# Using the Lyman-$\alpha$ Forest to Determine the Power Spectrum of Mass Density Perturbations

Alasdair Higgins

Doctor of Philosophy

The University of Edinburgh

2002

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

November 10, 2002

This thesis is dedicated to my landlady, good friend and guardian angel Rosalind Foggie.

# Acknowledgements

First of all I'd like to thank top bloke and cosmology ace John Peacock. Without his excellent advice and support I doubt I'd ever have finished at all. Also on the technical side (and of course for also being really nice people) thanks must go to Alan Heavens, Will Percival and Michael Brown for 'useful discussions', though to be honest useful explanations would probably be more accurate.

Next up much appreciation must go to my beloved office mates, both past and present, for creating a unique and hugely enjoyable working environment, so cheers to James, Michael (again), Lou, Helen and Liv.

Since this thesis is, embarrassingly enough, the result of nearly 4 years work, thanks must go to Em for being nice to me for most of that time.

Loads of people deserve thanks for looking after me, you all know who you are so thank you. That said some really went far beyond the reasonable call of duty, to list their individual acts would take far too long, but they all know why they are on this list, so massive debts of gratitude to my Mum, Fi, Keith, Rossy and of course the near infinitely patient and wonderful Olivia.

Finally I'd like to thank the mighty Alex McLeish for invaluable advice about the wacky worlds of football and women.

# Abstract

I present an analysis of techniques used to determine the mass power spectrum from observations of the intergalactic medium.

Simple Monte Carlo simulations are presented which clarify some of the problems which must be overcome when attempting any inversion process. The 1 dimensional flux power spectrum is dependent on both the clustering of the absorption features and on the scaling relations existing between the dark matter and baryon distributions.

More realistic simulated spectra are used to illustrate differences in the mean flux power spectrum for a range of cosmologies and the ratio of the 1 dimensional linear mass and flux power spectra. The mass distribution is much more dependent on cosmological parameters than the flux power spectrum, highlighting the very accurate measurements of the latter quantity required for an accurate recovery of the mass density power spectrum.

These spectra are further analysed by deconvolution into Voigt profiles. This technique is shown to be an excellent approximation, in spite of the current cosmic web paradigm of the Ly-$\alpha$ forest. However the power spectrum of the positions of these fitted lines is shown to be a very poor indicator of the underlying mass density field.

Inferring the 3 dimensional forms of power spectra from 1 dimensional data is shown to be problematic. This, coupled with analysis of the correlation matrix of the flux power spectrum, casts doubts on the reliability of the error analysis presented in the literature.

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction

One of the aims of observational cosmology is to determine the functional form of the power spectrum of initial mass density perturbations, which are believed to be the seeds for modern day structures. As well as being of interest in its own right, it could also provide a test for theories of cosmological structure formation and offer constraints on the values of cosmological parameters. Importantly for a Gaussian field, which the initial density perturbations are expected to be (e.g. Fischler, Ratra & Susskind 1985; Peebles 1993), the power spectrum offers a statistically complete description. Measuring the primordial mass power spectrum is unfortunately a difficult task since non-linear gravitational growth of the inhomogeneities causes substantial alteration of the power spectrum with time. Therefore the mass power spectrum seen today is a non-Gaussian field due to being distorted from its primordial form.

Advances in the understanding of the equation of state of the intergalactic medium (IGM) has meant that this has become an area of study which can now provide an illuminating companion to other, more traditional, methods of determining the mass power spectrum. In this thesis I will present work in this field. Firstly in this chapter an outline of some vital background material will be given. This will include basic cosmology, statistics, absorption physics and a review of the papers upon which the work in this thesis in based.

In chapter 2 I shall present results from simple computer simulations of the IGM. These results will be discussed with respect to certain difficulties they highlight in using the IGM to recover the mass power spectrum. Recovering the mass power spectrum by using Lyman-$\alpha$ (hereafter Ly-$\alpha$) absorbers as indicators of peaks in the mass density density field is also discussed. The lognormal approximation (Coles & Jones 1991) is introduced and applied to simulate mass density fields.

More sophisticated simulations will then be analysed in chapter 3. This chapter includes a study of the relationship between 1 dimensional mass and flux spectra for various cosmologies. Line lists of the spectra are also found and used to calculate the power spectrum of line positions and the relation to the mass power spectrum is discussed. The importance of the profiles of individual Ly-$\alpha$ absorbers is also highlighted through a technique where line positions are scrambled.

Chapter 4 contains work on the statistics of the techniques studied in the previous chapters. A study of the flux correlation matrix and the number of independent points being examined is presented. The difficulties in inferring the 3 dimensional flux power spectrum is also demonstrated, and an alternative of working with likelihoods based on the directly observable 1 dimensional data is presented. Also the relationship between the mass and flux variances for individual realizations is briefly discussed.

Finally all this work will be summarised and conclusions presented in chapter 5.

## 1.1 Basic cosmology

In this section I will outline some of the fundamental observations and theories in cosmology which provide the framework within which the work in this thesis is based. The paradigm for modern cosmology is the hot big bang inflationary model. In this model the Universe began from a singularity a finite time ago. A major triumph of this model is a prediction for relative abundances of light elements which matches the observations. The hot big bang model results in the Universe being statistically isotropic and homogeneous. Observationally isotropy is very well motivated, with structure distributions

being highly isotropic on large scales. The strongest evidence however comes from the cosmic microwave background (CMB). Once corrected to account for the movement of our galaxy, the CMB is observed to be the same to 1 part in $10^5$ in all directions (e.g. de Bernardis *et al.* 2000). To say more it becomes essential to invoke the Copernican Principle, which states that we are not privileged observers. Therefore it is postulated that if the Universe is isotropic about the Earth, then it is isotropic about all locations. Any mass distribution which is isotropic about all positions can be readily shown to be homogeneous. Another key feature about the Universe is that observations show it to be expanding (as will be discussed later). These three elements form the basis for modern day cosmology.

### 1.1.1 The Robertson Walker metric

The metric required to describe an isotropic, homogeneous and expanding universe can be completely inferred from symmetry arguments. This was first done independently by H.P. Robertson and A.G. Walker. The metric can be written in a number of forms, for example:

$$ds^2 = c^2 d\tau^2 = c^2 dt^2 - R^2(t) \left( \frac{dr^2}{1 - kr^2} + r^2 d\psi^2 \right). \tag{1.1}$$

There are a few things to note with the Robertson-Walker (RW) metric. First it is usual to establish a set of 'fundamental observers'. These are hypothetical observers who are at rest with respect to the local matter in their vicinity and are free from any strong gravitational fields. The time coordinate of the metric, which is referred to as cosmic time, uses the fact that the Universe is expanding to allow fundamental observers to synchronise their clocks. The idea is that because the density of the Universe is changing, it allows fundamental observers to use the value of density they see as a measure of time. In the spatial part, the $r$ and $\psi$ coordinates are comoving, meaning the coordinates expand with the Universe, for example two objects moving apart with the expansion but with no peculiar velocity with respect to each other will have the same comoving separation, $r$, at all times. The comoving coordinates between two points is related to the proper physical distance between them at any given time by

the scale factor $R(t)$. The coordinates are determined so that the proper and comoving distances are coincident now, so $R_0 = 1$. Finally the factor $k$ is determined by the curvature of the Universe and can take one of three values, $\pm 1$ or 0. Each of these values correspond to very different geometries.

$k = 1$ applies when space is 'closed', geometry corresponds to the surface of a sphere so any trajectory of fixed $\psi$ will always return to its starting point.

$k = 0$ applies when space is flat and geometry is Euclidean.

$k = -1$ applies when space is 'open' and infinite in extent.

### 1.1.2 Redshift

The form of the scale factor that appears in the RW metric can be found in terms of a fundamental observable in astrophysics, the redshift, $z$. Distances to objects will normally be expressed in terms of a redshift rather than proper or comoving distance since calculating the latter requires knowledge of certain cosmological parameters. The redshift is defined as

$$1 + z \equiv \frac{\lambda_o}{\lambda_e}, \tag{1.2}$$

where the subscripts o and e represent observed and emitted respectively. This expression can be linked to the RW metric via the expression obtained for radial light rays to show that

$$1 + z \equiv \frac{\lambda_o}{\lambda_e} = \frac{R(t_o)}{R(t_e)} = \frac{1}{R(t_e)}. \tag{1.3}$$

For small redshifts the recession velocity is related linearly to the redshift, and is given by

$$z \simeq \frac{v}{c}. \tag{1.4}$$

### 1.1.3 Hubble's law

A universe that expands uniformly results in distant objects recessing faster than nearby objects. In 1929 Edwin Hubble discovered a correlation between the recession velocities of galaxies and their distance from us. The resulting equation is known as Hubble's law

$$v = H_0 r_{\rm p}, \tag{1.5}$$

where $v$ is the velocity of recession, $r_{\rm p}$ is the proper distance to the galaxy and $H_0$ is the Hubble constant, usually measured in units of km s$^{-1}$Mpc$^{-1}$. The Hubble constant is the current day value of the Hubble parameter $H(t)$, which governs the rate of expansion at a given time. The Hubble parameter changes with time for any universe which is not empty, due to the pull of gravity slowing down the expansion.

For a small separation, where things are Euclidean the proper separation of two fundamental observers is just $R(t)dr$, which upon substitution into Hubble's law, and re-arranging gives

$$H = \frac{\dot{R}}{R}. \tag{1.6}$$

The value of $H_0$ has been the subject of some debate ever since Hubble's first estimate of its value to be of the order of 500 km s$^{-1}$Mpc$^{-1}$. The most recent measurements of The Hubble Space Telescope Extragalactic Distance Scale Key Project, which uses Cepheid variables to estimate $H_0$ give a value of $72 \pm 8$ km s$^{-1}$Mpc$^{-1}$ (Freedman *et al.* 2001). Due to the uncertainty in the value of $H_0$ it is common practice to use a parameter $h$ defined as

$$h \equiv \frac{H_0}{100 \text{ km s}^{-1}\text{Mpc}^{-1}}. \tag{1.7}$$

### 1.1.4 The density and fate of the Universe

The dynamics of the Universe can be calculated using the theory of general relativity. This results in the Friedmann equations

$$\ddot{R} = -\frac{4\pi G}{3}\left(\rho + \frac{3p}{c^2}\right)R + \frac{\Lambda R}{3}, \tag{1.8}$$

$$\dot{R}^2 = \frac{8\pi G\rho}{3}R^2 - kc^2 + \frac{\Lambda R^2}{3}. \tag{1.9}$$

In the equations above $\rho$ is the mass density at time $t$, $p$ is the pressure of matter and $\Lambda$ is the cosmological constant as described below. A simple Newtonian argument results in similar expressions to those above, but a full general relativistic treatment is needed to obtain the $\Lambda$ terms and the pressure term in equation 1.8.

$\Lambda$ is a constant of integration representing the energy of a vacuum. It was originally included by Einstein in his field equations, not because he thought a non-zero vacuum energy was well motivated, but because he wanted a solution to these equations which could result in a static universe. Vacuum energy requires a negative pressure equation of state resulting in a universal repulsive force. It also has massive implications for the flatness and fate of the Universe as will be shown later.

Looking at equation 1.9, it can be seen that the density of the Universe is related to its geometry (via the constant $k$). Further to this it is clear that for any given Hubble parameter there is a value of the density for which $k = 0$. The density corresponding to $k = 0$ in a universe without any cosmological constant contribution is known as the critical density. Often the density of the Universe is expressed as a dimensionless parameter $\Omega$ defined as

$$\Omega \equiv \frac{\rho}{\rho_c} = \frac{8\pi G\rho}{3H^2}. \tag{1.10}$$

It is usual to denote the current day value of this parameter as $\Omega_0$. In a matter only universe, the value of $\Omega$ has a clear meaning, if it is less than 1 then the universe is 'open' and the gravitational attraction of all the matter in the universe is not enough

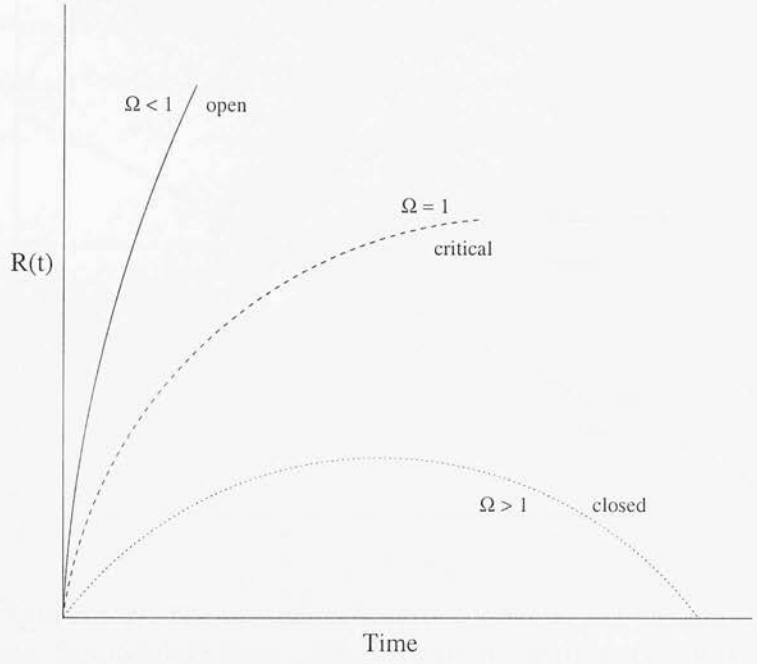


**Figure 1.1:** The scale factor as a function of time for a universe containing only matter. The scale factor in an open universe increases indefinitely. In a closed universe the scale factor will, after some finite time, become zero as everything collapses to a single point. A critical density universe loiters between these two extremes, with the scale factor asymptotically tending towards a constant at large times.

to ever halt the expansion. In this case the universe will continue to grow and cool indefinitely. If $\Omega$ is greater than 1 then the universe is 'closed', at some time in the future all the matter will eventually be pulled back, via gravitational interaction, to the same point (see figure 1.1).

More generally however contributions to the dimensionless density parameter from matter, vacuum energy and radiation must all be considered. In this case $\Omega$ without a subscript can be used for the sum of these contributions

$$\Omega = \Omega_m + \Omega_\Lambda + \Omega_r, \tag{1.11}$$

where $\Omega_\Lambda$ is given by

$$\Omega_\Lambda = \frac{\Lambda}{3H^2}. \tag{1.12}$$

It should be noted however that there is some ambiguity in the literature where $\Omega$

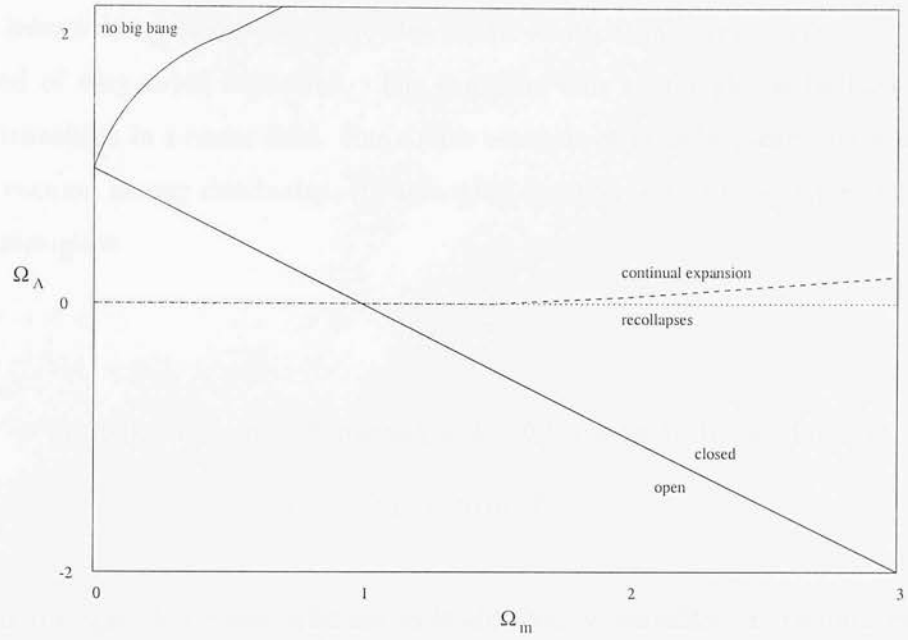


**Figure 1.2:** The fate and curvature of a matter and vacuum energy dominated universe. The curvature for a given cosmology is determined only by the sum of $\Omega_m$ and $\Omega_\Lambda$ not by the individual components. The fate of a universe depends on both $\Omega_m$ and $\Omega_\Lambda$. Any negative value of $\Omega_\Lambda$ will result in a universe which eventually recollapses. In some cosmologies a big bang singularity is excluded.

(rather than $\Omega_m$) is often used to represent solely the matter contribution. The fate of the Universe is affected by all the contributions to $\Omega$, however the different components effect the dynamics of the Universe in different ways as demonstrated in figure 1.2 (e.g. Peacock 1999).

Currently the preferred cosmological model is a flat model $\Omega_0 = 1$ with the modern day contributions of $\Omega_m \sim 0.3$ and $\Omega_\Lambda \sim 0.7$. The main evidence for this comes from the CMB (e.g. Balbi 2000) and observations of type 1a supernovae (e.g. Jha *et al.* 2001) as well as a host of other corroborative data. However question marks remain over the reliability of some of these results and the debate on this subject is far from over.

### 1.1.5 Inflation

Despite its success in many areas the hot big bang model gives rise to some interesting problems outlined below. A solution to these problems was proposed by Guth in 1981,

and is known as inflation. The basic idea is that in its infancy the Universe underwent a period of very rapid expansion. The cause for this is thought to be have been a phase transition in a scalar field. One simple example of an inflationary universe is one where vacuum energy dominates. Re-arranging equation 1.9 and assuming the $\Lambda$ term dominates gives

$$\frac{\dot{R}}{R} = \frac{\Lambda}{3}. \tag{1.13}$$

So for an inflationary period starting at $t = 0$ the scale factor at time $t$ is

$$R(t) = R(0)e^{\frac{\Lambda t}{3}}. \tag{1.14}$$

It is thought that when inflation ends the energy contained in vacuum energy is transformed into normal matter and radiation. Although the precise mechanism for the process of inflation remains unclear, it still presents a very attractive proposition since it solves the problems discussed below.

### The flatness problem

It can be shown that a universe of $\Omega = 1$ is unstable. Therefore small deviations from this value will quickly result in a big divergence from $\Omega = 1$. For a matter dominated universe (for the radiation dominated case deviations grow even more rapidly) the evolution of $\Omega$ with $z$ is given by

$$\frac{1}{\Omega(z)} - 1 = (1+z)^{-1}\left(\frac{1}{\Omega_0} - 1\right). \tag{1.15}$$

Since the value of $\Omega$ is close to unity now it must have been even more so in the past. In fact it can be shown that for sensible values of $\Omega_0$ the deviation from critical density at the Planck epoch must have been no larger than about 1 part in $10^{60}$. This required level of precision leads one to ask how is it that the Universe started with a density so close to critical density?

Inflation helps solve this problem since during an inflationary period the Universe is actually driven towards $\Omega = 1$. So regardless of the initial conditions a long enough

period of inflation will yield a universe which has a density extremely close to the critical value.

### The horizon problem

The earliest epoch that photons observed today can have been emitted is known as the recombination epoch. Prior to this the Universe was opaque due to light being scattered off free electrons. The size of casually connected regions at this time ($z \sim 1000$) corresponds to an observed angle on the sky of about 1 degree. However the CMB is seen to be isotropic over the entire sky. The question therefore is how is the CMB so smooth given that it corresponds to many supposedly distinct regions?

However since regions can be expanded by a huge amount during inflation it follows that the horizon problem can be solved. An initially small (and therefore casually connected) region can establish equilibrium before being inflated. The rapid expansion during inflation can then lead to the region becoming larger than the horizon scale. Therefore the observed isotropic CMB ceases to be a problem.

### The structure problem

Although statistically homogeneous and isotropic the Universe is clearly neither of these things on small scales. The structure seen is assumed to be the result of gravity acting on small primordial perturbations in density. The problem is explaining a mechanism to produce these perturbations.

In the inflationary model these initial perturbations in the density of the Universe can be attributed to quantum fluctuations. The uncertainty principle means that, even in a vacuum, the energy required to create a particle-antiparticle pair can be 'borrowed' for a short period of time. These pairs are continually phasing in and out of existence, and such an event is called a quantum fluctuation. During inflation quantum fluctuations can be frozen in, leading to small perturbations in density. After inflation has ended gravity acts on these perturbations to form the structure seen today.

Although offering a self consistent theory of the early Universe can inflation be corroborated observationally? Certainly some strong predictions arise, as previously discussed an inflationary period will lead to a spatially flat universe, also if initial perturbations were caused by quantum fluctuations then these should be scale invariant and Gaussian. The strongest piece of evidence to support inflation would be to observe the correct form of gravity-wave background. However this is way beyond current technology and will probably remain infeasible for some time. Clearly there is much still to understand about inflation, however despite this it provides an excellent foundation of understanding and is the canonical view of the astrophysical community today.

### 1.1.6 Structure formation and growth of perturbations

The structures seen in the Universe today are thought to be the result of gravity acting on small initial perturbations. A region of overdensity in a homogeneous background will, with time, become increasingly overdense. This is because any matter surrounding the perturbation will experience a gravitational attraction towards it. This process of regions growing ever more overdense does not continue indefinitely due to the effects of pressure and dissipative processes becoming important.

#### The growth of linear perturbations

Before perturbations become large enough to pass into the non-linear regime the treatment of their evolution can be approximated with a Newtonian treatment. This is because Birkhoff's theorem shows that a perturbed region will evolve, as governed by the Friedmann equations, as if it were an independent universe. To study the evolution of density perturbations it is useful to define a fractional density perturbation as

$$\delta \equiv \frac{\rho - \bar{\rho}}{\bar{\rho}}. \tag{1.16}$$

In a flat matter dominated universe the Friedmann equations for a perturbed region

(with density $\rho_{\rm p}$) and the background universe give

$$H^2 = \frac{8\pi G\bar{\rho}}{3} \qquad H_{\rm p}^2 = \frac{8\pi G\rho_{\rm p}}{3} - \frac{k_{\rm p}c^2}{R_{\rm p}^2}. \tag{1.17}$$

If this perturbation is small then the expansion rates and scale factors will be very similar for both cases and the resultant expression for $\delta$ is

$$\delta = \frac{k_{\rm p}c^2}{R^2H^2}. \tag{1.18}$$

So, since $H^2 \propto \rho \propto R^{-3}$, it can immediately be seen that in an Einstein de-Sitter universe

$$\delta \propto (1+z)^{-1}. \tag{1.19}$$

A similar argument for a general cosmology results in a much more involved calculation. This evolution of perturbations can be quantified by a linear growth factor defined as

$$D(a) \equiv \frac{\delta(a)}{\delta(0)}, \tag{1.20}$$

where

$$a \equiv 1/(1+z). \tag{1.21}$$

In general exact solutions for $D(a)$ can only be obtained by numerical integration of a function determined by the Friedmann equation (Heath 1977) given by

$$D(a) = \frac{5\Omega_m}{2a}\frac{da}{d\tau}\int_0^a \left(\frac{da'}{d\tau}\right)^{-3} da', \tag{1.22}$$

where

$$\left(\frac{da}{d\tau}\right)^2 = 1 + \Omega_m\left(\frac{1}{a} - 1\right) + \Omega_\Lambda(a^2 - 1). \tag{1.23}$$

It is often useful to quantify the modern day value of the linear growth for any given cosmology relative to the growth factor for an Einstein De-Sitter universe such that

$$g(\Omega_m, \Omega_\Lambda) = \frac{\delta(z=0, \Omega_m, \Omega_\Lambda)}{\delta(z=0, \Omega_m = 1, \Omega_\Lambda = 0)}. \tag{1.24}$$

Where the values of $\delta$ for the different cosmologies are equal at high $z$. An excellent approximation has been shown (Carroll, Press & Turner 1992) to be

$$g(\Omega_m, \Omega_\Lambda) \simeq \frac{5}{2}\Omega_m \left[\Omega_m^{4/7} - \Omega_\Lambda + (1 + \Omega_m/2)(1 + \Omega_\Lambda/70)\right]^{-1}. \quad (1.25)$$

### 1.1.7 Dark matter

Much of the matter in the Universe appears to be dark and its presence can only be inferred by its gravitational interaction. Galaxy rotation curves and gravitational lensing analysis for example show that the matter which is actually visible accounts for only a very small fraction of the total mass in galaxies. The important question is clearly what actually comprises this dark matter? If the dark matter is in the form of baryons then it is likely to be massive compact halo objects (MACHOs) such as white dwarfs or brown dwarfs. However using the ratio of hydrogen to deuterium to determine a value for the amount of baryons present in the Universe gives $\Omega_b = 0.019 \pm 0.0024$ (e.g. Tytler *et al.* 2000), whereas $\Omega_m$ is thought to be about 0.3. This implies that most dark matter is non-baryonic. There are two main candidates for the form of non-baryonic dark matter, cold and hot (meaning it is travelling relativistically at the time of recombination).

