \left[ \frac{\alpha(r_1)r_{12}}{r_1} \mu_1\mu_2^2 + \frac{\alpha(r_2)r_{12}}{r_2} \mu_1^2\mu_2 \right] - \frac{1}{3} \left[ \frac{\alpha(r_1)r_{12}}{r_1} \mu_{12}\mu_2 + \frac{\alpha(r_2)r_{12}}{r_2} \mu_1\mu_{12} \right] \left\{ + \frac{1}{6} \left[ \frac{\alpha(r_1)r_{12}}{r_1} \mu_1 + \frac{\alpha(r_2)r_{12}}{r_2} \mu_2 \right] \right\}, \quad (\text{A.65})$$

$$B_5 = \frac{4}{5}\mu_{12}^2 - \frac{4}{5} - \frac{1}{30} \frac{\alpha(r_1)\alpha(r_2)\mu_{12}r_{12}^2}{r_1r_2} - \frac{1}{10} \frac{\alpha(r_1)\alpha(r_2)\mu_1\mu_2r_{12}^2}{r_1r_2} + \frac{1}{2} \left[ \frac{\alpha(r_1)r_{12}}{r_1} \mu_1\mu_2^2 + \frac{\alpha(r_2)r_{12}}{r_2} \mu_1^2\mu_2 \right] + \frac{1}{5} \left[ \frac{\alpha(r_1)r_{12}}{r_1} \mu_{12}\mu_2 + \frac{\alpha(r_2)r_{12}}{r_2} \mu_1\mu_{12} \right] - \frac{1}{10} \left[ \frac{\alpha(r_1)r_{12}}{r_1} \mu_1 + \frac{\alpha(r_2)r_{12}}{r_2} \mu_2 \right], \quad (\text{A.66})$$

$$B_6 = -\frac{\alpha(r_1)\alpha(r_2)\mu_{12}r_{12}^2}{r_1r_2} \quad (\text{A.67})$$

Notice that the quantity in curly braces in equation A.65 is exactly the same as  $B_1$ .

Therefore, remove this term from equation A.65 and say that

$$\check{\xi}_1 = \left( \frac{\beta}{3} + \frac{\beta^2}{6} \right) \bar{\xi}(r_{12}). \quad (\text{A.68})$$

The final term of  $B_5$  also has the form of  $B_1$ . However adding this term to  $\check{\xi}_1$  would make the  $\beta$  term inseparable from the averages of  $\xi$ .

Although the shape functions and resulting  $\check{\xi}$ 's are not unique, the above set of functions is complete and has the desirable property of being separable into functions of the shape, functions of  $\beta$  and functions of averages of  $\xi^r$ . This set of functions will allow for a complete description of spherical redshift distortions in the linear regime.

## Appendix B

### Definitions of Terms and Symbols

To aid in understanding, I have provided a list of important terms with definitions. Following the list of terms is two lists of symbols (one Roman and one Greek) with definitions.

#### B.1 Terms

**brute force:** Introduced in Chapter 1. The brute force method, commonly used in analysis of CMB fluctuations, assumes that the density is Gaussianly distributed. It also assumes a Gaussian likelihood function. These methods work best on the largest scales: where the fluctuations are likely to be Gaussian and the computation costs are the least.

**classical method:** Introduced in Chapter 1. Introduced by Feldman et al. (1994), classical methods of analyzing redshift surveys require the position and the wavelength of the fluctuation to be measured simultaneously. It implies a weight scheme that depends only on the local density. Classical methods work best on size scales that will comfortably fit within the survey.

**decorrelation:** Introduced in Section 2.3. Decorrelation is the process of creating statistically uncorrelated parameters. That is to say, recombine a set of parameters in such a way that the new parameters have covariances of zero. Decorrelation is especially useful if the data is going to be used in some sort of least-squares fit.

**Fisher matrix:** Introduced in Chapter 2. The inverse of the full covariance matrix for all

parameters. The Fisher matrix is generated using a prior guess as to the values of all measured parameters. If the prior guess is similar to the measured value then the Fisher matrix describes the amount of information a particular data set has about the set of measurable parameters.

**linear (regime):** Introduced in Chapter 2. Refers to size scales which are large enough that overdensities (or underdensities) are small compared to the mean density. In linear regime, overdensities grow linearly without change of shape.

**non-linear (regime):** Introduced in Chapter 2. Refers to size scales where overdensities can become comparable to or larger than the mean density. The fact that density cannot become negative causes underdense regions to expand and overdense regions to collapse.

**pair weight method:** Introduced by this thesis. The pair weight method is a general method of analyzing redshift surveys. The pair weight method weights each pair of overdensities by a weight that is determined by the survey itself (and the prior power spectrum). The pair weight method is not limited by the shape or complexity of the survey nor is it limited to testing models which are Gaussian. The primary limitation of the pair weight method is the computational expense.

**power spectrum:** Introduced in Chapter 1. The power spectrum is the covariance, expressed in Fourier space, of fluctuations. It is a function of wavenumber  $k$ . The shape and amplitude of the power spectrum can be used to measure several cosmological parameters.