Cold dark matter (CDM) candidates consist mostly of weakly interacting massive particles (WIMPs) such as heavy neutrinos. Structure formation theories seem to demand that at least some of the dark matter is cold. A favourable WIMP candidate to make up this component is the neutralino, a particle which is predicted by supersymmetry.

Hot dark matter (HDM) refers to particles which are moving relativistically. An example of a HDM candidate is the light neutrino which from a particle physics perspective is far more desirable than a heavy one. However the properties of HDM do not correlate well with structure formation theories so both forms of dark matter bring with them their own pros and cons.

## 1.2 The correlation function and power spectrum

This section outlines the statistical definitions of the expressions encountered later. The correlation function and power spectrum defined in this section give a formal mechanism via which the inhomogeneities in the mass density field of the Universe can be described.

The correlation function of a density field can be defined as

$$\xi(\mathbf{x}) \equiv \langle \delta(\mathbf{x}')\delta(\mathbf{x}'+\mathbf{x}) \rangle, \tag{1.26}$$

where $\langle\rangle$ represents an average over some normalization volume $V$, and $\delta(\mathbf{x})$ is the fractional density perturbation defined in equation 1.16.

It is usual to then express $\delta(\mathbf{x})$ as a superposition of plane wave modes

$$\delta(\mathbf{x}) = \sum_{\mathbf{k}} \delta_{\mathbf{k}} e^{-i\mathbf{k}.\mathbf{x}}, \tag{1.27}$$

where

$$\mathbf{k} = \frac{2\pi}{L}(n_x, n_y, n_z) \qquad n_{x,y,z} = 1, 2, \ldots \tag{1.28}$$

Although this treatment is not strictly valid if space has curved geometry, the corrections needed are small on almost all scales and can be neglected.

Substituting equation 1.27 into 1.26 and noting that since $\xi(\mathbf{x})$ is a real field, one of the $\delta$'s can be replaced by its complex conjugate gives

$$\xi(\mathbf{x}) = \left\langle \sum_{\mathbf{k}} \sum_{\mathbf{k}'} \delta^*_{\mathbf{k}} \delta_{\mathbf{k}'} e^{i(\mathbf{k}-\mathbf{k}').\mathbf{x}'} e^{-i\mathbf{k}'.\mathbf{x}} \right\rangle. \tag{1.29}$$

Applying periodic boundary conditions (so all $\mathbf{k}' \neq \mathbf{k}$ terms average to zero) and expressing the equation as an integral rather than a sum gives the result

$$\xi(\mathbf{x}) = \frac{V}{(2\pi)^3} \int \langle |\delta_{\mathbf{k}}|^2 \rangle e^{-i\mathbf{k}.\mathbf{x}} \, d^3\mathbf{k}. \tag{1.30}$$

It can now be seen that the correlation function is the Fourier transform of $\langle|\delta_{\mathbf{k}}|^2\rangle$ which is referred to as the power spectrum, $P(\mathbf{k})$, defined such that

$$P(\mathbf{k}) \equiv \lim_{V\to\infty} \langle|\delta_{\mathbf{k}}|^2\rangle. \tag{1.31}$$

This expression can be simplified using the properties of isotropy. In an isotropic universe there can be no preferred direction and so only the modulus of $\mathbf{k}$ is important thus $P(k) = \langle|\delta_{\mathbf{k}}|^2\rangle = \langle|\delta_k|^2\rangle$. Therefore the angle integration can be performed (e.g. Gradshteyn & Ryzhik 1994) giving

$$\xi(x) = \frac{V}{(2\pi)^3}\int P(k)\frac{\sin kx}{kx}4\pi k^2\ dk. \tag{1.32}$$

Having defined these quantities it is straightforward to define their one dimensional counterparts. Isotropy requires the functional form of the correlation function to be the same in both cases so $\xi(x) = \xi(\mathbf{x})$. However by considering the 1D and 3D Fourier transforms of $\xi$, it can immediately be seen that the 1D and 3D power spectra must have different functional forms. The relationship between them was shown (Lumsden, Heavens & Peacock 1989) to be

$$P_{1\mathrm{D}}(k) = \frac{1}{2\pi}\int_k^\infty P(y)y\ dy, \tag{1.33}$$

$$P(k) = -\frac{2\pi}{k}\frac{d}{dk}P_{1\mathrm{D}}(k). \tag{1.34}$$

Throughout this thesis any power spectrum term without a subscript will refer to the 3D form and the 1D form will always be marked as such. The power spectrum is usually expressed in a dimensionless form as the variance of $\delta$ per logarithmic bin

$$\Delta^2(k) \equiv \frac{V}{(2\pi)^3}4\pi k^3 P(k) = \frac{d\langle\delta^2\rangle}{d\ln k}, \tag{1.35}$$

with the 1 dimensional equivalent given by

$$\Delta^2_{1\mathrm{D}}(k) = \frac{L}{\pi}kP_{1\mathrm{D}}(k). \tag{1.36}$$

A value of $\Delta^2(k) = 1$ corresponds to order unity density fluctuations from modes in the logarithmic bin about wavenumber $k$.

Up until now the discussion has concentrated solely on continuous distributions. In many real and simulated data sets, a continuous variable is sampled at regular intervals. Crossing from a continuous to a discrete distribution has a non trivial effect, where a Fourier transform is now replaced by a fast Fourier Transform (FFT). In the case of a real field, $x_i$, sampled at $N$ points where $i = 0, 1, 2, ..., N-1$, the FFT components, $\tilde{x}_k$, can be calculated by

$$\tilde{x}_k = \frac{1}{N} \sum_{j=0}^{N-1} x_j \times \exp\left(-i\frac{2\pi jk}{N}\right), \tag{1.37}$$

with the inverse relation

$$x_j = \sum_{k=0}^{N-1} \tilde{x}_k \times \exp\left(i\frac{2\pi jk}{N}\right). \tag{1.38}$$

For example later work frequently refers to the flux power spectrum which is obtained by taking an FFT of the spectra using $N$ flux points, $f_j$ and then squaring the resultant $\tilde{f}_k$'s to give $P_{\rm 1D,F}$.

It should also be noted that a correlation function can be defined for a distribution of discrete objects as well as a field. If there is a density of objects $n$, then $\xi(\mathbf{x})$ gives the excess probability of finding an object in an element $dV$ centred at $\mathbf{x}$ from another one. Once again isotropy means only the magnitude of $\mathbf{x}$ is important and the direction can be ignored

$$dp = n\,[1 + \xi(x)]\,dV. \tag{1.39}$$

Although both forms of $\xi(x)$ will be used the definition will be clear from the context.

## 1.3 The physics of absorption

This section contains an outline to the physics involved in the processes of absorption and emission within the classical regime. The quantum mechanical basis for these processes is reviewed in appendix A.

### 1.3.1 Emissivity and opacity

Classically the change of specific intensity $I_\nu$ of light passing through a thickness $dx$ of an absorber can be described by the equation

$$\frac{dI_\nu}{dx} = -\kappa_\nu I_\nu + \epsilon_\nu, \tag{1.40}$$

where $\kappa_\nu$ and $\epsilon_\nu$ are the opacity and emissivity of the matter in question. The emissivity relates to the Einstein $A_{\mathrm{ul}}$ coefficient (see Appendix A) since the energy emitted will be equal to the number of transitions occurring multiplied by the energy of each of the corresponding photons. Since this process is isotropic, dividing this energy by the total solid angle gives the specific intensity

$$\int \epsilon_\nu \, d\nu = n_{\mathrm{u}} A_{\mathrm{ul}} \frac{h\nu}{4\pi}. \tag{1.41}$$

The integration accounts for a small spread of energy about $\nu_{ul}$. This spread can be written as a function $\phi_\nu$ such that $\int \phi_\nu \, d\nu = 1$, giving

$$\epsilon_\nu = n_{\mathrm{u}} A_{\mathrm{ul}} \frac{h\nu}{4\pi} \phi_\nu. \tag{1.42}$$

Similarly the optical depth can be defined in terms of opacity as

$$\tau_\nu \equiv \int d\tau_\nu = \int \kappa_\nu \, dx. \tag{1.43}$$

By a similar argument to the one given for emissivity it can be shown that $\kappa_\nu$ is related to the Einstein $B_{\mathrm{lu}}$ coefficient, or alternatively the cross section by the equation

$$\kappa_\nu = n_{\mathrm{l}} \sigma \phi_\nu. \tag{1.44}$$

Often the $\epsilon_\nu$ term is negligible in which case equation 1.40 integrates to

$$I_\nu(x) = I_\nu(0)e^{-\tau_\nu}. \tag{1.45}$$

So far the arguments have referred to the density of absorbers $n_\mathrm{l}$. When studying absorption features what is important is the total number of absorbers encountered thus it is more usual to refer to the column density $N$ defined as

$$N = \int n_\mathrm{l}\, dx. \tag{1.46}$$

It is usual to measure column densities in $\mathrm{cm}^{-2}$ and to mark $N$ with a subscript to show the identity of the absorber. $N_\mathrm{HI}$ for example represents the column density of neutral hydrogen.

### 1.3.2 Line width

Any transition between states appears to involve a well defined energy difference $\hbar\omega_\mathrm{fi}$. However there will always be some spread around this energy due to the uncertainty principle which can be stated as

$$\Delta E.t \geq \hbar, \tag{1.47}$$

where $\Delta E$ is the uncertainty in the measurement and $t$ is the time available to make the measurement. Since $t \sim 1/A_\mathrm{ul}$, $\Delta E \sim \hbar A_\mathrm{ul}$ which gives a finite width to any absorption line resulting from a given transition, this width is known as the natural width. Applying perturbation theory to the atomic system in question allows the shape of the frequency distribution to be found (e.g. Landau & Lifshitz 1982). This calculation shows the natural response shape to be a Lorentzian, given formally by

$$\frac{dp(\omega)}{d\omega} = \frac{\Gamma}{2\pi}\frac{1}{(\omega-\omega_\mathrm{fi})^2 + (\Gamma/2)^2}. \tag{1.48}$$

Where $\Gamma$ is the de-excitation rate of the upper level and is equal to $A_\mathrm{ul}$.

Often line profiles are dominated entirely by Doppler broadening. In this case the random motion of gas particles gives rise to a Doppler shift so any frequency of light

within a small range of $\nu_{\rm ul}$ can be absorbed. The function $\tilde{\phi_\nu} d\nu$ then equals $P(v)dv$, the fraction of atoms, assumed to have the same mass, $m$, whose radial velocities lie within the range $dv$ of $v$. In thermal equilibrium the system has a Maxwellian distribution of velocities

$$\tilde{\phi_\nu}(\Delta\nu) = \lambda P(v) = \frac{\lambda}{\sqrt{\pi} b} e^{-\left(\frac{v}{b}\right)^2}, \tag{1.49}$$

where $\lambda$ corresponds to the wavelength associated with the absorption and $b$ is known as the Doppler parameter, which is given by

$$b = \left(\frac{2k_{\rm B}T}{m}\right)^{1/2}. \tag{1.50}$$

The absorption feature obtained from a Maxwellian distribution of velocities is known as a Voigt profile. This is just the absorption profile seen when the optical depth is given by the Voigt function. This function is found at each point by the convolution of a Gaussian, corresponding to the motion of the gas, and a Lorentzian centred at the point of interest, corresponding to the natural width. This results in an expression for the Voigt function which is given by

$$\frac{dp(\omega)}{d\omega} = \frac{\Gamma}{2\pi^{3/2} b} \int_{-\infty}^{\infty} \frac{e^{-v^2/b^2}\, dv}{(\omega - \omega_{\rm fi}(1 + v/c))^2 + (\Gamma/2)^2}. \tag{1.51}$$

Typically the width of the Gaussian far exceeds the natural width. The centres of Voigt profiles are often dominated by the motion of the gas and thus look Gaussian. However in high column density systems the Lorentzian will dominate the profile at the edges, leading to the presence of prominent damping wings.

When the true line profile cannot be fully resolved, a parameter known as the equivalent width is often used, defined by

$$W_\lambda \equiv \int \left[1 - \frac{I_\nu}{I_\nu(0)}\right] d\lambda = \frac{\lambda^2}{c} \int (1 - e^{-\tau_\nu})\, d\nu. \tag{1.52}$$

The integral extends over the line and $\lambda$ represents the wavelength at the centre of the line. $W_\lambda$ measures the fraction of the energy removed from the spectrum by the absorption line. It represents the width a line which extends to 0, centred on $\lambda$, would

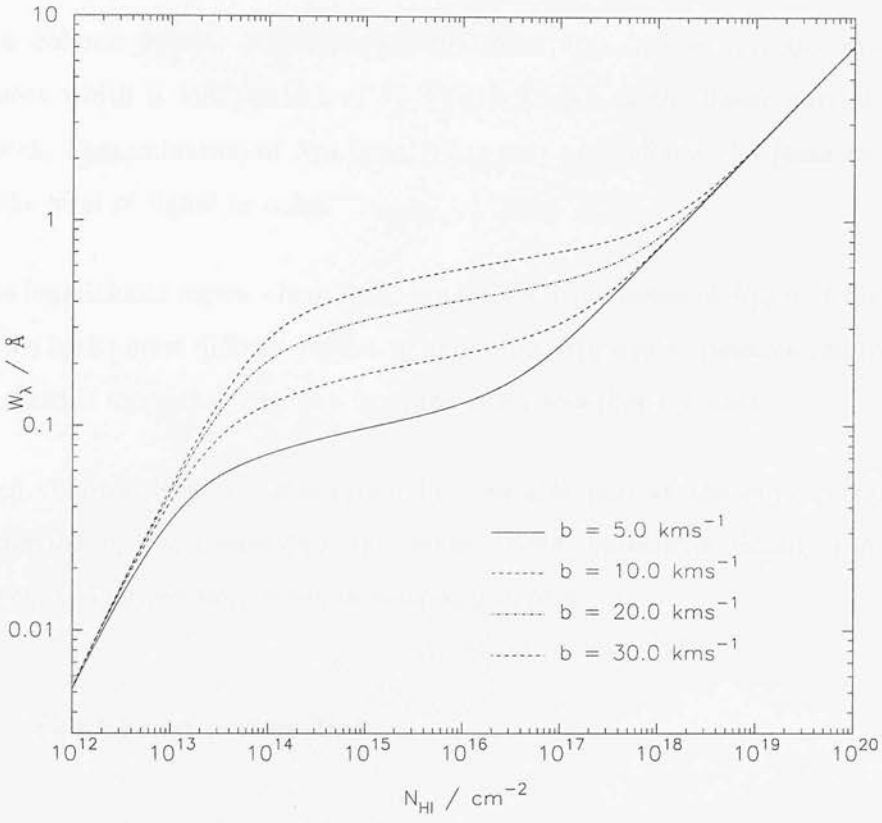


**Figure 1.3:** The curve of growth, or relation between Doppler parameter and equivalent width, for the Ly-$\alpha$ transition. There are three distinct regions. In the two extremes of column density the equivalent width has a single value irrespective of the value of the Doppler parameter. Conversely the equivalent width for features with intermediate column densities are highly dependent on the value of the Doppler parameter.

have if it removed the same fraction of energy as the actual absorption line. It can be readily seen that $W_\lambda$ is unaffected by changes in resolution, on the condition that each individual feature can still be resolved.

### 1.3.3 The curve of growth

The curve of growth is the name given to the relationship between the equivalent width of a line and its column density for different values of the Doppler parameter. The curve of growth for the Ly-$\alpha$ transition, for $b$ = 5, 10, 20 and 30 km s$^{-1}$, is shown in figure 1.3. As seen on the diagram there are three distinct parts to the curve:

Low column density - in this case the absorption line is optically thin and the equivalent width is independent of $b$. This is known as the linear part of the curve of growth. Determination of $N_{\rm HI}$ from $W_\lambda$ is easy and reliable, for features identified, given the level of signal to noise.

The logarithmic region - here there is a strong dependence of $N_{\rm HI}$ on $b$ for any given $W_\lambda$. This is the most difficult regime to determine $N_{\rm HI}$ and $b$. Reliable results can only be obtained if more than one line from the same absorber are used.

High column density - absorption lines in this part of the curve of growth are characterized by prominent damping wings. Here the column density can be found very accurately given $W_\lambda$, which is independent of $b$.

### 1.3.4 The Lyman-$\alpha$ transition

The Ly-$\alpha$ transition refers to the change from the first excited state to the ground state in the Hydrogen atom. This accounts for the vast majority of light absorbed by the IGM in quasar spectra as will be discussed in the following section. Applying the theory of quantum mechanics to a Hydrogen atom allows the calculation of both the restframe wavelength of this transition, 1216 Å, and of an expression for the optical depth (as demonstrated in appendix A).

$$\tau_\nu = 1.11 \times 10^{-2} N_{\rm HI} \phi_\nu. \tag{1.53}$$

## 1.4 The Lyman-$\alpha$ forest

Studies in astrophysics have a range of epochs which can theoretically be observed. Galaxy and cluster surveys offer information about the Universe today ($z \simeq 0$), whilst the CMB provides a view of the very early Universe ($z \sim 1000$). However the intervening epochs are not so easy to observe. As redshift increases the apparent magnitudes of objects tends to increase and they become more difficult to observe, so at high $z$ many things are impossible to see with current telescope technology. However very luminous objects which are still observable at high redshift ($z > 5.0$) such as quasars can indirectly offer information about the intervening redshifts. Although quasars are a huge area of study in their own right with many interesting features, here these issues will be ignored, and they will be regarded simply as high redshift bright light sources. Lynds (1971) discovered that quasar spectra contained hundreds of absorption lines from 'clouds' of neutral hydrogen gas, which comprise the IGM. Each cloud absorbs light at the wavelength of the redshifted Ly-$\alpha$ transition. These systems appeared to be uniformly distributed over the complete range between the lowest observable redshift up to that of the quasar being studied. This entire system of clouds has become known as 'The Ly-$\alpha$ Forest'.

Historically the Ly-$\alpha$ forest has been analysed in terms of blended line profiles. In this process different cloud profiles are fitted, to match the observed continuum, in order to gain information about the cloud system. Each cloud can be characterized by three parameters, its redshift, column density and Doppler parameter. Although this process of fitting lines can sometimes be degenerate, the technique has enjoyed considerable success. This picture has now been superceded by what is known as 'the cosmic web'. Here rather than there being discrete clouds along a line of sight there is an evolving network of filaments and sheets. The clouds, previously thought to be discrete, are simply the most overdense parts of the web. An illustration of this is shown in figure 1.4 which shows an $N$-body simulation of the Ly-$\alpha$ forest (Zhang *et al.* 1998).

The IGM is thought to be highly ionized (Gunn & Peterson 1965), with a neutral component 'the cosmic web' of predominately hydrogen but with some metals present.

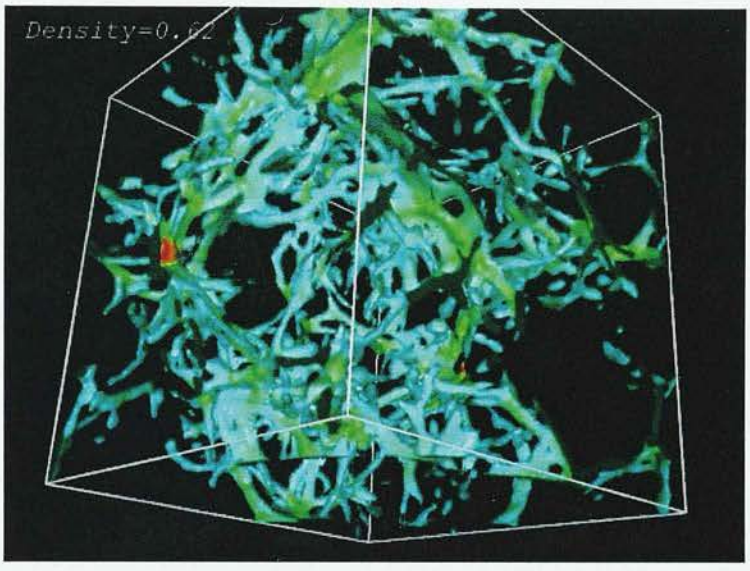


**Figure 1.4:** An $N$-body simulation by Zhang *et al.* (1998) showing the structure of neutral hydrogen in the Ly-$\alpha$ forest. Density variations are shown in colour.

Despite this picture of the IGM, treating the Ly-$\alpha$ forest as a set of discrete clouds is a good approximation, and remains a valuable method of analysis. There follows a brief review of some broadly accepted results of previous work on the Ly-$\alpha$ forest.

### 1.4.1 Flux decrement and mean optical depth

One observable property of the forest is a measure of the mean fraction of the quasars' continuum that is absorbed, $D$, or alternatively the mean optical depth, $\bar{\tau}$. These quantities are defined as

$$D \equiv \left\langle 1 - \frac{f_{\rm obs}}{f_{\rm cont}} \right\rangle = \langle 1 - e^{-\tau} \rangle \equiv 1 - e^{-\bar{\tau}}, \tag{1.54}$$

where $f_{\rm obs}$ and $f_{\rm cont}$ are the observed and continuum flux respectively. Although this may seem like a straightforward calculation there are some complexities. Firstly the quasar continuum flux is never known precisely and has to be estimated, therefore introducing errors when calculating the absorption. Also some of the absorption seen may be due to metal lines, which are hard to identify so are often wrongly included. Typically values are quoted for $D_A$, the flux decrement between Ly-$\alpha$ and Ly-$\beta$ (rest wavelength 1026 Å) transitions. Quoted values vary from $D_A(z = 3) = 0.36$ (Press,

Rybicki & Schneider 1993) to $D_A(z=3) \simeq 0.2$ (Zuo & Lu 1993).

### 1.4.2 Evolution with redshift

As the Ly-$\alpha$ forest is studied over a huge redshift range it is natural to expect to see evolution in the forest with redshift due to changing cosmological conditions. However the increase seen in the number of Ly-$\alpha$ absorption lines as $z$ increases is not consistent with a non-evolving cloud population. Observationally a strong $z$ dependence is seen with high redshift quasars showing very much more absorption than their low redshift counterparts. This dependence can be expressed in the form of a power law

$$\frac{dN}{dz} = \left(\frac{dN}{dz}\right)_0 (1+z)^{\gamma}. \tag{1.55}$$

Here the value of the first term on the right hand side is dependent on the column densities or equivalent widths of the clouds being considered. Calculating the exponent by line counting is a difficult task and a broad range of values have been quoted, some as low as $\gamma = 1.7 \pm 1.0$ for lines where the rest $W_\lambda > 0.2$ Å(Atwood, Baldwin & Carswell 1985). The authors find that the value of $\gamma$ is dependant on $W_\lambda$ but not by more than $\pm$ 0.5, as long as there are enough lines in the sample to overcome small number statistics. At the other extreme, values as high as $\gamma = 2.9 \pm 0.3$ (Cooke, Espey & Carswell 1997) have been calculated, where all values of $W_\lambda$ were included. Methods of finding the exponent without directly counting lines suggest a high value, $\gamma = 2.46 \pm 0.37$ (Press, Rybicki & Schneider 1993), or consider a broken power law to be a better fit to the data (Zuo & Lu 1993). It is now clear observationally that a single power law does not adequately describe the entire range of redshift from zero up to the highest quasar redshifts. Observations with the Hubble Space Telescope reveal that the steep decline in cloud number with decreasing redshift does not extend to zero, instead a much flatter evolution for $z < 2$ is seen (Impey *et al.* 1996). High redshift work $z > 4$ suggests the evolution accelerates with $\gamma$ becoming as high as 5.5 (Williger *et al.* 1994).

### 1.4.3 The column density distribution function

It was suggested (Carswell *et al.* 1984) that the number of clouds per unit $N_{\rm HI}$ interval fitted a power law distribution

$$\frac{dN}{dN_{\rm HI}} \propto N_{\rm HI}^{-\beta}, \quad \beta = 1.68, \quad 13 < \log N_{\rm HI} < 15. \tag{1.56}$$

Results from high column density surveys were studied and the distribution in these systems was found to be well fitted by the above relation with $\beta = 1.5$ (Tytler 1987). Work with Keck spectra seems to show that the power law can be extended over a range of ten orders of magnitude from $10^{12}$ to $10^{22}$ cm$^{-2}$ (Kirkman & Tytler 1997). It should be noted that these results are obtained assuming that the blending of the low column density lines is being handled correctly (Hu *et al.* 1995).

Evidence for departure from this single power law has been presented (Meiksin & Madau 1993; Giallongo *et al.* 1993) and cannot be discounted. Departure from a single power law could explain why individual high resolution spectra seem to suggest high values of $\beta$ (Rauch *et al.* 1992; Atwood, Baldwin & Carswell 1985).

### 1.4.4 Doppler parameter distribution

The width of an absorption feature can lead to a measurement of $b$, the Doppler parameter. Typically values of $b$ fall between about 15 - 45 km s$^{-1}$ with a median value of about 30 - 35 km s$^{-1}$ (Carswell *et al.* 1991; Rauch *et al.* 1992).

There has been some speculation as to a correlation between column density and Doppler parameter, where high values of $N_{\rm HI}$ correspond to high values of $b$. Initially this was explained as the effect of noise in the spectra distorting $b$ measurements leading to underestimation. Despite this, high resolution spectra seemed to suggest a correlation does exist where a minimum value of $b$ has a weak dependence on $N_{\rm HI}$ (Hu *et al.* 1995). Kirkman & Tytler (1997) also found a similar trend which is well fitted by

the expression

$$b_{\rm min} = 14 + 4 \times \log \left( \frac{N_{\rm HI}\ {\rm cm}^{-2}}{10^{12.5}} \right)\ {\rm km\ s}^{-1}. \tag{1.57}$$

### 1.4.5 The equation of state

The physical processes governing the behaviour of the IGM are thought to be well understood. $N$-body simulations offer results in close agreement with observations (e.g. Cen *et al.* 1994; Zhang, Annios & Norman 1995; Hernquist *et al.* 1996; Zhang *et al.* 1998) and so the essential physics are thought to be well captured. These simulations also demonstrate that the majority of the evolution of the Ly-$\alpha$ forest can be attributed to the effects of universal expansion. The key physical processes are heating by photoionization coupled with cooling from radiative processes and adiabatic expansion.

The temperature-density relation was studied semi-analytically by Hui & Gnedin (1997). A model was used where the density evolution was found from the Zel'dovich (1970) approximation, while the thermal and chemical evolution were solved numerically. This model was then tested against full hydrodynamic simulations and found to be in good agreement for the low densities found in the IGM. The physical processes occurring in the IGM lead to a tight correlation between temperature and baryonic density given as

$$T = T_0 (1 + \delta)^{\gamma - 1}. \tag{1.58}$$

The ranges of these parameters shown to be physically reasonable (Hui, Gnedin & Zhang 1997; Hui & Gnedin 1997) are, $10^3\ {\rm K} < T_0 < 10^{4.5}\ {\rm K}$, and $1.2 < \gamma < 1.7$. The values of these parameters depends on the cosmology and reionization history of the Universe. Effects such as shock heating and feedback from star formation are thought to be relatively unimportant. The most likely candidates to provide the ionizing radiation are quasars and young stars, so the evolution of quasars and the star formation history of the Universe are key factors in understanding the reionization history, and thus the evolution of the IGM.

### 1.4.6 Morphology

For over a decade now the properties of the Ly-$\alpha$ forest have been usefully predicted by hydrodynamical simulations. Furthermore irrespective of the cosmological parameters used, these simulations give rise to a generic picture of the Ly-$\alpha$ forest (e.g. Rauch 1998). At high redshifts the evolution of the IGM is largely governed simply by the Hubble expansion and change in the ionization rate of the gas. At these times the absorption is dominated by sheet-like structures. These sheets are similar to small pancakes arising in the Zel'dovich approximation, where collapse of triaxial structures along the shortest axis is predicted, but are subtly different. The sheets are actually present in the early geometry of the density fluctuations in 3 dimensions and the features are merely sharpened by non-linear dynamics (e.g. Bond *et al.* 1996).

At lower redshifts absorption arises not only from sheets but also filament structures. Low column density systems ($N_{\rm HI} < 10^{14}$ cm$^{-2}$) are associated with the sheet-like structures whose characteristic proper length scale is typically between about one hundred kpc and one Mpc.

Higher column density features are seen to correspond to filamentary structures. These filaments are of relatively uniform thickness ($\sim$ 40 - 100 proper kpc) and extend over distances of many megaparsecs. These filaments are typically found at the intersection of the sheets discussed above, resulting in a morphology of interlinked systems, leading to the terminology 'cosmic web' (first introduced by Bond *et al.* 1996) being used to describe the IGM. As the column density increases still further the geometry of the absorber tends to become more spherical. As column densities reach about $10^{16}$ cm$^{-2}$ the absorbers are almost completely spherical and correspond to minihalos. On scales of many Mpcs the halo systems tend to align along filaments in a similar way to the distribution of dark matter in $N$-body simulations.

A key result from the simulations which matches the observations well is the broken power law found for the evolution of the number of absorbers with redshift (e.g. Muecket *et al.* 1996) who found that for systems with $N_{\rm HI} > 10^{14}$ cm$^{-2}$, $\gamma \sim 0.6$ for $z < 1.5$ switching to $\gamma \sim 2.6$ for $1.5 < z < 3.0$. This break arises as the struc-

tures dominating the absorption change with time. The sheets which dominate high redshift absorption are expanding with time leading to a decrease in column density. As these systems drop below the detection threshold, the filaments begin to dominate the absorption. Although the column densities associated with these systems are also dropping, they were originally higher and so take longer to drop below the detection threshold.

## 1.5 Measuring the mass density power spectrum

This section contains a brief review of the most important ways in which the mass power spectrum can be determined from observational data.

### 1.5.1 Galaxy surveys

The mass power spectrum can be found from galaxy surveys (e.g Feldman, Kaiser & Peacock 1994), by assuming the galaxies offer Poisson indicators of mass. To do this the weighted galaxy distribution is converted to an overdensity distribution. The weighting is introduced so that small volumes containing large numbers of galaxies do not unduly affect the uncertainty of the power spectrum. The conversion to an overdensity distribution is done by subtracting off a synthetic random catalogue. Once this has been done the effect of the window function must be considered. Since the analysis attempts to measure a random field of infinite extent with a limited survey volume, what is actually measured, $P_{\rm obs}$, is a convolution of the actual power spectrum and the window function, $W_k$,

$$P_{\rm obs} = P_{\rm true} * |W_k|^2. \tag{1.59}$$

Therefore to obtain an accurate measurement of the power spectrum the geometry of the survey must be taken into account, and this deconvolution performed.

The most recent application of this method was performed by Percival *et al.* (2001) with a sample of $\sim$ 147 000 galaxies from the 2dF Galaxy Redshift Survey (2dFGRS). It was also shown, by comparisons with simulations, that the power spectrum is coincident to the linear density perturbations on certain scales ($0.02 \lesssim k \lesssim 0.15\ h{\rm Mpc}^{-1}$).

However there are difficulties associated with using galaxy surveys for this work. For example the discrete nature of galaxies gives rise to shot noise, though the large number of galaxies used in Percival *et al.* (2001) means the effects should be negligible. Observational difficulties such as redshift-space distortions can also influence the data. The radial position of a given galaxy is inferred by its redshift. To do this it is assumed

that the redshift comes about solely from the Hubble flow, although any peculiar velocity that a galaxy has will contribute to the redshift measurement, therefore introducing errors when inferring its distance.

More importantly the process of galaxy formation is not well understood and the assumption that galaxies offer a Poisson sampling of the mass field is not necessarily a good one. This introduces an uncertain, possibly scale dependent, bias parameter between the mass and galaxy distributions. However recent work (Verde *et al.* 2001; Lahav *et al.* 2001) suggests that the 2dFGRS does, on certain scales, offer a Poisson sampling of the mass field and thus this bias parameter is very close to unity. Therefore the recovered galaxy power spectrum should match the underlying mass power spectrum well.

### 1.5.2 Gravitational lensing

One of the predictions made by general relativity is that the path of light is affected by the local mass distribution. This effect is known as gravitational lensing which is a rapidly developing field which can be applied to determine cosmological parameters. For the study of large scale structure the weak lensing regime is important. Here rather than a large deviation in the path of light, caused by an isolated concentration of mass, being of interest, the small deviations caused by the large scale structure are studied. This effect is known as cosmic shear and a simulation by Colombi (2000) of this is shown in figure 1.5.

The correlations of shapes in observed distant galaxies can be used to infer the mass distribution through which the light has passed. The huge advantage offered by this method is that it probes the mass, regardless of its physical nature, directly. However there are also considerable drawbacks. Any residual distortion due to cosmic shear is intrinsically very small and the systematic errors in the observations are generally far more significant. Therefore very careful indentification and removal of the systematics is crucial. Furthermore, at the current time, the quality and quantity of the data is insufficient to impose stringent constraints on the mass power spectrum.

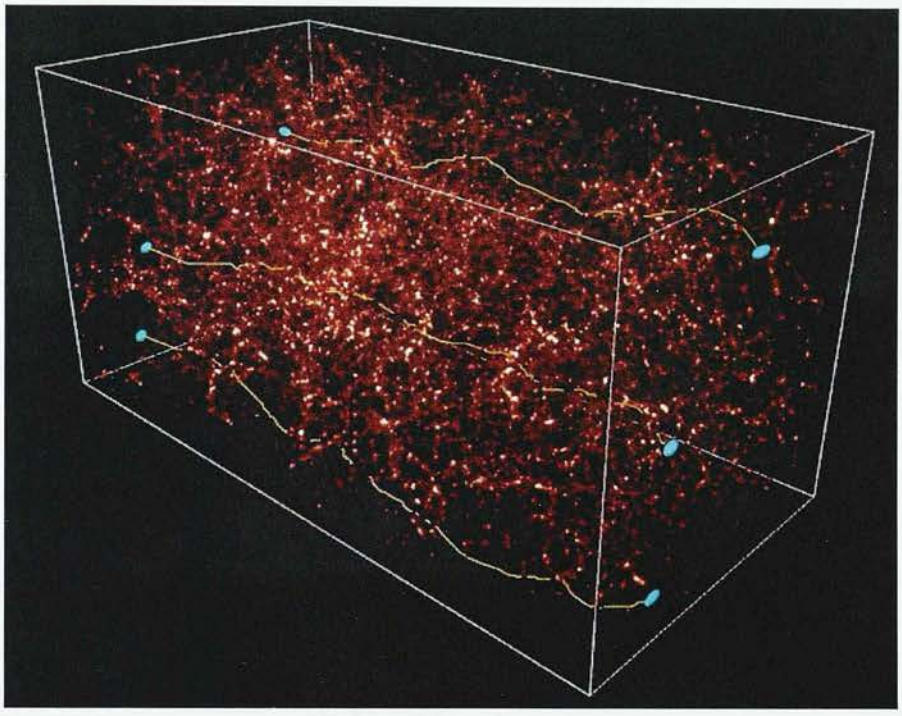

**Figure 1.5:** A simulation of the effect of cosmic shear on light from distant objects by Colombi (2000). The matter concentration of a region is indicated by its brightness, while the blue dots represent galaxies. Due to gravitational lensing the path of the light emitted is determined by the mass distribution through which it passes. The resulting distortions in the shapes of the observed galaxies may be used to infer the large scale structure of the Universe.

However, in the future, as the data improves and the analysis techniques are refined this field should become one of the most powerful probes of the matter distribution in the Universe.

### 1.5.3 The cosmic microwave background

The CMB allows the opportunity to study the correlations of inhomogenities in the very early Universe. The detected temperature inhomogenities (Smoot *et al.* 1982) can be used to set constraints on large scale density variations. The temperature power spectrum of the CMB can be related to the mass power spectrum since the temperature and mass perturbations correspond to each other.

The dominant large scale effect was originally derived by Sachs & Wolfe (1967)

after whom it is named. The perturbations in gravitational potential at the surface of last scattering has two effects. The first is that photons travelling from within a potential well are redshifted as they climb out. A second, more subtle, effect is that the gravitational potential perturbations cause varying amounts of time dilation, thus causing the age, and the temperature, of the Universe at the surface of last scattering to vary. These effects combine to give the expression

$$\frac{\delta T}{T} = \frac{\delta \Phi}{3c^2}. \tag{1.60}$$

The gravitational potential, $\Phi$, in the above expression can be related to be mass density via Poisson's equation

$$\nabla^2 \delta \Phi_k = 4\pi G \rho \delta_k. \tag{1.61}$$

Other processes such as the Doppler effect of photons scattering off moving plasma and adiabatic expansion must also be considered. The inferred mass power spectrum is dependent on several factors such as the Hubble constant, $\Omega$, the baryon fraction and the contribution of tensor fluctuations (White and Bunn 1996). For a given set of cosmological parameters a matter power spectrum can be calculated using an algorithm called CMBFAST (Seljak & Zaldarriaga 1996).

### 1.5.4 The Ly-$\alpha$ forest

This section contains a review of the papers which lay the foundation upon which the work in this thesis is based.

#### Croft *et al.* (1998)

This paper argues that the Ly-$\alpha$ forest could provide an elegant way to recover the shape of $P(k)$. It is then argued that once the shape is known, the amplitude can be determined by evolving a number cosmological simulations with varying amplitudes of $P(k)$. The correct amplitude is then the amplitude corresponding to the simulation which successfully reproduces the observed power spectrum of transmitted quasar flux.

Is is argued that this method circumvents many of the problems faced when measuring $P(k)$ by different methods, for example, there are none of the bias parameters or shot noise problems associated with calculating the power spectrum from galaxy surveys. It also has the advantage of probing the power spectrum at high redshifts. The effects of non-linearity at high redshift are less pronounced than at lower redshifts, and thus a direct measurement of the linear regime is possible for a wider range of scales.

The method for recovering the shape of the power spectra exploits the fact that the state of the gas in the IGM is largely governed by well understood processes, namely radiative cooling coupled with a further adiabatic cooling due to the expansion of the Universe, against heating due to photoionization by the ultraviolet (UV) background. This leads to a situation where the optical depth of the Ly-$\alpha$ forest is directly related to the underlying baryon density (Bi, Ge & Fang 1995; Hui, Gnedin & Zhang 1997). This is implied since the temperature seems very well fitted by a power law relation

$$T = T_0 \left( \frac{\rho_b}{\bar{\rho}_b} \right)^{\alpha}, \tag{1.62}$$

where $\rho_b$ is the baryonic density (Hui & Gnedin 1997). This relation only holds in the regime where $\rho_b/\bar{\rho}_b$ is small. When the $\rho_b/\bar{\rho}_b \simeq 10$ regime is reached there is a range of temperatures for any given density, depending on the extent that the gas has been shocked (Meiksin 1994; Theuns *et al.* 1998). The parameters $T_0$ and $\alpha$ depend on the reionization history of the Universe and the spectral shape of the UV background. The density regime of the Ly-$\alpha$ forest is such that any pressure gradients present should be small compared to gravitational forces. This leads to an environment where the gas traces the structure of the dark matter, meaning $\rho_b \propto \rho$. Since the gas is highly ionized and dominated by photoionization it follows that

$$n_{\rm HI} \propto \Gamma^{-1} \alpha(T) n^2, \tag{1.63}$$

where $n_{\rm HI}$ and $n$ represent the densities of neutral hydrogen and total gas, $\Gamma$ is the photoionization rate of the UV background and $\alpha(T)$ is the recombination coefficient. Inputting an expression for $\alpha(T)$ which is proportional to $T^{-0.7}$ (Osterbrock 1989) produces a tight relationship between optical depth and baryon density where by

$$\tau \propto \rho_b^2 T^{-0.7} \propto \rho_b^{\beta} \quad \text{where} \quad \beta = 2 - 0.7\alpha. \tag{1.64}$$

Although the argument neglects many effects, for example shock heating, collisional ionization and peculiar velocities it is claimed these do not affect the relation significantly. Given this relationship it becomes possible to map a continuous density field from the continuum of a quasar spectrum. This process cannot be achieved by direct inversion, firstly as some of the parameters are not known precisely and secondly because inferring an accurate value of $\tau$ from the flux becomes virtually impossible as the flux level approaches zero. Instead of attempting this inversion this paper proposes the use of Gaussianization. This is a process introduced by Weinberg (1992) which exploits the fact that the primordial density field is expected to a have a Gaussian probability distribution function (PDF). Although the density field evolves with time it tends to preserve its relative order, for example overdense regions remain overdense. Gaussianization exploits this to try and map a density field back to its original form. This process can be applied to quasar spectrum by ranking in order the normalized flux values and then assigning a density value which corresponds to a Gaussian PDF. The resulting density values obtained are only relative, and it is for this reason that only the shape of the power spectra can be recovered. Once the shape of $P_{1D}(k)$ has been recovered in this way, the three dimensional form, $P(k)$, can be obtained via differentiation (equation 1.34).

This procedure is applied to a quasar spectrum, but the statistical uncertainties are large since only one quasar is used. It is noted on scales below about $2\pi/k = 1.5\ h^{-1}$ Mpc (comoving) this method cannot be reliably applied as the effects of non-linear gravitational evolution, thermal broadening and peculiar velocities dominate. It is also questioned whether an upper limit for the reliability of recovery at large scales might be set by the discreteness of the sources that make up the UV background.

**Hui (1999)**

In this paper the affect of redshift-space distortions is considered in more detail. This problem arises since any comoving distances used are calculated from redshift information. Clearly any peculiar velocities will introduce an error when converting between these two parameters. In general these distortions affect the flux power spectrum to make it anisotropic. Therefore recovering the three dimensional mass fluctuation power

spectrum from the one dimensional flux power spectrum becomes a more involved process. This paper presents a modification to the Croft *et al.* method to account for this. A distortion kernel $W^{\mathrm{F}\rho}(k_{||}/k, k)$ is introduced so that the flux (denoted by subscript F) and mass power spectra are now related by the equation

$$P_{\mathrm{1D,F}}(k_{||}) = \frac{1}{2\pi}\int_{k_{||}}^{\infty} W^{\mathrm{F}\rho}(k_{||}/y, y)P(y)y\ dy, \tag{1.65}$$

where $k_{||}$ represents the wave vector along the line of sight and $k$, as usual, is the magnitude of $\mathbf{k}$. Recovery of the three dimensional mass fluctuation power spectrum now requires inverting equation 1.65. Hui argues that for practical purposes it is necessary to truncate the infinite vectors involved. However since $P(k)$ is in general a function which decays rapidly with $k$ at high $k$, a practical limit can be chosen for truncation which leads to a small error in the final answer. This paper shows that excluding the distortion kernel, which corresponds to $W^{\mathrm{F}\rho}$ = constant, tends to lead to an underestimation of the steepness of the mass power spectrum.

**Croft *et al.* (1999)**

This paper applies the method from Croft *et al.* (1998) to a sample of 19 quasar spectra. A result for the measurement of $P(k)$ is given between the range $2\pi/k \sim 450 - 2350$ km s$^{-1}$. The results are given in the form of the equation

$$P(k) = P_{\mathrm{p}}\left(\frac{k}{k_{\mathrm{p}}}\right)^{n}, \tag{1.66}$$

where $n = -2.25 \pm 0.18$ with an amplitude $\Delta^2(k_{\mathrm{p}}) = 0.57^{+0.26}_{-0.18}$, for the chosen pivot wavenumber $k_{\mathrm{p}} = 0.008$ (km s$^{-1}$)$^{-1}$.

It is also demonstrated that the effect of artificial correlations produced by variations in the UV ionizing background are negligible.

### McDonald *et al.* (2000)

In this paper the usefulness of the process of Gaussianization is called into question. It is claimed that this process can amplify noise effects, and that its role in improving the recovery of the mass fluctuation power spectrum is questionable. The authors therefore prefer to omit this step and first calculate the one dimensional power spectrum of the transmitted flux $P_{\rm 1D,F}(k)$ with the Lomb periodogram technique (Press *et al.* 1992). They then use numerical simulations to study the relationship between the flux power spectrum and the mass density power spectrum. They relate the flux and mass density power spectra with the formula

$$P_{\rm 1D,F}(k) = A_{\rm F}\frac{e^{-k^2 v_c^2}}{2\pi}\int_k^\infty P(y)y\,dy, \tag{1.67}$$

where the three dimensional power spectrum of mass density perturbations is given by a functional form, with modified parameters, which is known to fit a cold dark matter model well (e.g. Bardeen *et al.* 1986). The authors note that the equation used agrees well with $P_{\rm 1D,F}(k)$ for the low $k$ regime ($k < 0.04$ (km s$^{-1}$)$^{-1}$). Although it is noted that the 3D Gaussian smoothing used in this paper provides a better fit for the power spectrum on small scales than a 1D Gaussian smoothing, the 3D smoothing does not return the correct value of the spectral index $n$ of the model used in the simulation. Fits are obtained for the observations to this formula treating $A_{\rm F}$, $v_c$ and $n$ as free parameters. Applying this method to a flux power spectrum from simulations they find an error in recovery of the spectral index of $\Delta n = -0.02 \pm 0.07$ and assume that the same offset should be seen in the observations. They use this as a correction to the observed slope and add the error in quadrature. This method is applied to 8 quasar spectra, with the slope of $P(k)$ measured to be $n = -2.55 \pm 0.10$ with an amplitude of $\Delta^2(k) = 0.72 \pm 0.09$, at $k = 0.04$ (km s$^{-1}$)$^{-1}$.

### Feng & Fang (2000)

The usefulness of the Gaussianization process is again under scrutiny in this paper. The authors claim that the non-Gaussian behaviour of the flux still remains even after this process is applied. Any recovered mass power spectrum is shown to be systematically lower than the initial mass power spectrum, on certain scales, if this is the case. A

modification to the usual Gaussianization method is presented which overcomes this problem. The method involves using a discrete wavelet transform in conjunction with Gaussianization and a summary of the prescription given is:

1) Use the conventional Gaussianization technique to recover the mass field. Normalize this field so simulations reproduce the observed flux limits.

2) Calculate the wavelet function coefficients of the recovered density field on each scale.

3) For each scale make a Gaussian mapping of the coefficients to produce unnormalized wavelet function coefficients.

4) Normalize these new coefficients by requiring that the variance matches that of the original coefficients.

5) Randomize the spatial sequence of the Gaussianized coefficients.

6) Use the coefficients to reconstruct the mass density field to the scale given by the resolution of the flux.

**Zaldarriaga *et al.* (2000)**

In this paper the authors, rather than attempting any inversion technique prefer to simply compare the flux power spectrum with predictions for a range of cosmological models. The Ly-$\alpha$ simulations they use are generated with a PM code (e.g. Meiksin & White 2001), with the gas properties inferred from the dark matter distribution, via simple scaling relations.

The results of this paper indicate that the spectral index is poorly constrained when inferring the mass density field. The reason for this is that the smoothing scale linking the baryonic and dark matter distribution is left as a free parameter. Essentially differences in the power on various scales can be accounted for as being the result of either the value of the spectral index, or by a change in the temperature of the IGM. It is suggested that much of this uncertainty could be accounted for by the use of a

more sophisticated hydro-PM code (Gnedin & Hui 1998). Although it is claimed that the degeneracy cannot be completely removed due to uncertainties of the reionization history.

**Croft *et al.* (2000)**

This paper applies a modified version of the original method to a sample of Keck spectra (30 HIRES and 23 LRIS spectra). The results of the recovered $P(k)$ presented are a spectral index of $n = -2.47 \pm 0.06$ and amplitude $\Delta^2(k) = 0.54^{+0.14}_{-0.12}$, at $k = 0.03$ $(\mathrm{km\,s^{-1}})^{-1}$. Modifications to the original method are included to account for errors arising from redshift-space distortions, non-linearity and thermal broadening. In this modified method they calculate the one dimensional flux power spectrum using the Lomb periodogram method. A three dimensional flux power spectrum $P_{\mathrm{F}}(k)$ is then defined such that it is the power spectrum of the three dimensional flux field that would have a line-of-sight power spectrum $P_{\mathrm{1D,F}}(k)$ if it were isotropic

$$P_{\mathrm{F}}(k) = -\frac{2\pi}{k}\frac{d}{dk}P_{\mathrm{1D,F}}(k). \tag{1.68}$$

This flux power spectrum is then related to the mass density power spectrum with the assumption that they are related by a scale dependent bias factor such that

$$P_{\mathrm{F}}(k) = b^2(k)P(k). \tag{1.69}$$

The calculation of $b(k)$ is performed with 'normalizing simulations', which rely on the fact that the form of $P(k)$ is in good agreement with a low density ΛCDM model, which can therefore be used in these simulations. The bias parameter is then calculated using the equation

$$b(k) = \left[\frac{P_{\mathrm{F}}^{\mathrm{sim}}(k)}{P^{\mathrm{sim}}(k)}\right]^{1/2}. \tag{1.70}$$

The simulation used is the one for which the value of $\Delta^2_{\mathrm{F}}(k)$ most closely matches the observations. Any systematic uncertainties in the recovered $P(k)$ will now relate directly to the bias parameter. They argue that if the assumption of Gaussian fluctuations is correct then the shape and amplitude of $P(k)$ will be the important feature to

determine the bias parameter for any cosmological model. However the authors then claim the dependence of $b(k)$ on the shape of the power spectrum is weak since what is computed in the simulations is $b^2(k) = P_{\rm F}(k)/P(k)$ rather than $P_{\rm F}(k)$ itself. Therefore it is claimed that the uncertainty in $b(k)$ associated with adopting a ΛCDM model for the normalizing simulations should be small. An extensive study is presented showing that reducing the uncertainties is limited by a lack of knowledge about the parameters $T_0$, $\alpha$ (see equation 1.62) and $\bar{\tau}$. It is argued that as these parameters are determined more precisely the errors in this modified method will be reduced.

**Jamkhedkar *et al.* (2001)**

The problem of correctly normalizing the flux power spectrum is addressed in this paper. Since the effect of mean optical depth and continuum fitting affect the inferred amplitude of the mass power spectrum, the authors present a method of independent normalization. This is done with a discrete wavelet transform of the flux spectrum.