**prior:** Introduced in Chapter 2. The prior includes all prior assumptions including *e.g.*, homogeneity and isotropy. In particular, the prior often refers to the prior guess for the power spectrum.

**selection function:** Introduced in Chapter 1. The expected number of galaxies at a particular location given the selection criteria of the survey.

## B.2 Symbols

$b$ : Introduced in the Abstract.  $b$  is the linear bias parameter.  $b$  is the ratio of galaxy overdensities to matter overdensities ( $\delta_{\text{gal}} = b\delta_{\text{mass}}$ ). If  $b = 1$  then galaxies are an unbiased tracer of the density.

$B_{\alpha ij}$ : Introduced in Section 2.1.2. The shape function. In the case with no redshift distortions this restricts the separation of parcels  $i$  and  $j$  to be  $r_\alpha$ . In the more general case,  $B_{\alpha ij}$  also includes a shape (contained within the index  $\alpha$ ) which describes one of the components of the redshift distortions.

$C_{ab(\alpha)}$ : Introduced in Chapter 2. The covariance of the modes  $a$  and  $b$ . If there are indices  $\alpha$ , then it is the derivative of the covariance of the modes with respect to the parameter  $\alpha$ .

$\mathcal{C}_{ab}$ : Introduced in Section 4.3.5. This is the band power matrix which disentangles  $P(k)$  from  $\beta P(k)$  from  $\beta^2 P(k)$  and correlates neighboring points within each of the power spectra.

$\mathfrak{C}_{ijkl}$ : Introduced in Section 2.1.2. The covariance of the correlation functions  $\langle \Delta \xi_{ij} \Delta \xi_{kl} \rangle$ . In the most general, non-Gaussian, case this can depend on 3-point terms and 4-point terms as well as 2-point terms (in the general case  $\mathfrak{C}_{ijkl}$  depends on the locations of parcels  $i, j, k$ , and  $l$  as well as their separations). For most of the Dissertation this is reduced to the more restricted Gaussian case where  $\mathfrak{C}_{ijkl} = C_{ik}C_{jl} + C_{il}C_{jk}$ .

$D_{\text{dist}}$ : Introduced in Chapter 4. This is the distortion operator. It describes the manner in which overdensities in real space translate into overdensities in redshift space.

$\mathcal{D}_{ab}$ : Introduced in Section 4.3.5. This is the band-power matrix which disentangles  $P(k)$  from  $\beta P(k)$  from  $\beta^2 P(k)$  and decorrelates neighboring points of each (however, at each  $k$  the three different power spectra are correlated). Same as  $\mathcal{C}$  except decorrelated not correlated.

$f$ : Introduced in Chapter 1.  $f (\approx \Omega_m^{0.6})$  is the linear growth rate. It is the ratio of the velocity of a parcel of matter to the comoving distance it has traveled (in Hubble units).

$F_{\alpha\beta}$ : Introduced in Section 2.1.1. The Fisher matrix. See Tegmark et al. (1997) for a nice explanation. This is the inverse of the covariance matrix with parameters  $\alpha$  and  $\beta$ . The Fisher matrix is also called the Fisher information matrix because it is a measure of how much information is contained within the data set about the particular set of parameters.

$\mathcal{F}$ : Introduced in Section 4.3.5. This matrix is used to correlate and disentangle the power spectra ( $P(k)$ ,  $\beta P(k)$ , and  $\beta^2 P(k)$ ).

$G_{\alpha\beta}$ : Introduced in Section 2.3.1. This is the scaled Fisher matrix. That is to say that it is the Fisher matrix multiplied by the prior values of  $\xi_\alpha$  and  $\xi_\beta$ .

$h$ : Introduced in Chapter 1. The Hubble constant  $H_0$  in units of  $100 \text{ km s}^{-1} \text{ Mpc}^{-1}$ .

$H_0$ : Introduced in Chapter 1.  $H_0$  is the Hubble constant at the present epoch. The Hubble constant is a measure of the rate of expansion of the universe.

$\check{J}_i(kr_{12})$ : Introduced in Chapter 4. The generalized Fourier transform window. These are the window functions necessary to transform  $\check{\xi}_i(r)$  to  $g_i(\beta)P(k)$ . For indices 0, 2, and 4 the generalized window is just the spherical Bessel function  $j$ .

$J_{3\alpha}$ : Introduced in Section 2.3.1. The volume integral of the two-point correlation function out to radius  $r_\alpha$ . In other words, the number of additional galaxies surrounding a galaxy than would be expected if galaxies were uniformly distributed. This is used as an FKP constant for the separation  $r_\alpha$ .

$\mathcal{L}$ : Introduced in Section 2.1.1. The likelihood function. The likelihood function is high where the data fits the model and low where the data does not fit very well.

$\bar{n}_i$ : Introduced in Chapter 2. The expected density of galaxies (in the absence of fluctuations) at location  $\mathbf{r}_i$ . Also referred to as the selection function.