First the flux is decomposed into the background and fluctuation information as

$$F(\lambda) = \bar{F}(\lambda) + \bar{F}(\lambda)\delta(\lambda). \tag{1.71}$$

It is then argued that a wavelet analysis has the advantage that if the background flux is correlated with the fluctuations this effect is quantified by the spatial information. This technique also allows simultaneous estimation of the normalization of the background and the calculation of the power spectrum. The mathematical formalism of this technique is omitted for reasons of conciseness.

**Gnedin & Hamilton (2001)**

This paper investigates systematic errors of inferring the mass power spectrum from the Ly-$\alpha$ forest. They find the main sources of error concern the assumed equation of state and mean optical depth.

Recent observations (Schaye *et al.* 2001; McDonald *et al.* 2001) suggest somewhat

different values of the parameters $T_0$ and $\gamma$ than those used by Croft *et al.*. Although finding the effect of altering $\gamma$ is small, the amplitude of the mass power spectrum is reported to be strongly dependent on $T_0$. This effect is quantified by the equation

$$P_{\rm lin}(k, T_0) = P_{\rm lin}(k, 2 \times 10^4 \text{ K}) \left( \frac{2 \times 10^4 \text{ K}}{T_0} \right). \tag{1.72}$$

Similarly they find that the value of the mean optical depth used will also have an effect on the amplitude of the recovered power spectrum such that

$$P_{\rm lin}(k, \bar{\tau}) = P_{\rm lin}(k, 0.349) \left( \frac{0.349}{\bar{\tau}} \right)^{0.75}. \tag{1.73}$$

These two effects therefore mean the errors in the amplitude of any recovered mass power spectrum are highly effected by the uncertainties in these parameters. However other effects such as inhomogeneities in the ionizing background, shock heating of gas and the form of the prior mass power spectrum are argued to be less significant than the random errors.

**Zaldarriaga *et al.* (2001)**

This paper discusses the effects of non-linearity on the recovery of the linear mass power spectrum. It is argued that the non-linear effects drive the power spectrum to a power law of $k^{-1.4}$, regardless of the initial conditions. The authors therefore claim that previously quoted uncertainties in the shape of the linear power spectrum are underestimated.

It is argued that the important quantity is the spectral index in the weakly non-linear scale. For a power spectrum which is initially a power law (and, it is claimed, a good approximation of CDM spectra), one loop perturbation theory gives the non-linear power spectrum as

$$P(k) = P_{\rm lin}(k) \left[ 1 + \alpha(n) \left( \frac{k}{k_{\rm nl}} \right)^{n+3} \right]. \tag{1.74}$$

Where $\alpha(n)$ monotonically decreases with $n$. The critical value is $n_c = -1.4$ since for a spectral index below this value the non-linear spectrum is less steep than the

linear one. When $n > -1.4$ the non-linear corrections are negative, and non-linear power spectrum grows more slowly than the linear case leading to a spectrum steeper than the linear one. When $n \simeq -1.4$ the power spectrum retains its linear shape. Therefore it is claimed that all cases are driven towards the critical index regardless of their initial linear shape (Scoccimarro & Frieman 1996).

The authors claim this process occurs on scales of interest in the Ly-$\alpha$ forest at $z \sim 3$. For example in the case of a model with a low value for the shape parameter and large amplitude they claim the non-linear corrections have a strong effect at $k \sim 4 \times 10^{-3}$ (km s$^{-1}$)$^{-1}$. Therefore they conclude that the constraints on the initial linear power spectrum that can be set with the present data are poor.

# Chapter 2

# Monte Carlo simulations

This chapter contains analysis of $\Delta^2_{\rm 1D,F}(k)$ for rudimentary simulations, where spectra are modelled with blended Voigt profiles. This work is then discussed in the context of the difficulties it highlights when attempting to infer the mass field from quasar spectra. The power spectrum obtained from the positions of discrete objects are also considered, and applied to line centres of Ly-$\alpha$ absorption spectra. Results are presented for cases where the line centres are randomly placed and when they are clustered according to an underlying density field. Additionally the formalism used to generate the density fields used in the latter case is discussed in detail.

## 2.1 Monte Carlo simulations

### 2.1.1 The model

This set of simulations is one of the simplest that can be studied. A spectrum is generated simply by assuming that there is a set of Ly-$\alpha$ absorption lines which are randomly positioned in accordance with a given redshift distribution. Since the pixels used in these simulations are small (see below) it can be assumed that there is no variation in $dN/dz$ across each pixel, thus the probability of a Ly-$\alpha$ cloud being centred

in each pixel is given by

$$dp = \frac{dN}{dz}\Delta z, \tag{2.1}$$

where both quantities on the right hand side represent the local values corresponding to the pixel in question. A random number generator is then used, in conjunction with this probability, to determine which pixels contain the centres of absorption features. This probability, $dp$, is small enough that the chance of there being more than one absorption feature centred in each pixel can be neglected. For each line present a column density and Doppler parameter are then assigned from simple distributions. Once these parameters have been assigned the absorption effects are calculated by assuming each feature has a perfect Voigt profile. Complications such as noise have been ignored while discussion of the effects of peculiar velocities is saved until §3.6.

### Redshift distribution

The redshift distribution function used is simply the power law relation described in equation 1.55, which is assumed to be independent of $W_\lambda$. The exponent $\gamma$ can be chosen to have any value and the constant in the equation is then determined, in conjunction with the other line parameters, by imposing the mean optical depth (equation 1.54) in a small redshift interval to have a selected value. The simulations analysed in this chapter were chosen to have $\bar{\tau}(z = 3.0) \simeq 0.35$ in accordance with Meiksin, Bryan & Machacek (2001) and references therein.

### Column density distribution

The column density of each cloud is assigned in accordance with the distribution described by equation 1.56. The exponent $\beta$ can be chosen to have any value. A lower limit of column density must be set since extending this power law relation towards zero would result in the probability of obtaining a high column density line becoming negligible. The lowest column density that can be assigned is set to be $1 \times 10^{12}$ cm$^{-2}$, since lines weaker than this do not have a significant optical depth and thus below this

limit the distribution of $N_{\rm HI}$ cannot be measured directly. There is no constraint on the highest possible column density, but any sensible value of the exponent in the distribution function ensures that very high column density systems will always be exceedingly rare.

#### Doppler parameter distribution

The Doppler parameter for any feature is drawn at random from a truncated PDF with the distribution above the truncation limit matching a Gaussian distribution with a selected mean, $\bar{b}$, and standard deviation, $\sigma_b$. This is consistent with the distribution of Doppler parameters observed by Kirkman & Tytler (1997). The truncation value for each feature is set by a minimum allowable value of Doppler parameter which is dependent on its column density. Again this is in accordance with the observations reported in Kirkman & Tytler (1997), and the minimum required value is found using the expression given in equation 1.57. In the case when a Doppler parameter drawn from the Gaussian is less than $b_{\rm min}$ this value is simply ignored and a new value is assigned, this process is repeated until a suitable value of $b$ is obtained. It should be noted that due to the constraint on $b_{\rm min}$ the mean and standard deviation of the resulting Doppler parameters will not equal the input values, $\bar{b}$ and $\sigma_b$.

### 2.1.2 Range and resolution

Due to the simplicity of these simulations they allow a large number of spectra to be generated corresponding to any arbitrary redshift range. A reasonably small range is chosen so that when an underlying power spectrum is added the cosmological evolution can be ignored. The range of comoving $r$ corresponding to any fixed redshift interval is obviously dependent on cosmology. When real data are considered the uncertainty in the values of cosmological parameters means the distance coordinate $r$ cannot be directly computed. Therefore it is most convenient to work with the directly observable variable, the velocity $v$, and inverse velocity space for the wave vectors. Working with velocities also makes it possible to compare results from different cosmologies and other

published data. The simulations analysed in this chapter are chosen to have a redshift range of $z = 2.9 - 3.0$. Using a convention where $v = 0$ at the lowest redshift limit it is sensible to define a redshift relative to this

$$1 + z_{\rm rel} = \frac{1 + z_{\rm up}}{1 + z_{\rm low}}. \tag{2.2}$$

Now if $z_{\rm rel} \ll 1$ then the velocity is given by

$$\frac{v}{c} \simeq z_{\rm rel}. \tag{2.3}$$

Alternatively if $z_{\rm rel}$ is large enough to introduce substantial errors with this method then Hubble's law can be used to approximate the velocity using the equation

$$v = H(\bar{z}) r_{\rm p}, \tag{2.4}$$

where $\bar{z}$ is the mean redshift. However this too is only a reasonable approximation if the redshift interval is small. In cases where the redshift difference is large the use of velocity as a measure of distance breaks down, since on these scales the two are not linearly related. The above calculation of equation 2.4 for $z = 2.9 - 3.0$ gives a velocity range of $v = 0 - 7600$ km s$^{-1}$.

The resolution of these simulations has been set at 2 km s$^{-1}$ per pixel. This decision has been made to match observations made with the Keck telescope (e.g. Kirkman & Tytler 1997). Using this pixel size and spectra length means the data can be studied on scales of $k$ running from $8.0 \times 10^{-4} - 1.57$ (km s$^{-1}$)$^{-1}$.

### 2.1.3 Analysis of spectra

#### Flux power spectra

The spectra generated by the above method were then analysed in the following ways. First an FFT was taken of the flux and the resulting components used to calculate the flux power spectrum $P_{\rm 1D,F}(k)$. It was found that $P_{\rm 1D,F}(k)$ for individual spectra

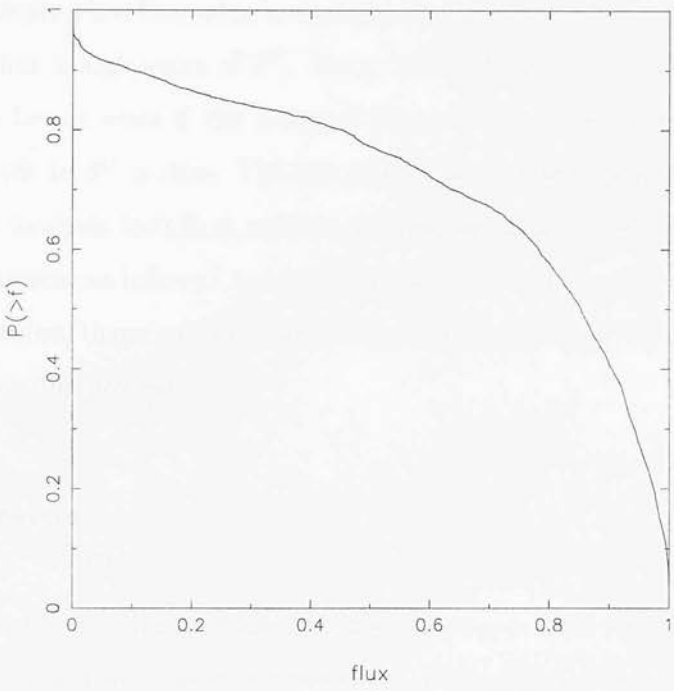


**Figure 2.1:** The probability of finding a flux level greater than a given value. The simplistic nature of these simulations leads a situation where large parts of the spectrum are very close to or equal to 0 or 1. This accounts for the sharp turn off seen at the extreme ends.

were very noisy, therefore a number of realizations were performed and the flux power spectra were averaged. This expression $\bar{P}_{\rm 1D,F}(k)$ was then fitted by a cubic spline to remove any remaining noise before being multiplied by $v_{\rm max}k/\pi$ to give $\Delta^2_{\rm 1D,F}(k)$. This process was performed for various values of the input parameters.

Additionally the effect of Gaussianization on the shape of the power spectrum was studied. This process is performed by assigning each flux point a $\delta^G$ value corresponding to a Gaussian PDF. First the cumulative probability distribution for the flux of a given spectrum is calculated (e.g figure 2.1). Now for the flux value in the $i^{\rm th}$ pixel, $f_i$, the associated probability $p_i$ can be assigned. A 'Gaussian' fluctuation value can then calculated using

$$p_i = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\delta_i^G} e^{-\frac{y^2}{2}}\, dy. \tag{2.5}$$

This $\delta^G$ value gives a relative measure of the density corresponding to a given flux level. For example a low flux value in a particular pixel has a high probability associated with it and thus a high value of $\delta^G$. Using this definition of $\delta^G$ requires a numerical truncation to be set since if the assigned probability is exactly equal to 0 or 1 then this corresponds to $\delta^G = \pm\infty$. The Gaussianization process was introduced by Croft *et al.* into the analysis to infer a relative underlying mass density from flux levels in a spectrum. However an inferred underlying mass density from a saturated flux region is a lower limit value, therefore this truncation is consistent with the physical motivation for introducing this process.

### Line power spectra

The power spectrum of line positions was also recovered. This technique is analogous to using galaxy surveys to recover the power spectrum which means, unfortunately, that many of the problems encountered in that field apply to this method as well. This will be discussed fully later. Here rather than the point objects of interest being galaxies they are the absorption line centres. Each pixel is assigned a value $\delta_\ell = n_i/N_\ell$ where $N_\ell$ is the total number of lines in the realization, and $n_i = 0$ or 1. Each pixel containing a cloud centre is set to $n_i = 1$, and all other pixels are assigned a value of $n_i = 0$. An FFT was then performed on this binary distribution and the power spectrum, labelled $P_{1\mathrm{D},\ell}(k)$, was calculated. As with the flux equivalent this was then used to calculate the dimensionless power spectrum $\Delta^2_{1\mathrm{D},\ell}(k)$. When calculating the power spectrum for a discrete set of objects the shot noise must be accounted for. The Fourier modes given by the FFT convention defined in equation 1.37 lead to the expression

$$\langle|\delta_{\ell,k}|^2\rangle = \frac{1}{N^2N_\ell^2}\left\langle\left(\sum_{i=0}^{N-1} n_i e^{ikx_i}\right)\left(\sum_{j=0}^{N-1} n_j e^{-ikx_j}\right)\right\rangle$$

$$= \frac{1}{N^2N_\ell^2}\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}\langle n_i n_j\rangle e^{ik(x_i-x_j)}, \tag{2.6}$$

where $N$ is the number of pixels. Now in the absence of clustering all $i \neq j$ terms

will vanish leaving just

$$\langle |\delta_{\ell,k}|^2 \rangle = \frac{1}{N^2 N_\ell^2} \sum_{i=0}^{N-1} \langle n_i^2 \rangle = \frac{1}{N^2 N_\ell}, \tag{2.7}$$

where the final term follows since the only occupancy numbers allowed are 0 or 1 meaning that $n_i^2 = n_i$. This shot noise is present in all $k$-modes and must be subtracted when calculating $P_{1D,\ell}(k)$.

### 2.1.4 Results and discussion

#### Flux and Gaussianized power spectra

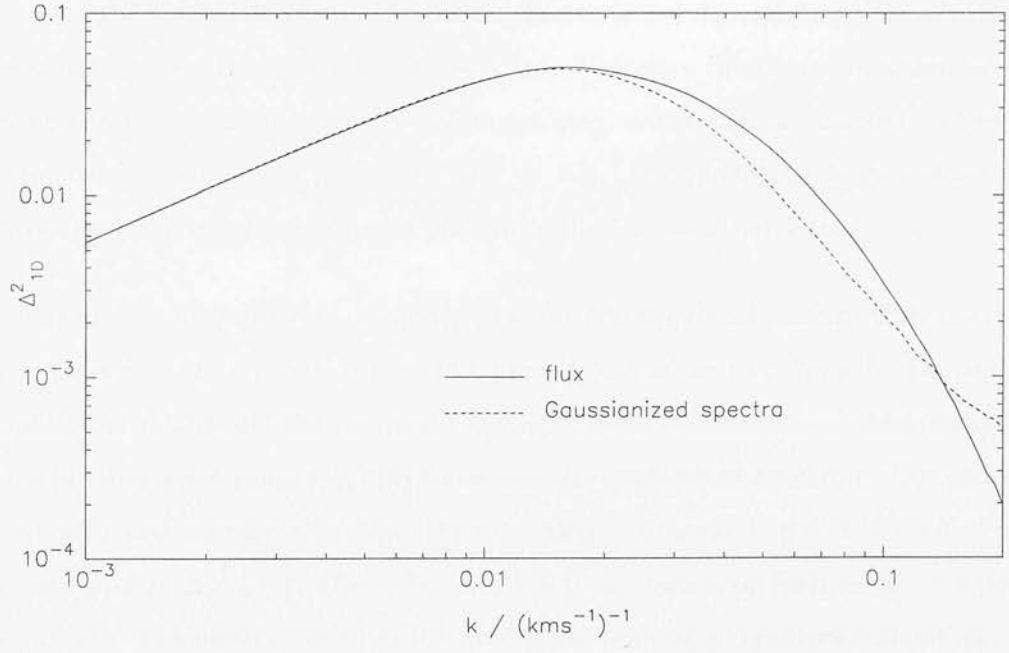


**Figure 2.2:** The dimensionless power spectra for the flux and Gaussianized values. The Gaussianized power spectra has been renormalised to match the flux at $k = 1.0 \times 10^{-3}$ $(\mathrm{km\,s^{-1}})^{-1}$. In the regime of $k < 0.01$ $(\mathrm{km\,s^{-1}})^{-1}$ the lines are almost identical. On smaller scales the lines diverge with the Gaussianized case showing a flatter drop off than the flux counterpart.

The process of Gaussianization only produces a power spectrum of relative am-

plitude. Since at this stage it is the shape, not the absolute magnitude, that is of interest the Gaussianized power can be renormalised to match the flux counterpart for easy comparison. Figure 2.2 shows the averaged power spectra of the flux and Gaussianization values where the latter has been renormalised to equal the former at $k = 1.0 \times 10^{-3}$ (km s$^{-1}$)$^{-1}$. In the region $k < 0.01$ (km s$^{-1}$)$^{-1}$ the power spectra are an excellent match, though at higher $k$ values the Gaussianized case drops from the maximum value with a flatter decline than the flux power spectrum. The important issue is whether, when analysing more advanced simulations or real data, one of these methods offers better recovery of $\Delta^2_{\rm 1D}(k)$ than the other. If the technique of Croft *et al.* (2000) is applied then the results obtained should not depend on whether the Gaussianization process is applied or not. The reason for this is that the change in the shape of the power spectrum from Gaussianization will be counteracted by a change in the shape of $b(k)$, the parameter used to link the flux and mass power spectra. Additionally since the amplitude of a Gaussianized spectra is not dependent on the actual flux levels, information is actually being discarded. Therefore the Gaussianization process can be considered an unnecessary additional step, which can be omitted. Therefore the following analysis was performed only on $\Delta^2_{\rm 1D,F}(k)$, though on large scales this is equivalent to considering the power spectra of the Gaussianized values.

Figure 2.3 shows the recovered $\Delta^2_{\rm 1D,F}(k)$ for 500 simulated spectra with the input parameters $\gamma = 2.5$, $\beta = 1.5$, $\bar{b} = 23.0$ km s$^{-1}$ and $\sigma_b = 14.0$ km s$^{-1}$. Despite the simplicity of this model the results do highlight some considerations which should be borne in mind when using $P_{\rm 1D,F}(k)$ to recover the mass power spectrum. The recovery of $P_{\rm 1D,F}(k)$ and subsequently $\Delta^2_{\rm 1D,F}(k)$ show that even a random distribution of lines gives a non-zero $\Delta^2_{\rm 1D,F}(k)$. This is because the Ly-$\alpha$ absorption features are not point-like objects. The natural width of the absorption features guarantees correlations will be seen on scales corresponding to a few tens of km s$^{-1}$. It does though highlight the fact that, in general and on some scales, the flux power spectrum will not necessarily be linearly related to the mass power spectrum (consider equation 1.69 for $P(k) = 0$).

Varying some of the input parameters affects the recovered shape of $\Delta^2_{\rm 1D,F}(k)$. Altering the redshift distribution parameter $\gamma$ does not have a significant affect on $\Delta^2_{\rm 1D,F}(k)$ since the constraint on mean optical depth ensures that changing $\gamma$ will alter

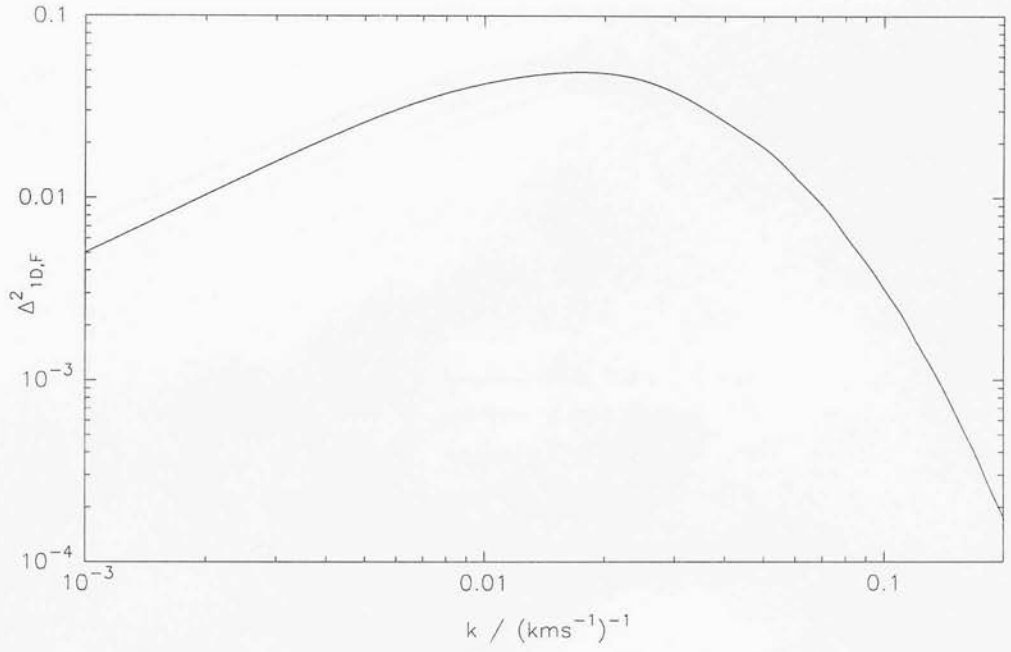


**Figure 2.3:** The dimensionless flux power spectrum for a Poisson line distribution. Although the placement of the features is unclustered the flux power spectrum has non negligible Fourier components. These correlations in the flux spectrum exist since Ly-$\alpha$ absorbers are not point like objects, but instead have a finite width.

$dn/dz|_0$ to ensure approximately the same number of lines still occur in each spectrum. Since the shape of $\Delta^2_{1D,F}(k)$ is dependent on the line profiles and all these features have the same intrinsic statistical properties the flux power spectrum remains largely unchanged.

Conversely varying the parameters which determine the absorption profiles can be seen to be significant. Figure 2.4 shows the affect of varying the column density distribution parameter $\beta$. It can clearly be seen that the magnitude of the peak of the recovered $\Delta^2_{1D,F}(k)$ is dependent on $\beta$. A high value of $\beta$ corresponds to systems of high $N_{HI}$ being rare. When $\beta$ is large the constraint on mean optical depth means that there will be more absorption features than for a low value of $\beta$, but these features will tend to have lower column densities. Figure 2.4 shows that fewer lines with a wider range of column densities have bigger Fourier components than more lines with a small range of $N_{HI}$. This can best be understood if an extreme case is considered

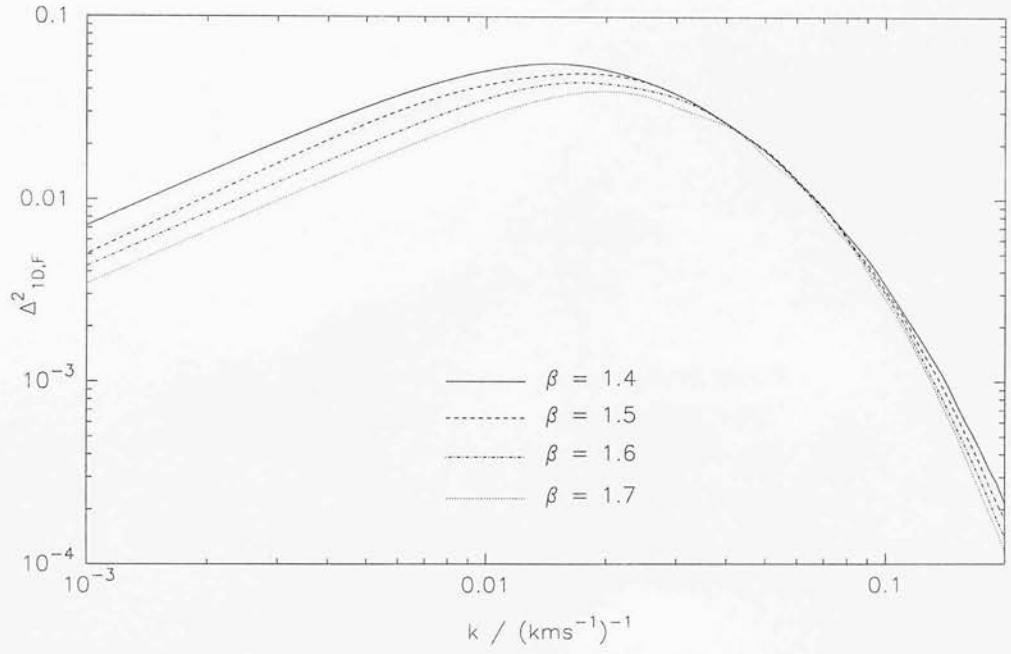


**Figure 2.4:** The dimensionless flux power spectrum for various values of $\beta$. This parameter determines the relative abundances of high and low column density features. As the statistical properties of the lines change so does the resulting flux power spectrum. The peak value decreases as high column density features become rarer.

where $\beta$ tends to infinity and therefore all the lines present tend to have the same, low, column density. In this case satisfying the $\bar{\tau}$ constraint requires that the spectrum has a small, almost constant, amount of absorption at each point across its entire length. Performing an FFT on this approximately constant flux would then result in negligible Fourier coefficients.

Figures 2.5 and 2.6 show the affect on $\Delta^2_{\mathrm{1D,F}}(k)$ of different values of $\bar{b}$ and $\sigma_b$. Changing the distribution of $b$ can be seen to have a small but noticeable effect on the recovered $\Delta^2_{\mathrm{1D,F}}(k)$. Looking at the region of $k > 0.1$ $(\mathrm{km\,s^{-1}})^{-1}$ these graphs show that increasing $\bar{b}$ or $\sigma_b$ has the effect of slightly narrowing the feature recovered in Fourier space. This is to be expected since high Doppler parameters correspond to broad absorption lines and so correlations can be expected on larger scales.

When more realistic models are considered the aim is to use $\Delta^2_{\mathrm{1D,F}}(k)$ to recover its

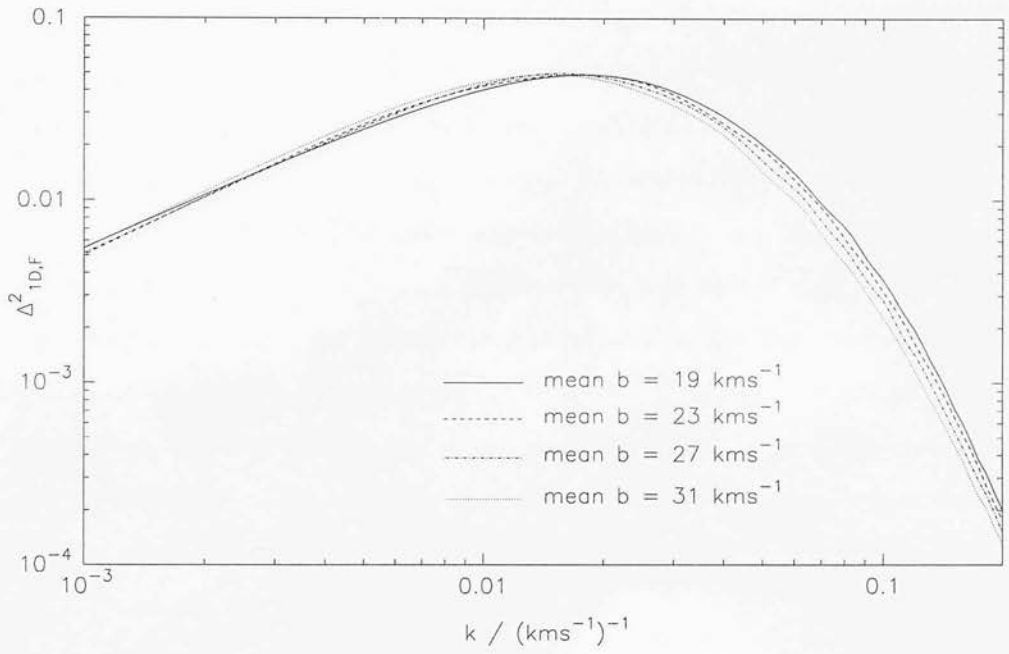


**Figure 2.5:** The dimensionless flux power spectrum for various values of $\bar{b}$, the mean Doppler parameter. For a given line, the width depends on the value of the Doppler parameter. As the mean Doppler parameter is increased a slight narrowing of the feature seen in Fourier space is apparent.

mass counterpart $\Delta^2_{1D}(k)$. The Croft *et al.* (2000) method suggests that this can be done with a bias parameter, $b(k)$. Therefore the change in the intrinsic shape of $\Delta^2_{1D,F}(k)$ given by a Poisson distribution of lines should be accounted for by this bias parameter. The important point is that $b(k)$ is calculated from the ratio of these power spectra in simulations. If the simulations correctly account for the factors that give rise to parameters governing the line statistics then the precise shape of a Poisson distribution of lines should be unimportant. Regardless of this the values of these parameters may give some indications of the suitability of this method, which involves the concept of inferring a continuous mass field from a quasar spectrum.

Certain values of $\beta$, which governs the column density distribution, for example may affect the fundamental assumption that the optical depth corresponding to a spatial point can be used to infer the density of dark matter at that point. A low value of $\beta$ would mean that absorption features are more likely to be saturated than a high value.

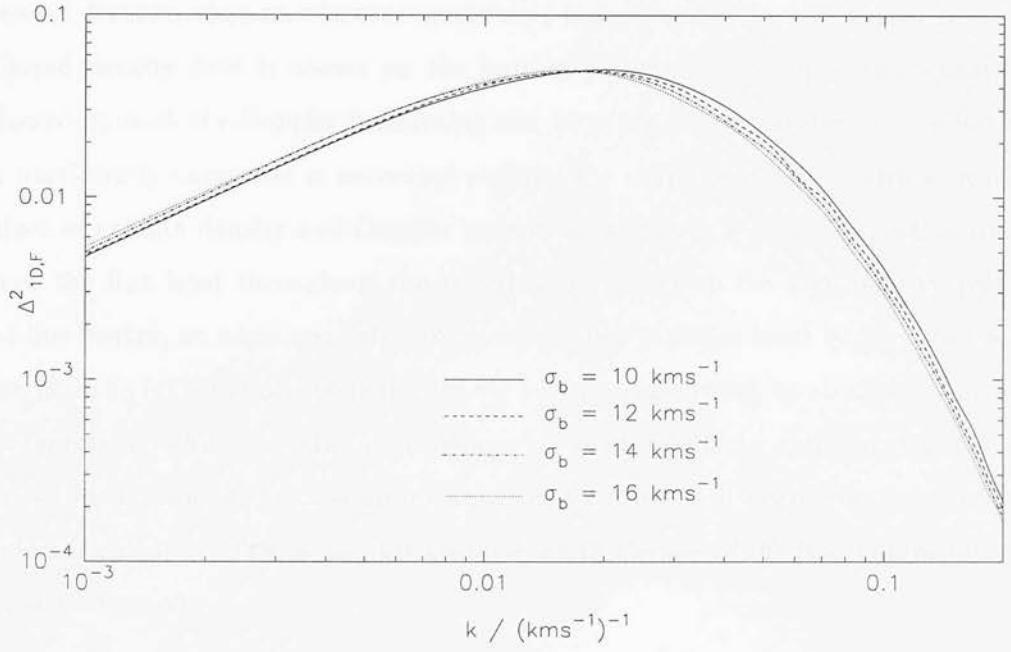


**Figure 2.6:** The dimensionless flux power spectrum for various values of $\sigma_b$, the variance in the Doppler parameter distribution function. As $\sigma_b$ increases, the likelihood of a feature having a high Doppler parameter is also increased. This is due to the truncation of the Gaussian distribution function, arising from the presence of a constraint on the minimum allowable value of $b$. Therefore the width of the flux power spectrum is weakly dependent on $\sigma_b$.

Strong absorption lines in spectra could be problematic when attempting to reconstruct the mass field. This follows since the inferred density at the centre of these saturated regions is only a lower limit. The reason for this is that the spectra cannot be used to accurately determine even the relative density in such regions since any additional obscuring material above some threshold limit will not cause any additional absorption.

The extent of the blending of lines, which depends on $\beta$, $\bar{b}$ and $\sigma_b$, can also be significant. When the lines are strongly blended the optical depth at any point may have contributions from many regions around that point. The Doppler broadening of lines confuses the link between the absorption seen and the spatial position it corresponds to. Figure 2.7 demonstrates this effect, in the underlying density field (shown in the top plot) there are two peaks close together. In the middle graph the profiles of the

absorption features corresponding to these peaks are shown as solid lines while the blended feature, that an observer would see, is represented by the dotted line. The inferred density field is shown on the bottom plot and shows how this quantity is affected by both the Doppler broadening and blending of the features. This effect may be particularly important in saturated regions. For example a feature with a very high values of column density and Doppler parameter will have a broad saturation trough. Since the flux level throughout the trough is zero, due to the high density point at the line centre, no additional absorption associated with the local baryons can occur. Therefore no information about the density in these regions can be obtained. Any areas of a spectrum which contain line blending, of which this is an extreme example, lead to the breakdown of the assumption that the local value of $\tau$ gives a measure of the underlying density. This is another obstacle which the use of the bias parameter, $b(k)$, hopes to overcome.

The distribution of Doppler parameters also raises another issue concerning the extent to which the baryons trace the dark matter. The Doppler parameter is intimately linked with the temperature of the gas. Since the underlying dark matter is thought to be collisionless, a key difference in the behaviour of the baryons is that at high densities it becomes pressure dominated and hot. The differing physical properties of these two types of matter will inevitably lead to differences in their distributions. The pressure forces arising in the baryonic distributions will, on certain scales, effectively oppose gravitational collapse. The critical length below which collapse is halted for a given system is known as the Jeans length. This scale arises since if the free fall collapse time of a structure is less than the time it takes for sound waves to travel across it then the pressure is unable to halt the collapse. The time taken for sound waves to cross a structure of size $x$ is just given by $t = x/c_s$. The free fall collapse time can be approximated by considering the simplified case where a particle undergoes constant acceleration starting from rest. This results in a collapse time given by $t_{\rm ff} = \sqrt{1/G\rho}$. Equating these two time scales then gives the critical length as $x = c_s\sqrt{1/G\rho}$, a result remarkably similar to the exact expression for the Jeans length which is

$$\lambda_{\rm J} = c_s\sqrt{\frac{\pi}{G\rho}}. \tag{2.8}$$

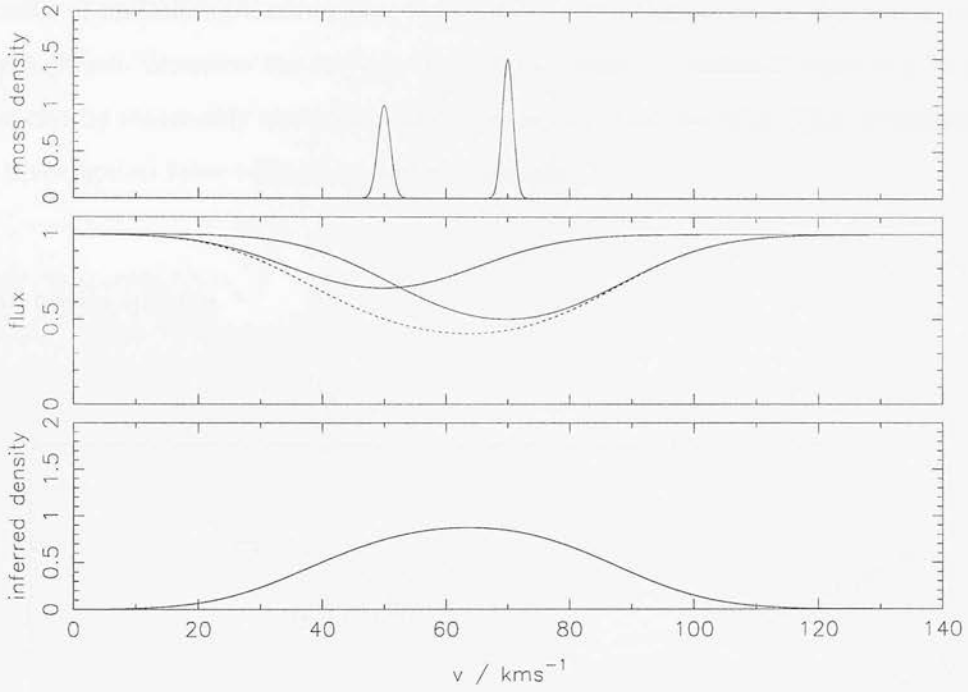


**Figure 2.7:** The effects of Doppler broadening and line blending on an inferred density field. The top and bottom plots show the actual and inferred density profiles respectively. The middle plot shows the absorption features associated with the density peaks and the blended profile. The inferred mass density shows that in addition to inferring a wide density profile, due to Doppler broadening, that one of the peaks is obscured by the line blending.

The effect on the power spectrum of the baryonic pressure forces can be modelled with the equation (e.g Fang *et al.* 1993)

$$P_b(k) = \frac{P_{\rm DM}(k)}{(1 + x_b^2 k^2)^2}, \tag{2.9}$$

where $x_b$ represents the Jeans length and is given by

$$x_b(z) = \frac{1}{H_0} \left[ \frac{2\gamma k_{\rm B} T_0(z)}{3\mu m_p \Omega_{m0}(1+z)} \right]^{1/2}, \tag{2.10}$$

where all the symbols have their usual meanings and $T_0$ is the temperature of the IGM at the mean baryon density and $\mu$ is the mean molecular weight of the IGM. Since $x_b$ is proportional to the square root of the temperature, $T_0$, increasing this quantity leads to a greater divergence between these distributions. The difference of the clustering is therefore likely to be greatest at the places where the gas is hottest.

In summary the assumption that absorption flux can give an indication of the density of underlying dark matter works best in the regime where any gas is cold and fairly diffuse. However the hope is that all the effects introduced when this is not the case can be reasonably accounted for by the bias parameter $b(k)$. This assumption will be investigated later with more realistic models.

### Line power spectra

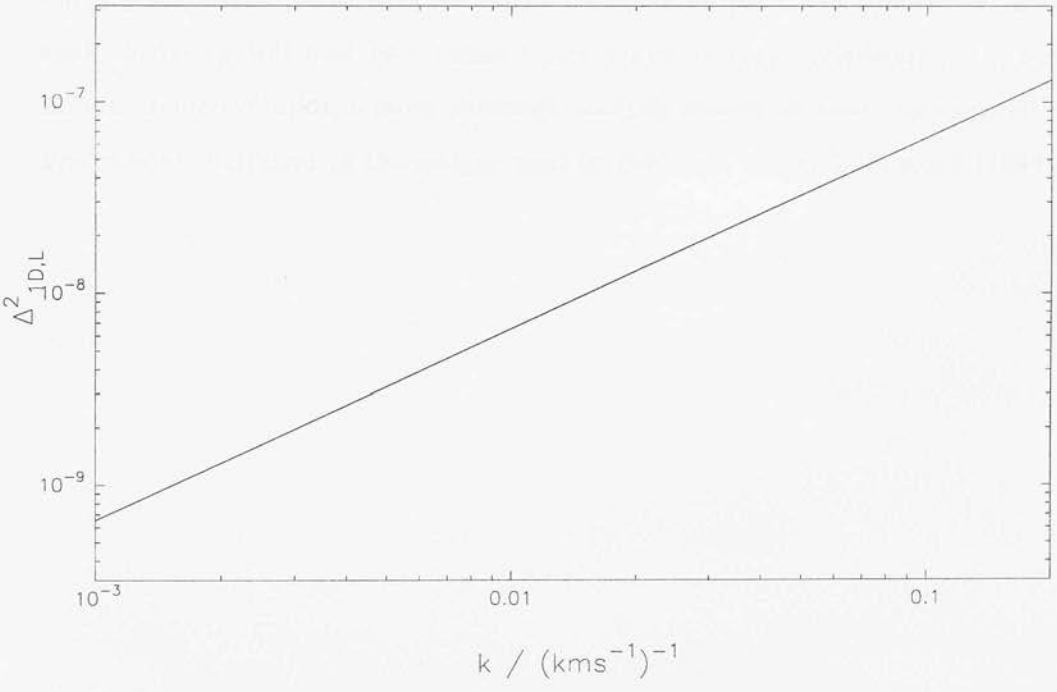


**Figure 2.8:** Dimensionless line power spectrum for a Poisson line distribution. For an infinite number of realizations the expected power spectrum would be zero on all scales. The large scatter on individual realizations leads to the non-zero result seen. Even when a signal is present it is, as expected, largely scale invariant.

The power spectrum of line positions does reasonably recover the Poisson result $\Delta^2_{1D,\ell}(k) = 0$ as shown in figure 2.8. The result does not vanish completely as $\Delta^2_{1D,\ell}(k)$ will only be zero for an infinite number of realizations. The scatter on the individual realizations is very large and the shot noise subtraction can lead to calculating a $P_{1D,\ell}(k)$

which is not non-negative definite. However averaging over the many realizations used here gives a best fit line where the recovered signal is weak and positive. Any signal that is recovered should not have a scale dependence, meaning $\Delta^2_{1D,\ell}(k)$ should be a straight line. The recovered form of $\Delta^2_{1D,\ell}(k)$ is an excellent fit to the straight line expected. The recovered signal is consistent with a Poisson distribution even despite the evolution, with redshift, of the line population. This effect must be considered since $P_{1D,\ell}(k)$ has been calculated from an expression which assumes that $\bar{n}$ is a constant. However the evolution across the spectra is not very dramatic and so the result still offers an excellent approximation. Comparison between the case where the line population evolves and where $\bar{n}$ is constant show that the errors are small on the scales considered, and even weak clustering will lead to a signal far in excess of these systematics. In cases of more extreme evolution a more thorough analysis should be used, for example the 1 dimensional equivalent of the scheme used by Feldman, Kaiser & Peacock (1994).

## 2.2 Adding an underlying density field

Once a density field has been generated (using the method described below) the selection of lines can be modified to position lines preferentially at places of high density. This will lead to a situation where the line centres will offer a Poisson sampling of the density field. To do this equation 2.1 is replaced with

$$dp = \frac{dN}{dz}[1 + \delta(x)]\Delta z. \tag{2.11}$$

The Monte Carlo simulations are then performed in exactly the same way as before except for this modification of line placement.

### 2.2.1 The form of the baryonic power spectrum

The input power spectrum used in this chapter was calculated as follows. The initial form of the dark matter perturbations was assumed to be a featureless power law such that $P_{\rm DM}(k) \propto k^n$, where $n$ is known as the spectral index and determines the relative power on large and small scales.

Various processes (gravitational interaction, dissipation etc) cause the initial perturbations to change, altering the functional form of the power spectrum. This alteration is quantified by the transfer function, $T(k)$, which is the ratio of the modern day amplitude of each mode and its initial value. Bardeen *et al.* (1996) found that the transfer function for CDM cosmologies was well fitted by the formula

$$T(q) = \frac{[\ln(1 + \alpha_1 q)/\alpha_1 q]}{[1 + \alpha_2 q + (\alpha_3 q)^2 + (\alpha_4 q)^3 + (\alpha_5 q)^4]^{1/4}}, \tag{2.12}$$

where $q \equiv k/\Gamma$, and $\Gamma$, known as the shape parameter, is dependent on $\Omega_m$, $\Omega_b$ and $h$. The $\alpha$ coefficients also depend on cosmological parameters and the values used here are the results of Ma (1996) corresponding to $\Omega_b = 0.05$ ($\alpha_1 = 2.205$, $\alpha_2 = 4.05$, $\alpha_3 = 18.3$, $\alpha_4 = 8.725$ and $\alpha_5 = 8.0$). The modern day linear power spectrum is then given by $P_{\rm DM}(k) \propto k^n T^2(k)$. The constant of proportionality is constrained using the modern day value of $\sigma_8$, which refers to the rms density fluctuations averaged over spheres of

radius 8 $h^{-1}$ Mpc. Observationally $\sigma_8$ is found to have a value close to unity. So with this prescription a modern day linear power spectrum can be generated for given values of just 3 parameters ($n, \Gamma$ and $\sigma_8$).

This power spectrum must now be renormalized to correspond to the redshift appropriate to the simulations. This renormalization corresponds only to a change in amplitude due to the evolution of linear perturbations, not a change in functional form. Since the redshift range of these simulations is small, it is a reasonable approximation to generate the power spectra at one redshift ($z = 3.0$ in this case) and use this across the entire length of the spectra. In linear theory, the power spectrum evolves according to

$$P_{\rm DM}(k, z) = D^2(z) P_{\rm DM}(k, 0), \tag{2.13}$$

where $D(z)$ is the linear growth factor for the cosmological parameters being used and is given in equation 1.22.

This linear form of $P_{\rm DM}(k, z)$ can then be used to calculate the more general $P_{\rm DM,NL}(k, z)$ in which the effects of the non-linear regime on the power spectrum are also included. This is done using the Peacock & Dodds (1996) method, which is an extension of work carried out by Hamilton *et al.* (1991). First the scales of the non-linear and linear regimes can be linked since non-linear density contrasts arise from the collapse of linear perturbations on linear scales. Once the scales corresponding to the two regimes have been established it has been shown that it is a good approximation to assert that the non-linear perturbations are a universal function of linear ones. These relationships are given by the expressions

$$k_{\rm L} = \left[1 + \Delta^2_{\rm NL}(k_{\rm NL})\right]^{-1/3} k_{\rm NL}, \tag{2.14}$$

$$\Delta^2_{\rm NL}(k_{\rm NL}) = f_{\rm NL}[\Delta^2_{\rm L}(k_{\rm L})]. \tag{2.15}$$

The function $f_{\rm NL}(x)$ is found from $N$-body simulations and is given by

$$f_{\rm NL}(x) = x \left\{ \frac{1 + B\beta x + (Ax)^{\alpha\beta}}{1 + [(Ax)^{\alpha} g^3(\Omega)/(V x^{1/2})]^{\beta}} \right\}^{1/\beta}, \tag{2.16}$$

where the coefficients are

$$A = 0.482(1 + n/3)^{-0.947},$$

$$B = 0.226(1 + n/3)^{-1.778},$$

$$\alpha = 3.310(1 + n/3)^{-0.244},$$

$$\beta = 0.862(1 + n/3)^{-0.287},$$

$$V = 11.55(1 + n/3)^{-0.423},$$

and the function $g(\Omega)$ accounts for the difference of the growth rate of linear perturbations in different cosmologies and is given in equation 1.24.

Once the non-linear dark matter power spectrum at the appropriate redshift has been calculated, it can be smoothed to find the corresponding baryonic $P_b(k)$. This accounts for the difference of the physical properties of baryons and dark matter which will inevitably affect the clustering. As previously discussed the appropriate scale of the baryonic smoothing corresponds to the Jeans length and can be expressed as

$$P_b(k, z) = \frac{P_{\mathrm{DM,NL}}(k, z)}{(1 + x_b^2 k^2)^2}, \tag{2.17}$$

with $x_b$ given by

$$x_b(z) = \frac{1}{H_0} \left[ \frac{2\gamma k_{\mathrm{B}} T_0(z)}{3\mu m_p \Omega_{m0}(1 + z)} \right]^{1/2}, \tag{2.18}$$

where the parameters $T_0$ and $\gamma$ were predicted for a given redshift by Hui & Gnedin (1997).

Finally, since these simulations are one dimensional, a transformation from the 3D power spectrum must be made using the relation given in 1.33. Figure 2.9 shows the 1D dark matter and baryonic power spectra obtained at $z = 3.0$ for the parameters $n = 1.0$, $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$, $\Gamma = 0.2$ and $\sigma_8 = 1.0$.

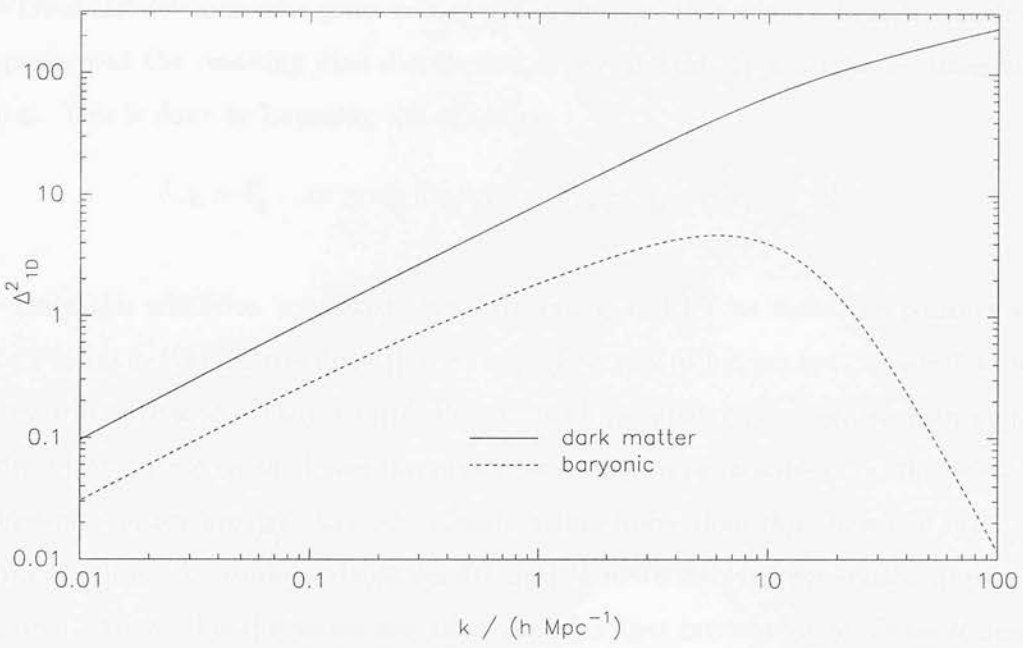


**Figure 2.9:** The baryonic and dark matter 1D power spectra with parameters $n = 1.0$, $\Gamma = 0.2$ and $\sigma_8 = 1.0$ for the cosmology $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$. The effect of the baryon smoothing is to reduce the amplitude of the 3D power spectrum on small scales. This is due to baryon distributions being dominated on small scales by pressure processes. Since the 1D power spectrum is an integral to infinity of the 3D form the 1D baryonic power spectrum has a lower amplitude on all scales than the dark matter one.