$K_{ab}$ : Introduced in Section 2.1.2. The integral of the covariance of correlation functions weighted by window functions  $a$  and  $b$ .

$M_{\alpha\beta}$ : Introduced in Chapter 2.1.2. A matrix defined in introducing the Pair Weight method. It becomes clear that  $M$  is, in fact, the Fisher matrix ( $F$ ).

$\mathcal{M}$ : Introduced in Section 4.3.5. This matrix is used to disentangle the power spectra ( $P(k)$ ,  $\beta P(k)$ , and  $\beta^2 P(k)$ ) while decorrelating neighboring points of each of the power spectra.

$n_s$ : Introduced in Chapter 1. The “tilt” of the power spectrum.

$P_k$ : Introduced in Chapter 1. The power spectrum at wavenumber  $k$ .

$r$ : Superscript  $r$  shows a quantity that is in real space.

$s$ : Superscript  $s$  shows a quantity that is in redshift space.

$W$ : Introduced in Equation 2.1. A window function. In Equation 2.1 the indices of  $W_{ij}$  refer to the locations of the two positions. Thereafter, the first index (usually  $a$  or  $b$ ) refers to the window number. In particular, there may be separate windows for different separations and, within those windows, there may be different windows for different shape functions. The following indices (usually some combination of  $i$ ,  $j$ ,  $k$ , and  $l$ ) refer to locations. In general, the windows are chosen to try to make the best use of the data.

$\hat{x}$ : Introduced in Chapter 2. A mode amplitude. This is the overdensity weighted by a linear window function. The hat signifies that it is a measured quantity.

$\hat{X}$ : Introduced in Chapter 2. A quadratically weighted mode. This is the overdensities weighted by a quadratic window (one with two indices of location). The hat signifies that it is a measured quantity.

### B.3 Greek Symbols

$\alpha$ : Introduced in Chapter 4. This quantity ( $\alpha = \partial \ln r^2 \bar{n} / \partial \ln r$ ) is 2 + the partial derivative of the logarithm of the selection function with respect to the logarithm of the distance. It determines how radial variations of the selection function affects redshift distortions.

$\beta$ : Introduced in the Abstract.  $\beta(= f/b)$  is the amplitude of the linear redshift distortions.

$\delta_i$ : Introduced in Chapter 2. The overdensity (of galaxies) at location  $\mathbf{r}_i$ . (If  $\delta_i = 1$  then that means the density at  $\mathbf{r}_i$  is twice the mean density.)

$\delta_{ij}$ : Used in Chapter 2. This is a 3-dimensional delta-function.

$\xi$ : Introduced in Chapter 2. The 2-point correlation function in some space. In  $k$ -space this is the power spectrum.

$\xi^r$ : Introduced in Chapter 4. This is the correlation function in real space.

$\xi^s$ : Introduced in Chapter 4. This is the correlation function in redshift space.

$\hat{\xi}_\alpha$ : Introduced in Section 2.1.1. The hat indicates that this is a measurement of the quantity rather than a prior guess of the quantity.

$\check{\xi}$ ,  $\bar{\xi}$ , and  $\bar{\bar{\xi}}$ : Introduced in Chapter 4. Averages of the two point correlation function over different functions of  $r$ . These averages become interesting due to redshift distortions.

$\check{\check{\xi}}$ : Introduced in Chapter 4. Measurable quantities in the case of redshift distortions. Each contains a function  $g$  of  $\beta$  and a  $\Xi$ . Each  $\check{\check{\xi}}$  can be Fourier transformed into  $g(\beta)P(k)$  (using the proper generalized Fourier transform window).

$\Xi_i$ : Introduced in Chapter 4. Functions of  $\xi$ ,  $\check{\xi}$ ,  $\bar{\xi}$ , and  $\bar{\bar{\xi}}$  produced by redshift distortions. Each of the  $\Xi_i$ 's can be Fourier transformed into  $P(k)$  (using the proper generalized Fourier transform window).

$\rho_a$ : Introduced in Chapter 1. The density of constituent  $a$ .

$\rho_{\text{crit}}$ : Introduced in Chapter 1. The critical density of the Universe. If the sum of the densities of the sum of all constituents is the critical density then the universe is flat.

$\Omega_a$ : Introduced in Chapter 1. The ratio  $\rho_a/\rho_{\text{crit}}$  of the density of constituent  $a$  ( $\rho_a$ ) to the critical density  $\rho_{\text{crit}}$ .

$\Omega_b$ : Introduced in Chapter 1. The ratio  $\rho_b/\rho_{\text{crit}}$  of the density of baryons to the critical density  $\rho_{\text{crit}}$ .

$\Omega_\Lambda$ : Introduced in Chapter 1. The ratio  $\rho_\Lambda/\rho_{\text{crit}}$  of vacuum energy density to the critical density  $\rho_{\text{crit}}$ .

$\Omega_m$ : Introduced in the Abstract. The ratio  $\rho_m/\rho_{\text{crit}}$  of the density of matter to the critical density  $\rho_{\text{crit}}$ .