### 2.2.2 The lognormal approximation

A density field for each realization can be generated by taking Gaussian fluctuations of the known input power spectrum $P(k)$. To do this it is noted that the corresponding $\delta_k$'s can be split into real and imaginary parts such that

$$\langle|\delta_k|^2\rangle = \langle|\delta_k^{(r)}|^2\rangle + \langle|\delta_k^{(i)}|^2\rangle = P(k). \tag{2.19}$$

Now assuming the phases of different modes are random and uncorrelated implies

$$\langle|\delta_k^{(r)}|^2\rangle = \langle|\delta_k^{(i)}|^2\rangle = P(k)/2, \tag{2.20}$$

meaning a density field with a Gaussian PDF around a known power spectrum can

be generated by taking $\delta_k^{(r)} = u \times \sqrt{P(k)/2}$ and $\delta_k^{(i)} = u' \times \sqrt{P(k)/2}$, where $u$ and $u'$ are Gaussian deviates with mean = 0, $\sigma = 1$. To ensure that when a Fourier transform is performed the resulting $\delta(x)$ distribution is a real field, $\delta_k$ must be an Hermitian series. This is done by imposing the condition

$$\delta_{-\mathbf{k}} = \delta_{\mathbf{k}}^* \quad \text{or more explicitly} \quad \delta_{-\mathbf{k}}^{(r)} = \delta_{\mathbf{k}}^{(r)} \quad \delta_{-\mathbf{k}}^{(i)} = -\delta_{\mathbf{k}}^{(i)}. \tag{2.21}$$

Once this condition is applied then performing an FFT on these components will give a suitable $\delta(x)$ distribution. However applying this technique as it stands does not necessarily give a physically sensible density field for input power spectra with amplitudes which are not small. Using this prescription offers no guarantees that the resulting values of $\delta(x)$ are greater than $-1$. Clearly values lower than this threshold make no physical sense. To ensure a viable density field, a more involved generation process is required. To do this the lognormal approximation first introduced by Coles & Jones (1991) is used. This model is a good approximation to non-linear matter distributions, such as those found in the IGM. This distribution has been tested against hydrodynamical simulations and been found to be in reasonable agreement (Bi & Davidsen 1997). It has also been used to model the IGM and found to have reproduced the observations well (Bi 1993; Bi & Davidsen 1997).

In order to use the lognormal approximation it is asserted that given a set of density perturbations $\delta_{\mathrm{G}}(x)$ (which have been marked with a subscript G because they are Gaussian fluctuations of the generating power spectrum $P_{\mathrm{G}}(k)$), the density of baryons is given by

$$\rho_b(x) = Ae^{\delta_{\mathrm{G}}(x)}, \tag{2.22}$$

where $A$ is a normalization constant. It should be noted that valid values of $\rho_b(x)$ are generated for any real value of $\delta_{\mathrm{G}}(x)$, thus the distribution of $\delta_{\mathrm{G}}(x)$ is not subject to the same constraint as the usual fractional density parameters that $\delta \geq -1$. The value of the constant $A$ can be found by invoking the properties of Gaussian fields. Taking the mean value of both sides of equation 2.22 gives

$$\langle \rho_b(x) \rangle \equiv \bar{\rho}_b = A\langle e^{\delta_{\mathrm{G}}(x)} \rangle. \tag{2.23}$$

Now since a Gaussian field is being used it is known (see Appendix B) that

$$\langle e^{\delta_{\rm G}(x)} \rangle = e^{\frac{\Delta^2}{2}}, \tag{2.24}$$

where

$$\Delta^2 = \langle \delta_{\rm G}^2(x) \rangle . \tag{2.25}$$

Hence

$$A = \bar{\rho}_b e^{-\frac{\Delta^2}{2}}. \tag{2.26}$$

Substituting equation 2.26 in to equation 2.22 gives

$$\rho_b = \bar{\rho}_b \exp\left(\delta_{\rm G}(x) - \frac{\Delta^2}{2}\right). \tag{2.27}$$

Figure 2.10 shows the distributions of a set of Gaussian fluctuations and the baryon densities generated from these fluctuations using the lognormal approximation. It can be seen that the distribution of the resulting baryon field peaks at low values, but has a long tail at the high end. This is to be expected since individual values have the constraint that $\delta_b \geq -1$, while the sum all values must remain equal to zero. It is therefore clear that applying the lognormal approximation to calculate baryon densities will lead to a distribution of $\delta_b$'s which has a different power spectrum to the corresponding input Gaussian fluctuations. Therefore a formalism is needed to calculate the power spectrum $P_G(k)$ which will provide suitable Gaussian fluctuations to generate a baryonic density field with the required distribution for a known power spectrum $P_b(k)$. The first step is to recall that the correlation function is given by

$$\xi_b(x) = \langle \delta_b(x') \delta_b(x' + x) \rangle. \tag{2.28}$$

Now from equations 1.16 and 2.27 it can be seen

$$\delta_b(x) = \exp\left(\delta_{\rm G}(x) - \frac{\Delta^2}{2}\right) - 1, \tag{2.29}$$

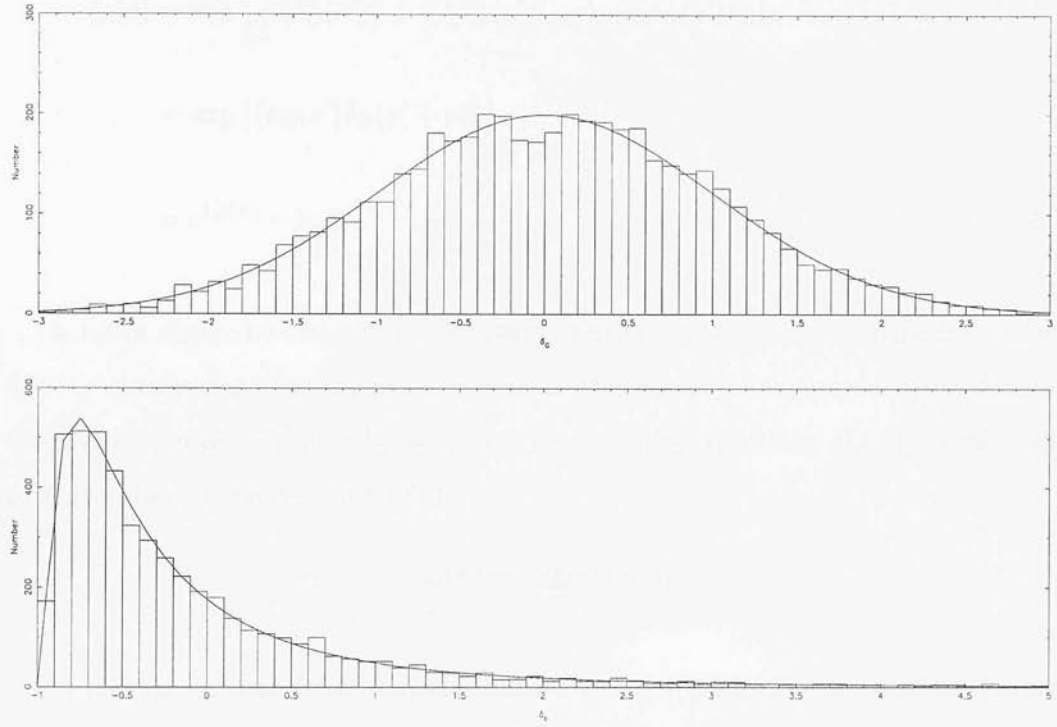


**Figure 2.10:** Histograms showing the distribution of a set of Gaussian fluctuations, $\delta_G$, and the baryon field resulting from applying the lognormal approximation to these fluctuations, $\delta_b$. The numbers expected in each bin are shown by the overlying lines. It can be seen that the resulting baryonic density field is non-Gaussian, being skewed towards low density values whilst having a long tail for high values.

which upon substitution to equation 2.28 gives

$$\xi_b(x) = \left\langle \left\{ \exp\left[\delta_{\rm G}(x') - \frac{\Delta^2}{2}\right] - 1 \right\} \left\{ \exp\left[\delta_{\rm G}(x'+x) - \frac{\Delta^2}{2}\right] - 1 \right\} \right\rangle . \tag{2.30}$$

Multiplying these terms out gives

$$\xi_b(x) = \left\langle \exp\left[\delta_{\rm G}(x') + \delta_{\rm G}(x'+x) - \frac{2\Delta^2}{2}\right] \right\rangle - \left\langle \exp\left[\delta_{\rm G}(x') - \frac{\Delta^2}{2}\right] \right\rangle$$

$$- \left\langle \exp\left[\delta_{\rm G}(x'+x) - \frac{\Delta^2}{2}\right] \right\rangle + 1. \tag{2.31}$$

Now using the properties of Gaussian statistics as shown in equation 2.24 the result simplifies to

$$\xi_b(x) = \exp\left\{ \frac{1}{2} \left\langle \left[\delta_{\rm G}(x') + \delta_{\rm G}(x'+x)\right]^2 \right\rangle - \Delta^2 \right\} - 1$$

$$\xi_b(x) = \exp\left[\frac{1}{2}\left\langle\delta_{\rm G}^2(x')\right\rangle + \frac{1}{2}\left\langle\delta_{\rm G}^2(x'+x)\right\rangle + \left\langle\delta_{\rm G}(x')\delta_{\rm G}(x'+x)\right\rangle - \Delta^2\right] - 1$$

$$= \exp\left[\left\langle\delta_{\rm G}(x')\delta_{\rm G}(x'+x)\right\rangle\right] - 1$$

$$= e^{\xi_{\rm G}(x)} - 1. \tag{2.32}$$

A result shown by Coles & Jones (1991). Thus it appears that to generate a baryon density distribution with the power spectrum $P_b(k)$ from the lognormal approximation, the original set of $\delta_{\rm G}$'s must be generated from a power spectrum $P_{\rm G}(k)$ corresponding to $\xi_{\rm G}(x)$ which is related to $\xi_b(x)$ by

$$\xi_{\rm G}(x) = \ln[\xi_b(x) + 1]. \tag{2.33}$$

From this equation it can be seen that for low amplitudes the forms of the baryonic and generating power spectra will be coincident (since for small $x$, $\ln(1+x) \simeq x$). It should also be noted that this prescription will break down if $\xi_b(x)$ ever has a value of, or lower than, $-1$. Another problem which can be encountered is that even for a given correlation function $\xi_b(x)$ which does correspond to a well defined $\xi_G(x)$, there is no guarantee that the resulting $P_{\rm G}(k)$ will be non-negative definite.

Even when $P_{\rm G}(k)$ is well defined and positive this mechanism only works within certain limits if $P_{\rm G}(k)$ is large. For example figure 2.11 shows the average recovered $\Delta^2_{{\rm 1D},b}(k)$ for 500 realizations using this technique, with the length and resolution parameters previously discussed, for the baryonic power spectrum shown in figure 2.9. The discrepancy seen between the input and recovered power spectra can be eliminated by using a longer baseline to calculate the density field. Even in this case care must taken, although the mean recovered $\Delta^2_{{\rm 1D},b}(k)$ may be in good agreement with the desired value, the variances of the densities for each realization may still be very erratic. This is demonstrated in figure 2.12. The top histogram shows the variances calculated from the Fourier components when Gaussian fluctuations are applied to the baryonic power spectrum. The second histogram shows the variances recovered from the lognormal approximation. In both cases the solid line shows the value of the mean from the Gaussian fluctuations of the baryonic power spectrum, while in the bottom

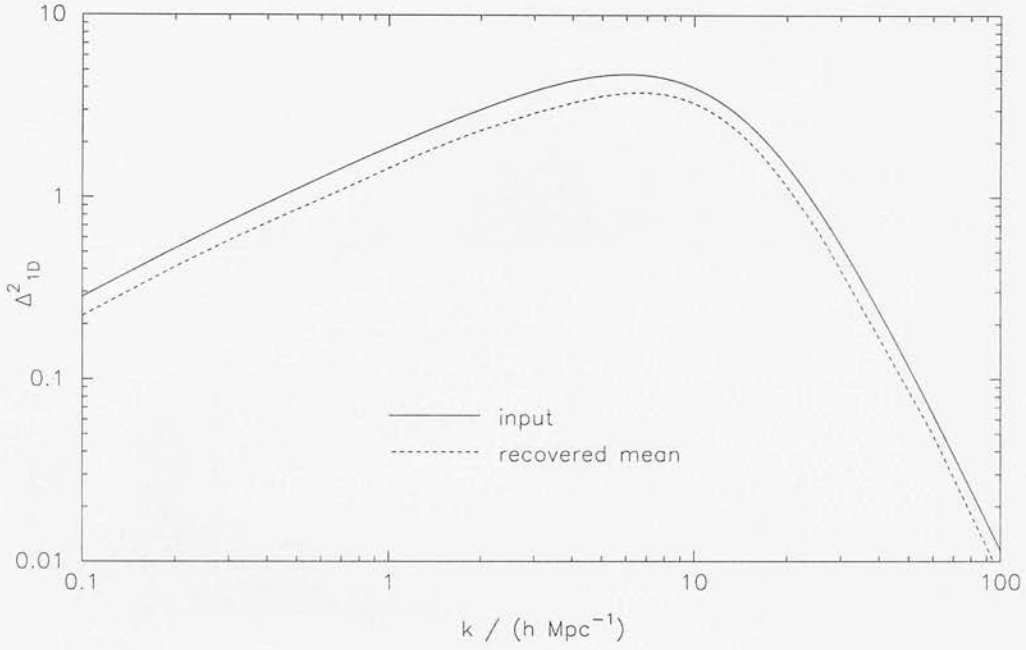


**Figure 2.11:** Input and recovered dimensionless baryonic power spectra when using the lognormal approximation. The method breaks down if the base line is not of sufficient length, as shown by the discrepancy between the two lines. Although the shape of the input power spectrum is well recovered the amplitude is too low.

plot the dashed line shows the mean of the variance when using the lognormal approximation. The simulations used to generate these density fields have a baseline two orders of magnitude greater than those of the simulations previously discussed. It can be seen that although the average variance of each distribution is very similar, the distributions themselves are radically different. The mean variance of the density fields from the lognormal approximation is strongly affected by a few realizations with very large variances. This effect can be reduced by increasing the length of the baseline still further, however since only a subsection of any density field generated on a large baseline is used there still remains the possibility that this subsection will contain some points with very high densities, and thus a large variance. This should not present a problem however since these density fields are used either, in this chapter, to allocate a line position or, in chapter 3, to calculate the optical depth at some point. Therefore the flux variance of any spectra will not be unduly affected by these rare points of very

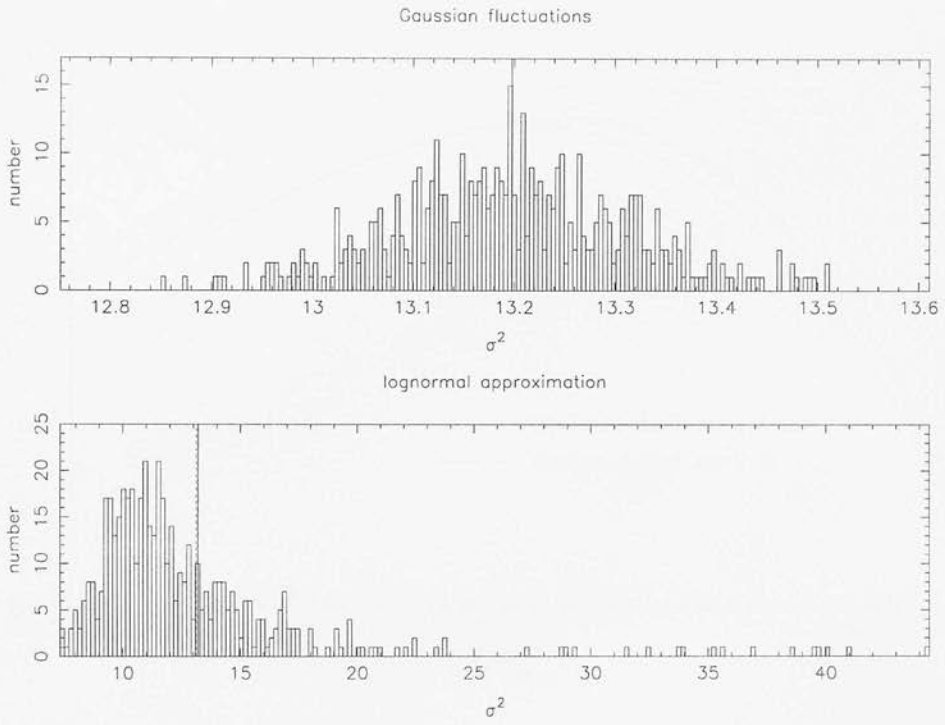


**Figure 2.12:** Distribution of variances when Gaussian fluctuations are applied to the baryonic power spectrum (top plot), and from realizations of the lognormal approximation (bottom plot). Care must be taken since although the mean values of the distributions are very similar the distributions themselves are radically different. The solid and dashed vertical lines show the mean value of the variances for the Gaussian fluctuated and generated baryonic density fields respectively.

high overdensity.

### 2.2.3 Results and discussion

#### Flux power spectra

Figure 2.13 shows the recovered $\Delta^2_{\mathrm{1D,F}}(k)$ for the two cases where the line distribution is Poisson and clustered according to an underlying density field. The value of $\Delta^2_{\mathrm{1D,F}}(k)$ is affected by the preferential placing of lines in regions of high density. The intrinsic properties of the absorption features are identical in both cases so the difference can be attributed solely to the positioning of the lines. On large scales the flux power

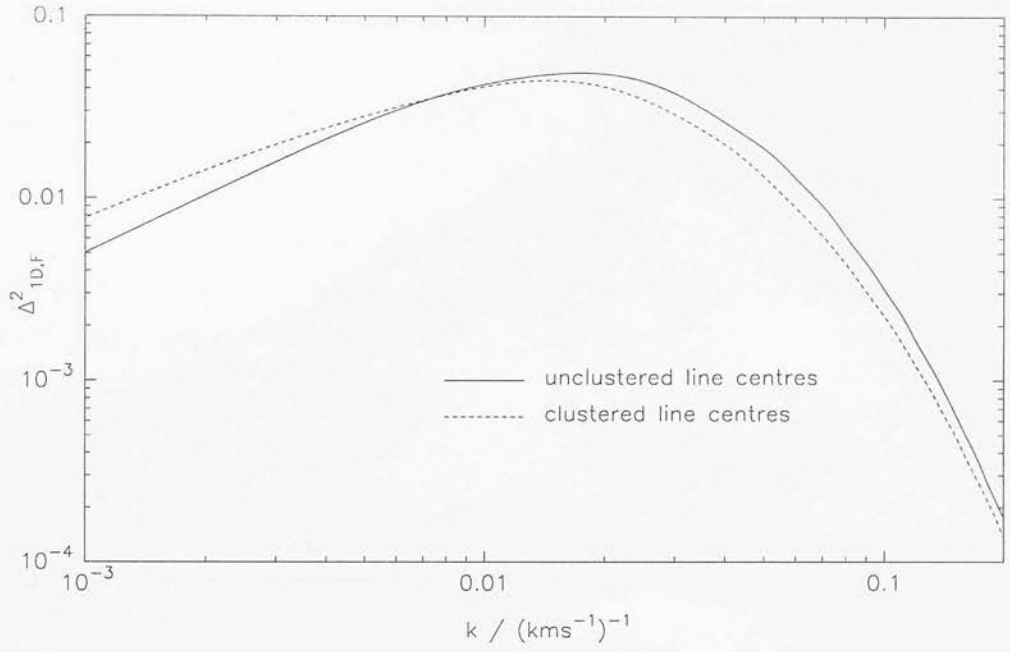


**Figure 2.13:** The dimensionless flux power spectra for both clustered and unclustered lines. The presence of clustering has an effect on the flux power spectrum even though the statistical properties of the absorbers are the same. On large scales the power is boosted by the presence of clustering and conversely on small scales it is repressed.

spectrum of the clustered lines is higher than the Poisson distribution. On these scales the absorption features can be thought of as isolated markers of regions of high density leading to an increase in power. More interestingly there is also a suppression of power on small scales where the Voigt profiles themselves should dominate. The amount of line blending must account for the difference between the two power spectra. This result coupled with those in §2.1.4 demonstrates that the flux power spectrum is influenced both by the intrinsic statistical properties of any absorption features and their positions. In observed data these two factors may be degenerate and are not necessarily easily disentangled since the underlying mass density will have an effect on both properties.

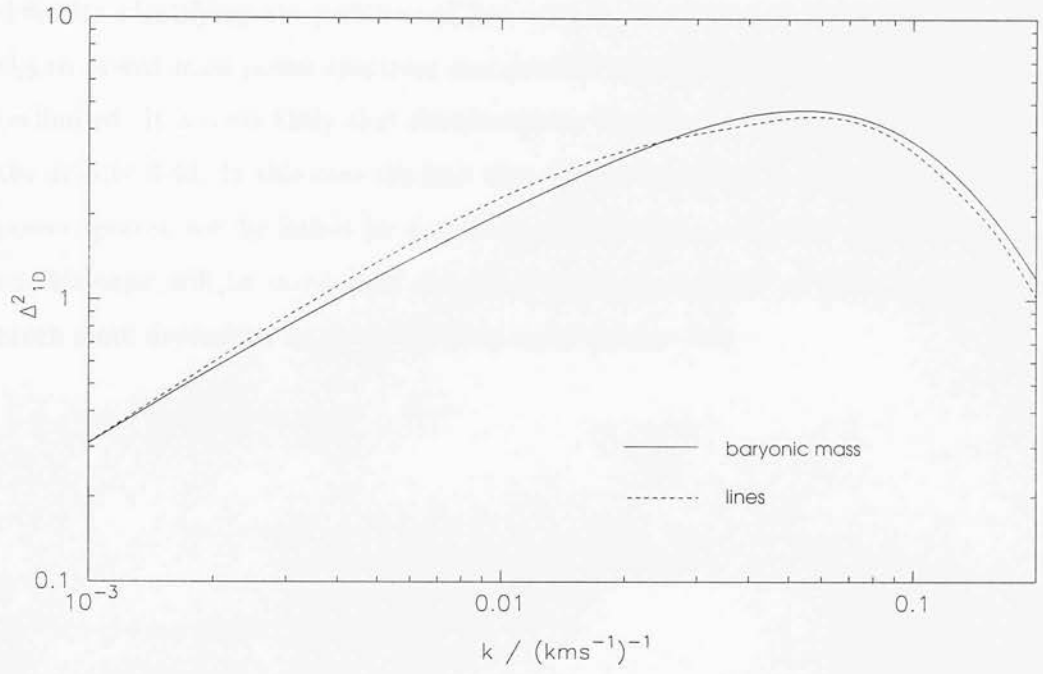


**Figure 2.14:** The dimensionless line power spectrum compared with its mass counterpart. As would be expected if the lines offer a Poisson sampling of the mass field the shape of recovered line power spectrum closely matches the mass equivalent. The amplitudes will in general be different and here the line power spectrum has been renormalised to match the mass power spectrum at $k = 1.0 \times 10^{-3}$ $(\mathrm{km\,s^{-1}})^{-1}$.

### Line power spectra

Figure 2.14 shows the line and mass power spectra where the line power spectra has been renormalised to match the mass one at $k = 1.0 \times 10^{-3}$ $(\mathrm{km\,s^{-1}})^{-1}$. The results show that in the unlikely situation that the line centres offer a Poisson sampling of the density field then the shapes of $\Delta^2_{1\mathrm{D},\ell}(k)$ and $\Delta^2_{1\mathrm{D}}(k)$ should be very similar. Therefore the power spectrum of line positions could be used to recover the shape, but not the magnitude, of the mass density power spectrum. Not much can be read in to this result since the method is so basic. It does however suggest that if line centres tend to be in high density regions then this may provide a useful aid in the recovery of the mass field. Even if this is the case then using this method does reintroduce the sorts of problems

which the Croft *et al.* method avoids. Shot noise for example appears and there may be difficulty identifying the positions of line centres. Furthermore the scales upon which the recovered mass power spectrum matches the underlying linear power spectrum will be limited. It is more likely that the absorption lines do not offer a Poisson sampling of the density field. In this case the best that can be hoped for is that the line and mass power spectra will be linked by a scale-dependent bias parameter. Further discussion on this topic will be saved until chapter 3 where the method of generating spectra is much more dependent on the underlying mass density field.

# Chapter 3

# Realistic spectra

In this chapter power spectra analysis is performed on spectra simulated by more sophisticated methods than in the previous chapter. Despite these simulations not being as rigorous as $N$-body simulations, they do have the advantage of requiring far less computing power and time thus allowing a wide range of parameter space to be studied. As before the flux and line power spectra will be obtained from these spectra. Additionally the relationship between the flux and mass power spectra will be analysed for various cosmologies. Finally a technique will be applied to highlight the importance of the gas properties on the flux power spectra. This technique involves scrambling the spatial information of a line list obtained for a given spectrum and then constructing a new spectrum from this altered absorber configuration. The flux power spectra of both the original and scrambled arrangement are then compared.

## 3.1 Simulated spectra

### 3.1.1 The model

The first step in these simulations is to generate a baryonic density field using the prescription outlined in §2.2.1. However, unlike previously where absorption lines were just laid down on this density field, the spectral features are obtained by summing the

contributions to the optical depth from each pixel. The line of sight spectrum can then be calculated using the prescription that the optical depth at any velocity along the line of sight is given by (e.g Meiksin & White 2000)

$$\tau(v) = \int \tau(x)\, dx, \tag{3.1}$$

where

$$\tau(x) \propto \rho_b(x)^2 T(x)^{-0.7} b^{-1} \exp\left[-\frac{(v(x) - v_0)^2}{b^2}\right]. \tag{3.2}$$

The temperature at any point can be computed from the equation of state given by

$$T(x) = T_0 \left(\frac{\rho_b}{\bar{\rho}_b}\right)^{\gamma - 1}, \tag{3.3}$$

where $\gamma$ and $T_0$ can be obtained from calculations performed by Hui & Gnedin (1997). This power law relation strictly only holds true in the regime where $\rho_b/\bar{\rho}_b \lesssim 10.0$. However, since points of density higher than this are rare, extending this relation to all points should not introduce a large error. Now substituting expressions for $b$ (equation 1.50) and $T$ (equation 3.3) into equation 3.2 gives

$$\tau(x) \propto \rho_b(x)^{3.2 - 1.2\gamma} \exp\left[-\frac{m_{\rm H}(v(x) - v_0)^2}{2k_{\rm B}T_0(\rho_b/\bar{\rho}_b)^{\gamma - 1}}\right]. \tag{3.4}$$

The constant of proportionality can be obtained by imposing the condition that the mean flux decrement of the spectra obtained have a value which matches the observations. As with the previous simulations, this is chosen to be $\bar{\tau}(z = 3.0) \simeq 0.35$. Once the optical depth has been calculated at each pixel the flux level is simply given by $e^{-\tau}$.

### 3.1.2 Range and resolution

These simulations must have a small redshift range in order to justify using a power spectrum which does not evolve across the length of the spectrum. As with the previous chapter the redshift range from $z = 2.9 - 3.0$ (velocity scale of $0 - 7600$ km s$^{-1}$) is studied, meaning once again the scales of $k$ which can be probed run from $8.0 \times 10^{-4} - 1.57$ (km s$^{-1}$)$^{-1}$.

## 3.2 The ratio of flux to baryonic power spectra

An important part of an accurate recovery of the mass density power spectrum using the most recent Croft *et al.* method is the role of the bias parameter, $b(k)$, as defined in equation 1.70. Rather than studying $b(k)$ directly, the work in this section will examine a related quantity, the ratio of the 1 dimensional flux and linear baryonic power spectra. The quantity that is calculated directly from the observations is $\Delta^2_{\rm 1D,F}(k)$ not the 3 dimensional equivalent. Croft *et al.* define an isotropic $\Delta^2_{\rm F}(k)$ from this quantity and compare this with its mass counterpart $\Delta^2(k)$. The true flux power spectrum is not expected to be isotropic, and so this constructed quantity should not be confused with the actual form of $\Delta^2_{\rm F}(\mathbf{k})$, which in general cannot be inferred from $\Delta^2_{\rm 1D,F}(k)$. Therefore, keeping in line with the observations, I prefer to work towards determining $\Delta^2_{\rm 1D}(k)$. If the 3 dimensional form is to be recovered then clearly a transform form via differentiation (see equation 1.34) must still be applied. However, since this transformation is not necessarily straightforward, it seems prudent to work with the data in its original form rather than inferring the 3 dimensional equivalent at the outset. Therefore studying the ratio of power spectra, as in the analysis below, can be considered as the equivalent of investigating the bias parameter, $b(k)$, in a 3 dimensional analysis.

The effect of cosmological parameters on the ratio between flux and linear mass power spectra must be considered, as accurate recovery of the latter is reliant on this ratio being not only a universal function, but also independent of cosmology. In this section a number of simulations are studied with various cosmological parameters and the aforementioned ratio is calculated.

Table 3.1 gives a list of the parameters used for each input baryonic power spectrum (calculated using the prescription described in §2.2.1). The first power spectrum is a fiducial ΛCDM model. Small alterations in the parameters (assuming a flat universe, so $\Omega_m$ and $\Omega_\Lambda$ are not independent variables) are then applied to obtain the following 4 power spectra. Power spectrum 6 has the same parameters as those used for the normalizing simulations of Croft *et al.* (2000). Finally an Einstein-de Sitter model is also considered. The flux power spectra for 1000 realizations are averaged for each cosmology, and the ratio calculated.

| Number | $\Omega_m$ | $\Omega_\Lambda$ | $\Gamma$ | Spectral Index $n$ | $\sigma_8$ |
|---|---|---|---|---|---|
| 1 | 0.30 | 0.70 | 0.20 | 1.00 | 1.00 |
| 2 | 0.35 | 0.65 | 0.20 | 1.00 | 1.00 |
| 3 | 0.30 | 0.70 | 0.30 | 1.00 | 1.00 |
| 4 | 0.30 | 0.70 | 0.20 | 0.90 | 1.00 |
| 5 | 0.30 | 0.70 | 0.20 | 1.00 | 0.80 |
| 6 | 0.40 | 0.60 | 0.26 | 0.95 | 0.74 |
| 7 | 1.00 | 0.00 | 0.20 | 1.00 | 1.00 |

**Table 3.1:** Parameters of power spectra to be considered

Clearly the amplitude of the flux power spectra, and therefore the ratio calculated, is dependent on the value of $\bar{\tau}$. In some limits this parameter will also affect the shape of the flux power spectra. This effect must be studied since if the shape of $\Delta^2_{\mathrm{1D,F}}(k)$ varies then not only the amplitude, but also the form of the recovered linear mass density power spectrum will be affected. Therefore $\Delta^2_{\mathrm{1D,F}}(k)$ is also obtained for the first power spectrum using the higher mean optical depth constraint of $\bar{\tau}(z = 3.0) \simeq 0.45$ (Press, Rybicki & Schneider 1993), as used by Croft *et al.*, and compared with the results for the lower mean optical constraint.

### 3.2.1 Results and discussion

Figure 3.1 shows the effect of altering the mean optical depth constraint on the flux power spectra. Although, as expected, the amplitudes of the resulting $\Delta^2_{\mathrm{1D,F}}(k)$ are different, their shapes match very closely. In this case, as $\bar{\tau}$ is increased so is the amplitude of the flux power spectrum. The reason for this is because any correlations present are more pronounced due to the higher levels of absorption. This trend is not general and some value of $\bar{\tau}$ will correspond to a maximum amplitude of $\Delta^2_{\mathrm{1D,F}}(k)$. This is the case since the two extreme cases ($\bar{\tau} = 0$ or $\infty$) both correspond to $\Delta^2_{\mathrm{1D,F}}(k) = 0$. So increasing $\bar{\tau}$ from zero increases the flux correlations, this effect continues until some maximal level is reached when the saturation of absorption features starts to decrease the correlations seen in the flux. However the crucial point is that in the regime

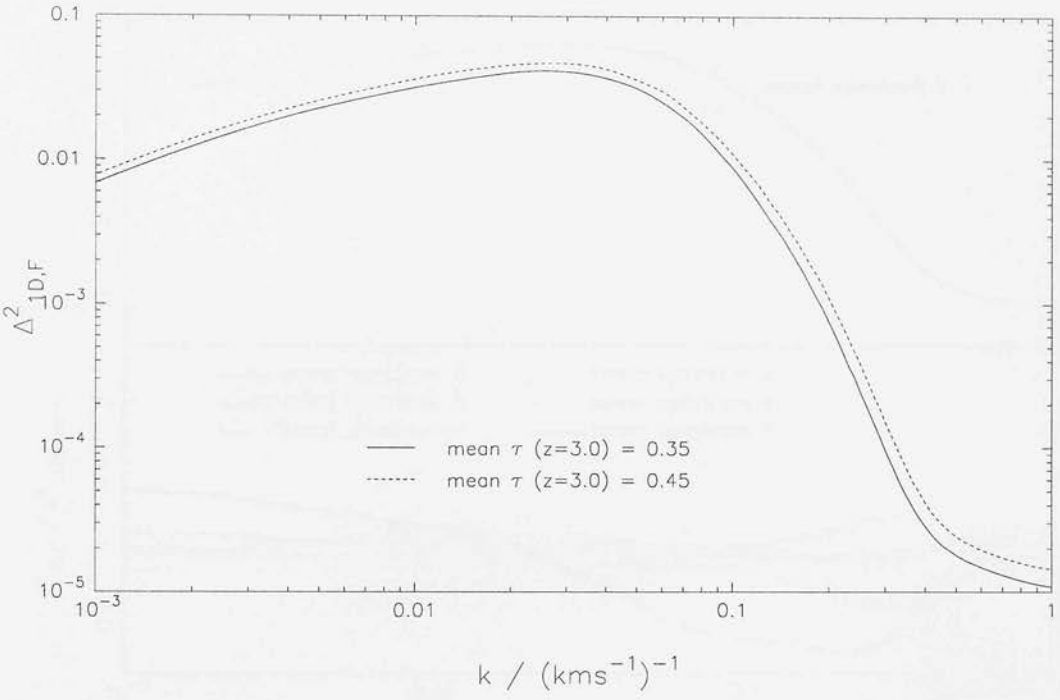


**Figure 3.1:** The effect of mean optical depth on the dimensionless flux power spectrum. Although there is an increase in amplitude when using the higher mean optical depth constraint, the shape of the flux power spectrum is the same in both cases.

of the observed values of mean optical depth the shape of the flux power spectrum remains unaltered by the value of this quantity. It must though be remembered that, as demonstrated by Croft *et al.*, the correct normalization of any recovered mass power spectrum does require using an accurate value of $\bar{\tau}$.

Figure 3.2 shows the average dimensionless flux power spectra obtained for the 7 different cosmological models. Since the flux power spectra are very similar the data has been plotted as a ratio of the first fiducial case, where $\Delta^2_{\rm 1D,F}(k)$ for this case is shown in the top plot. The bottom half of figure 3.2 demonstrates that the flux power spectra for all the ΛCDM models are the same to within about 10% in the regime of $k < 0.1$ (km s$^{-1}$)$^{-1}$. On smaller scales the lines separate out as the flux power spectra start to level out on different scales and at different values. The Einstein De-Sitter case (model 7) clearly differs from the other cases in a large region of $k$ space, showing

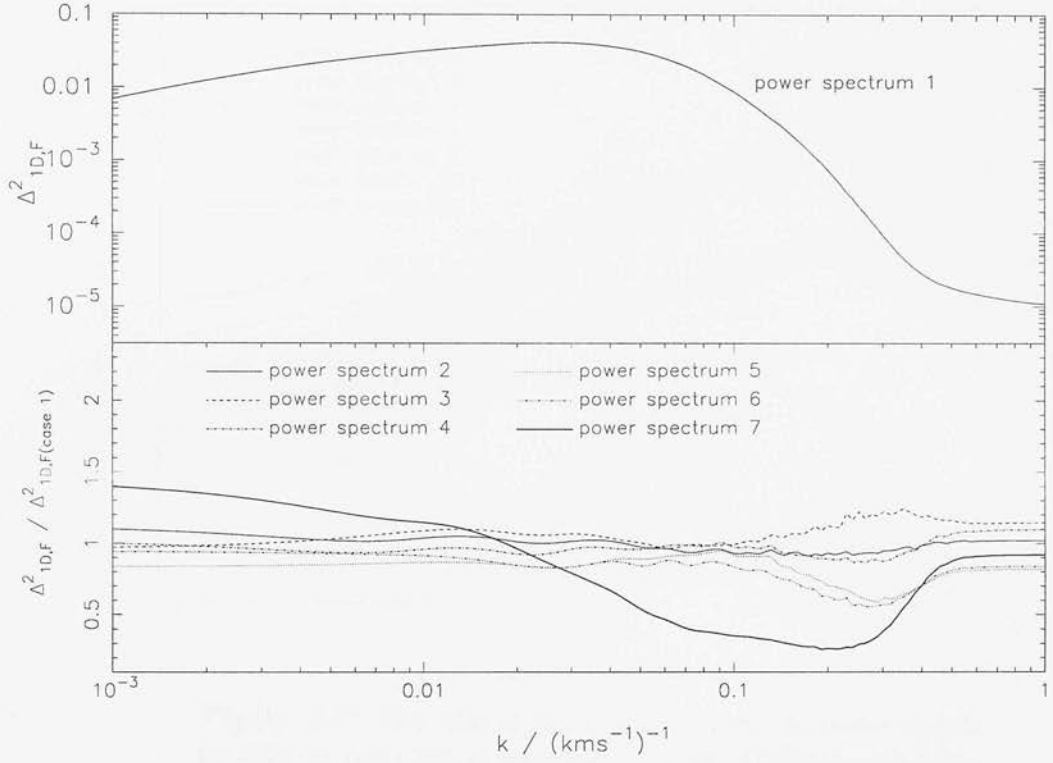


**Figure 3.2:** The dimensionless flux power spectra for different cosmological models. The top plot shows the recovered flux power spectrum for the model with parameters $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$, $\Gamma = 0.2$, $n = 1.0$ and $\sigma_8 = 1.0$. The second plot shows the ratio of the flux power spectra for each model (parameters given in table 3.1) relative to the flux power spectrum of the first case. While the matter dominated case (power spectrum 7) clearly deviates from the others, the ΛCDM models are all similar on scales $k < 0.1$ (km s$^{-1}$)$^{-1}$ before deviating slightly as they level out at high $k$.

more power on large scales and peaking at a lower value of $k$. Subsequently it has less power on intermediate scales, before levelling off in the same way as the vacuum energy dominated models at small scales.

The ratio of the flux and linear baryonic power spectra for each of the cases is shown in figure 3.3. The analysis will now focus only on the regime where $k < 0.1$ (km s$^{-1}$)$^{-1}$. This limit is chosen for several reasons, first it has been noted that on scales smaller than this results for $\Delta^2_{1D,F}(k)$ obtained from observational data diverge (Croft *et al.* 2000). This is thought to be the result of differing signal to noise levels. Secondly McDonald *et al.* (2000) found that for $k$ values larger than this the results were sensitive to the

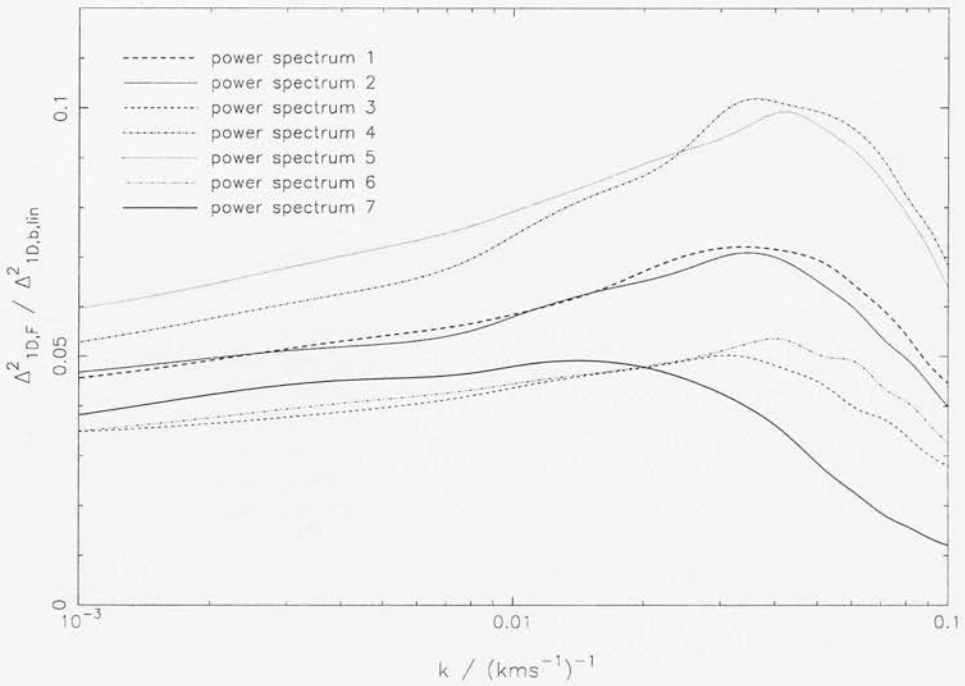


**Figure 3.3:** The ratio of flux and linear baryonic power spectra for different cosmological parameters. All the ΛCDM models follow the same trend where the ratio increases from low $k$ to a maximum at $k \sim 0.04$ $(\mathrm{km\,s^{-1}})^{-1}$, although the magnitudes are clearly model dependent. The matter dominated case shows less of an increase and peaks earlier. On large scales the ratio is relatively flat for all cases.

correct removal of contaminating metal lines in the spectra.

Since the flux power spectra are so alike the differences in the ratio of each case is strongly dependent on the linear baryonic $\Delta^2_{\mathrm{1D}}(k)$. As with the flux power spectra all the models, with the exception of the Einstein De-Sitter case, show the same trend. They rise from low $k$ to peak at $k \sim 0.04$ $(\mathrm{km\,s^{-1}})^{-1}$, then rapidly decline at larger $k$. Although exhibiting a similar overall form, the amplitude of this ratio is clearly model dependent, with the extreme cases differing by a factor of about 2 on all scales. Therefore there is no universal function which, when multiplied by any given $\Delta^2_{\mathrm{1D,F}}(k)$, gives the correctly normalized, corresponding linear mass power spectrum.

In the 3 dimensional analysis since the form of the flux power spectrum is found by differentiation, the gradient of any line rather than its absolute magnitude is the

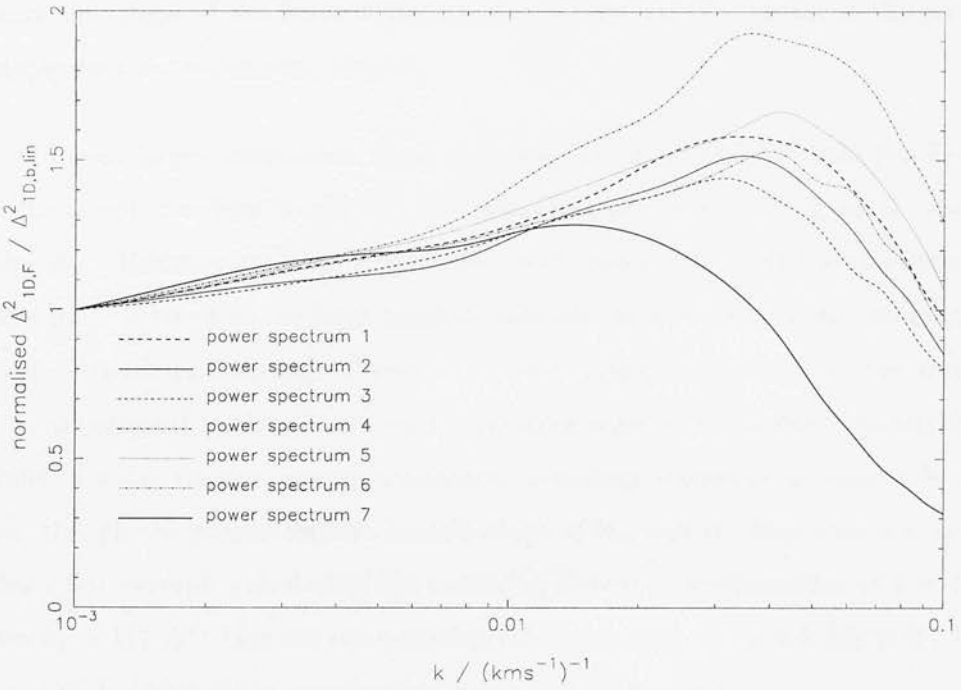


**Figure 3.4:** The renormalized ratio of flux and linear baryonic power spectra. Although showing considerable differences on small scales, on large scales each case is very similar.

important feature. However since 1D flux power are so similar, the inferred $\Delta^2_{\rm F}(k)$ will in turn be very alike. Therefore the bias parameter, $b(k)$, advocated by Croft *et al.* which is found by dividing the inferred flux power spectrum by the known mass power spectrum will still be subject to these differences in magnitude.

Leaving aside the issue of normalization, the relatively flat behaviour of the ratio of flux and linear baryonic power spectra on large scales does suggest that, in this regime, a mechanism aimed at recovering only the shape of the mass power spectrum may be successful. If recovery of only the shape of linear mass density power spectrum is attempted then the amplitude of the ratio found is unimportant. Figure 3.4 shows the ratio of the flux and linear baryonic power spectra for each case where this ratio has been renormalized such that $\Delta^2_{\rm 1D,F}(k = 1.0 \times 10^{-3}\ ({\rm km\ s^{-1}})^{-1}) = 1.0$. The difference between the lines is small on very large scales suggesting that a reasonably accurate recovery of the shape of the linear mass spectrum on these scales is possible. However the range over which this is valid is limited, with even the normalized ratios showing

significant deviations on scales of where $k \gtrsim 0.03$ $(\mathrm{km\,s^{-1}})^{-1}$. Even in the regime where the shape of the linear mass power spectrum can be recovered the problem of independent normalization remains.

On even larger scales than those analysed here it is expected that the flux power spectrum will continue to offer a good indication of the underlying mass density distribution. However probing larger scales with pencil beam surveys presents several difficulties. First of all for large redshift intervals there is not a linear relationship between distance and velocity. Therefore to set a distance scale requires the assumption of a cosmological model. Assuming the distance scale given by the currently favoured model of a flat vacuum energy dominated cosmology should be a reasonable assumption, though the precise relative contributions of $\Omega_m$ and $\Omega_\Lambda$ does have a considerable effect. For example calculating the comoving distance corresponding to $z = 2.9 - 3.1$ gives $r_c \sim 117\ h^{-1}\,\mathrm{Mpc}$ for the cosmological parameters $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$ rising to $r_c \sim 134\ h^{-1}\,\mathrm{Mpc}$ if the combination is $\Omega_m = 0.4$, $\Omega_\Lambda = 0.6$.

Another difficulty is that on very large scales it must be remembered that the mass power spectrum is evolving. On large scales where the correlations being measured are changing in amplitude over space this will complicate matters. Therefore, with this is in mind, any attempt to determine the mass density power spectrum on very large scales using the Ly-$\alpha$ forest must be performed with caution.

## 3.3 Deconvolution of spectra into Voigt profiles

Spectra generated from the prescription outlined in §3.1 can also be analysed by recovering line lists from them. A given spectrum can be fitted with blended Voigt profiles using the algorithms outlined below, which give a list of cloud positions, column densities and Doppler parameters. It should be noted that the 'cosmic web' paradigm for the IGM does not demand that the profiles found fully account for all the absorption seen. This is because the Ly-$\alpha$ absorption seen is accounted for as arising from a continuous fluctuating medium, rather than discrete clouds, and is therefore intrinsically non-Voigt (Outram, Carswell & Theuns 2000). The same applies to these simulations where the optical depths calculated from the underlying density field will not guarantee Voigt like features. However despite this, spectra (both real and simulated) can be well fitted using this technique since the majority of absorption does correspond well to these blended profiles. Once these line lists have been compiled they can then be studied in the ways detailed later in sections 3.4 and 3.5.

### 3.3.1 Line fitting algorithms

Extracting a list of absorption parameters for each spectrum has been done with two methods, SPECFIT and AutoVP, there follows a brief description of each algorithm. Although neither algorithm recovers the spectra perfectly they both provide acceptable fits to the data as is demonstrated later. The majority of fitting is done with SPECFIT which is by far the quicker of the two methods and AutoVP is applied to some data sets as a check that the results are not strongly affected by the line identification process. Codes for both of these algorithms were provided by Avery Meiksin.

#### SPECFIT

SPECFIT performs line fitting with the following steps:

1) First the spectrum is filtered using a wavelet transform. The smallest coefficients

are discarded to ensure a reduced $\chi^2$ of 1 between the filtered and unfiltered spectra.

2) Line candidates are identified as inflection points using the smoothed second derivative of the spectrum.

3) A spectral region is defined to fit around each line candidate, as a contiguous region with the flux smaller than a given $e^{-\tau_{\rm min}}$.

4) Overlapping regions are merged into a single continuous region.

5) A non-linear least squares fit is performed of candidate lines to the original spectrum.

### AutoVP

Below is a concise description of this algorithm, a more detailed account can be found in Davé *et al.* (1997).

1) Detection regions are identified by calculating an equivalent width per pixel across the spectrum as defined at the $i^{\rm th}$ pixel as

$$e_i = \Delta\lambda \left(1 - \frac{F_i}{C_i}\right), \tag{3.5}$$

where $\Delta\lambda$ is the pixel size and $F_i$ and $C_i$ represent the flux and calculated continuum in pixel $i$. The quantity $\sigma_{e_i}$ is then given by

$$\sigma_{e_i} = \frac{\Delta\lambda\sigma_{F_i}}{C_i}, \tag{3.6}$$

where $\sigma_{F_i}$ is the $1\sigma$ error in flux at pixel $i$. The $e_i$ values are then summed across 2 pixels either side of $i$ to give $E_i$, which has a standard deviation given by

$$\sigma^2_{E_i} = \sum_{n=i-2}^{i+2} \sigma^2_{e_n}. \tag{3.7}$$

Any point where $E_i > 8\sigma_{E_i}$ is identified as a detection region.

2) If the region is non-saturated, a single Voigt profile is placed at the minimum flux level in the detection region. $N_{\rm HI}$ and $b$ are adjusted by small decrements from large initial values until the feature is above the minimum flux everywhere. In saturated regions a feature is placed at the centre of saturation and $N_{\rm HI}$ and $b$ are adjusted to fit five pixels either side of the saturation.

3) The feature profiles from step 2 are then subtracted from the data to give a residual flux. Steps 1 and 2 are then repeated and a new residual flux calculated. This process is repeated until no new detections are found.

4) The resulting features are then subjected to simultaneous $\chi^2$ minimization of the three parameters $v_{\rm central}, N_{\rm HI}, b$.

5) A fit of the detection region is then attempted with one less line. If this results in a lower $\chi^2$ then the line is omitted, otherwise the fit retains the original feature list. This step ensures that the region is fitted with as few lines as possible whilst still maintaining the constraint that $\chi^2$ is minimised.

6) If $\chi^2 \lesssim 2$ per pixel then the algorithm is complete. If $\chi^2$ is above this level then a line is added to the point where the local contribution to $\chi^2$ is highest. This step is repeated until $\chi^2$ is suitably small.

### 3.3.2 Accuracy of line fitting

Once the line fitting has been performed the first step is to ensure that it is a good fit to the data. For each case a new spectrum is constructed where the Voigt profiles of the given line list are the only source of absorption. This new spectrum is subtracted from the original and the difference in the square of the flux at each point is summed to quantify the quality of the fit.

SPECFIT provides an excellent fit to many of spectra, one example of which is shown in figure 3.5, where the value of the fit parameter, described above, for this spectra is 2.33 $\times 10^{-4}$ per pixel. The top and middle graphs show the original and

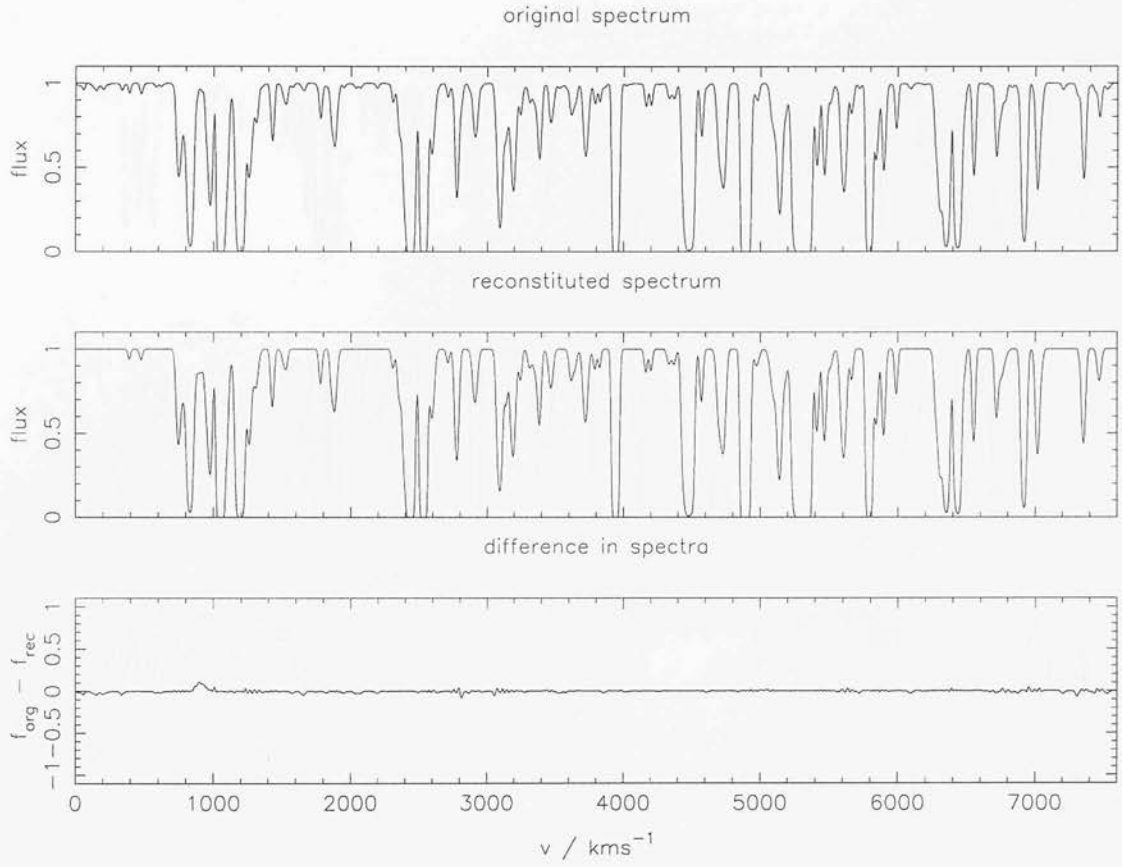


**Figure 3.5:** An example of a spectrum which is well fitted with Voigt profiles by SPECFIT. The top two plots show the original and reconstituted spectra respectively. The bottom plot shows the difference in flux levels between these two spectra at each point. The only source of discrepancy between the two corresponds to low level absorption in the original spectrum.

reconstituted spectra respectively, while the lower plot shows the difference in the flux of the two. As can be seen in figure 3.5 in the cases of the best fits, the only discrepancy between the original and reconstituted spectra arise in areas where the original spectrum has very low levels of absorption.

At the other extreme some spectra are very badly fitted by this algorithm. Figure 3.6 shows the realization with the best example of this and has a fit parameter of 0.207 per pixel. Clearly large areas of the reconstituted spectrum fail to match the features seen in the original. Instead in these regions the reconstituted spectrum have massive regions of saturation. This is thought to be a failure of the non-linear, least square fit

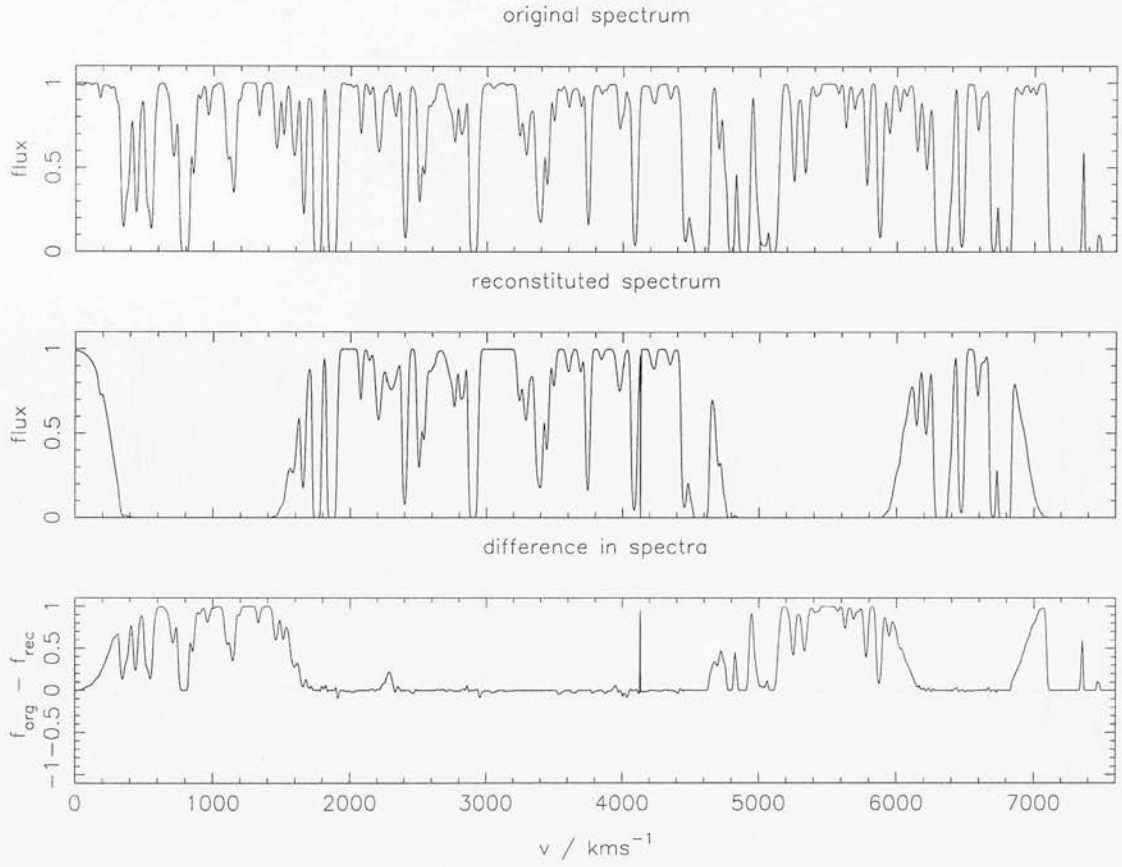


**Figure 3.6:** An example of a spectrum which is badly fitted with Voigt profiles by SPECFIT. Although some regions are well fitted, the reconstituted spectrum contains some large regions erroneously dominated by massive saturated features. In these regions the reconstituted spectrum bears no resemblance to the original.

routine. These regions are flagged by SPECFIT as having an unreasonably large value of $\chi^2$.

Inspection of the line lists produced by SPECFIT shows the list of absorbers contains some systematic errors. In some spectra this algorithm will spuriously fit some regions with features with very large column densities and Doppler parameters. In addition to these saturated features, SPECFIT also tends to include some tiny features, with very low column densities and Doppler parameters, which cover only a few pixels. An example of one of these features can be seen in figure 3.6 at $v \sim 4150$ km s$^{-1}$. These lines are removed since they are highly unphysical and their presence cannot be justified, and additionally, they often do not improve the fit to the original data.

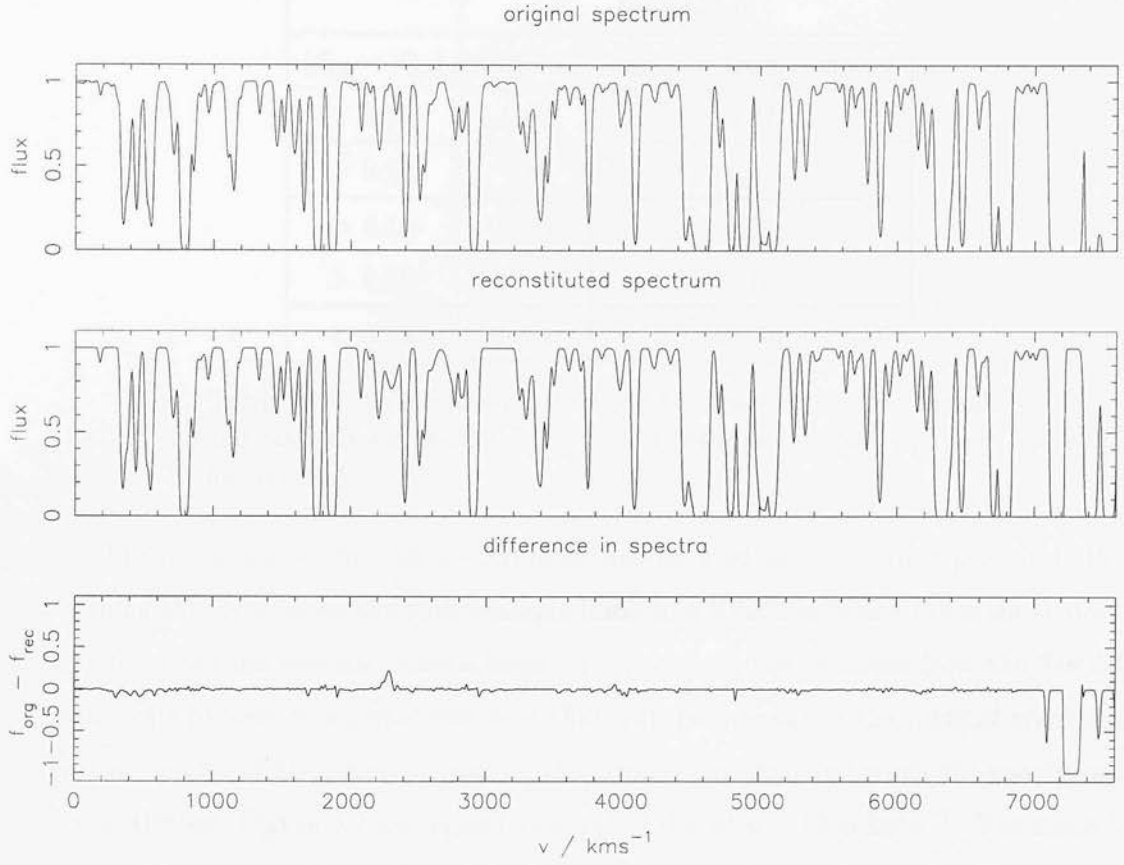


**Figure 3.7:** An example of the improvement in the Voigt profile fitting process using SPECFIT when spurious lines are identified and removed. The large saturated features are no longer present and the regions they covered are generally well fitted. The only remaining area of discrepancy not corresponding to low level absorption in the original spectrum is where one blended feature is only partially fitted.

Examining each spectrum and removing these spurious features from the line list improves the fit considerably. To do this an algorithm is applied where lines are excluded if their Doppler parameters lie outside a given range. It was found empirically that rejecting lines where $b < 10$ km s$^{-1}$ or $b > 100$ km s$^{-1}$, removed all spurious features while also retaining valid ones. Alternatively SPECFIT gives the option of specifying, at the outset, a range of $b$ values which can be considered acceptable. The result of rejecting spurious lines can be seen in figure 3.7 which shows the reconstituted spectrum, after this line removal process has been applied, for the same original spectrum as in figure 3.6.

| | % of pixels | |
|---|---|---|
| $\lvert f_{org} - f_{rec} \rvert$ | Original | Spurious lines removed |
| $> 0.50$ | 5.82 | 0.87 |
| $> 0.25$ | 7.84 | 1.62 |
| $> 0.10$ | 10.96 | 3.67 |
| $> 0.05$ | 15.19 | 7.10 |
| $< 0.05$ | 84.81 | 92.90 |

**Table 3.2:** A table showing the flux differences between the original and reconstituted spectra for SPECFIT, with and without spurious line removal.

The fit parameter for this spectrum is then reduced to 1.81 $\times 10^{-2}$ per pixel. Performing this process on the entire sample leads to a situation where the main errors in any fit arise from spectra where a large saturated feature is removed from the line list. After this process it is sometimes seen that features present in the original spectrum, in the vicinity of the removed feature, are erroneously excluded when the spectrum is reconstituted. Figure 3.7 has a good example of this at $v \sim 7200$ km s$^{-1}$. The mean fit parameter for the full sample of 100 spectra is decreased from to 4.59 $\times 10^{-2}$ per pixel down to 7.92 $\times 10^{-3}$ per pixel after the removal of these lines.

Table 3.2 shows the percentage of pixels with various differences in absolute flux levels between the original and reconstructed spectra for 100 realizations when using the original output from SPECFIT, and when spurious lines are removed. This table coupled with the fit parameter shows better fits are obtained when the spurious features are removed. This is also confirmed by eye when comparing figures 3.6 and 3.7.

In addition to quantifying the fits in the way previously discussed it is important to consider $\Delta^2_{1D,F}(k)$. Since this fundamental observable is used in the data analysis it is essential to ensure that this quantity in the reconstituted spectra closely matches the original value. Figure 3.8 shows this quantity for the sample when using SPECFIT and when the anomalous features have been removed. When the spurious lines are not removed there are considerable differences in both the high and low $k$ regime between the two flux power spectra. After the line removal process the original and recovered

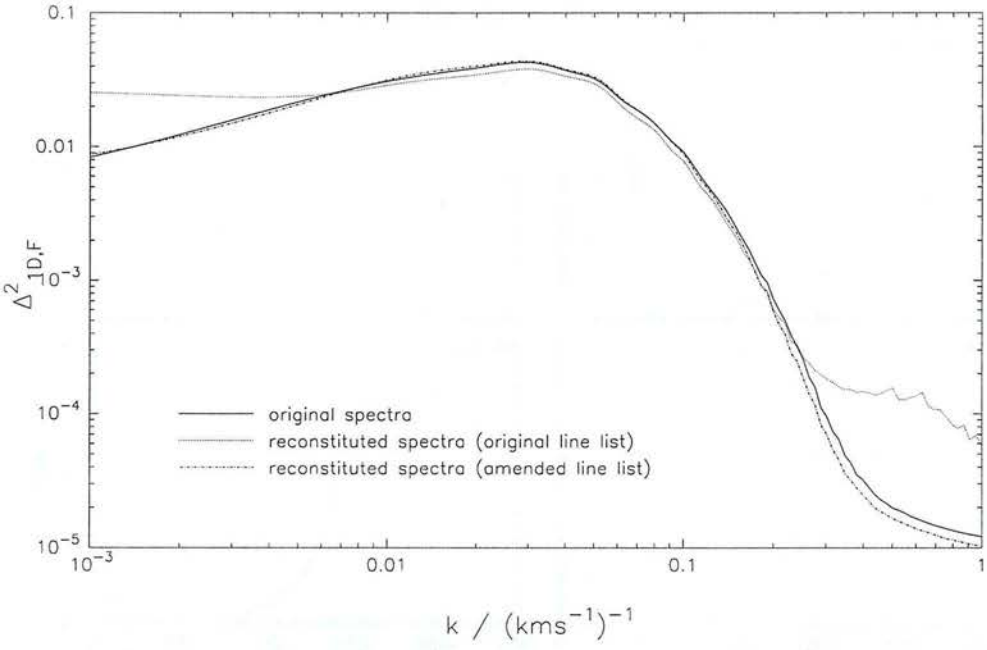


**Figure 3.8:** Averaged dimensionless flux power spectra for original and reconstructed spectra using SPECFIT. Using the original line list the flux power spectra match well on intermediate scales but there is a discrepancy of nearly an order of magnitude on small scales. On large scales there is a less pronounced yet still considerable difference. Once the spurious features have been removed there is a much better fit on all scales, though suppression of small scale power is seen.

$\Delta^2_{1D,F}(k)$ are in agreement for low $k$ values. At large $k$ the amended version of SPECFIT does not recover the flux power spectrum exactly but instead suppresses small scale power. This is not significant however since recovery of the mass power spectrum on these scales is not practical due to observational constraints (e.g McDonald *et al.* 2000; Croft *et al.* 2000). Therefore the errors introduced when using SPECFIT do not affect the data strongly in the regime of interest and using these line lists in the analysis is justified.

### 3.3.3 Comparison of line fitting algorithms

As previously discussed, to ensure that the line list data obtained is not dependent on the line fitting algorithm used, some of the spectra are also fitted using AutoVP. In

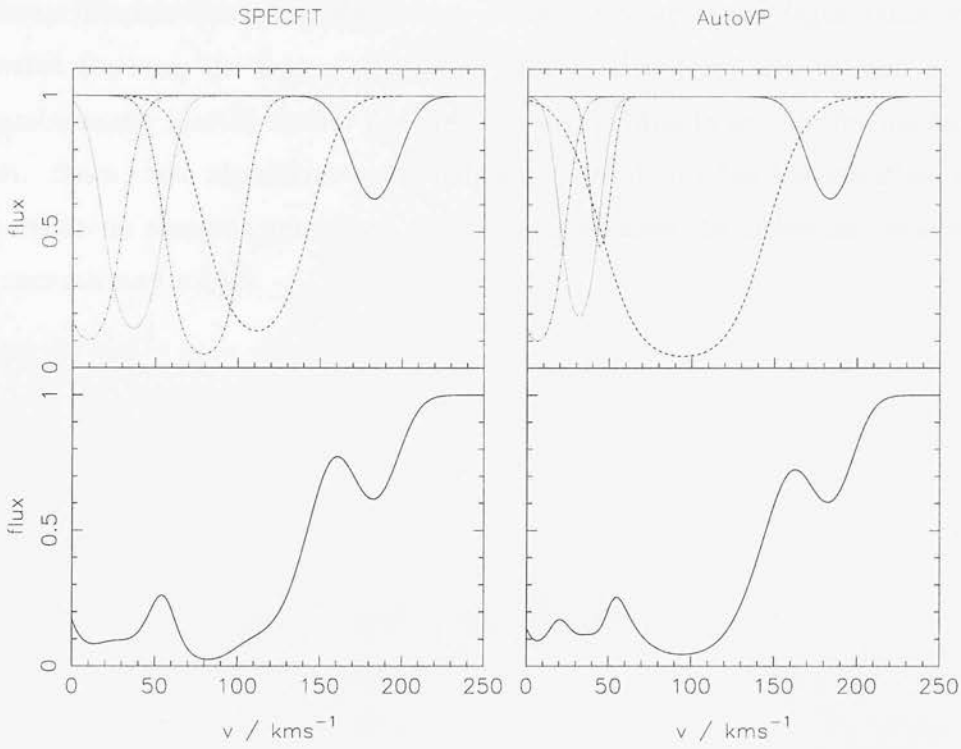


**Figure 3.9:** An example of degeneracy when fitting a blended line profile. The top plots show the individual absorption components and the bottom plots show the resultant blended profile. Although the component features are different the resultant profile is very similar.

general the line lists obtained from SPECFIT and AutoVP are similar, although some differences are evident. Unsurprisingly there are minor differences in the parameter values of the lines given, although these are usually small and unimportant. In addition the line lists obtained from AutoVP are invariably longer than those obtained by SPECFIT. Nearly all of these additional features have low column densities and correspond to areas of low level absorption in the original spectrum. As previously discussed these regions are are not accurately recovered by SPECFIT.

Significantly some discrepancy arises between the line lists where heavily blended features are fitted differently. The fits for heavily blended features can be degenerate as demonstrated in figure 3.9. This does not present a problem however since, in later analysis of the line positions, the individual components of these blended lines still mark regions of high underlying baryon density.

Despite the presence of these discrepancies there is in general an excellent agreement

between the two line fitting algorithms. With exception of the degeneracies found in blended features, the lists of absorption lines produced by AutoVP and SPECFIT unambiguously identify similar line positions and profiles to account for the flux levels seen. Since these algorithms apply different methods for line identification and the results are in excellent agreement, the line lists produced by either can be considered as accurate and robust.

## 3.4 Power spectrum of line positions

### 3.4.1 Method

Once a line list has been obtained it is straightforward to calculate $P_{\mathrm{1D},\ell}(k)$ using a binary distribution of $\delta_\ell$ as discussed in §2.1.3. Using a grid of the same resolution as the spectrum is generated on for this calculation can lead to a situation where more than one line centre is found in some given pixel. Since the only occupancy numbers allowed in the binary distribution are 0 or 1 this is potentially a problem. When this occurs the grid can be made finer though, in this case, finding two absorption lines so close together is rare enough (approximately one occurrence in every 10 spectra) that ignoring the presence of one of the lines has a negligible effect on the calculated $P_{\mathrm{1D},\ell}(k)$. If the centres of absorption lines offer either a Poisson sampling of the mass field or act as reliable markers of the peaks in this field then $\Delta^2_{1D,\ell}(k)$ should provide a good way to recover the shape of the underlying mass density field. The amplitude of the recovered field however would require some independent form of normalization. This technique was applied to a sample of 500 spectra where the line fitting was performed with SPECFIT.

### 3.4.2 Results and discussion

The first thing to note is that the small number of objects, inherent in using absorption line centres in this way, means the scatter on individual realizations is very large. In spite of this the number of realizations used here allows a suitable line to be fitted to the data. Since the normalization of the recovered $\Delta^2_{1D,\ell}(k)$ is arbitrary its amplitude is set to be equal to that of the underlying mass distribution at $k = 1.0 \times 10^{-3}$ $(\mathrm{km\ s^{-1}})^{-1}$. Figure 3.10 shows this renormalized $\Delta^2_{1D,\ell}(k)$ compared with $\Delta^2_{\mathrm{1D}}(k)$. It is clear that as a method of recovering the mass field this technique fails dramatically on scales of $k > 3 \times 10^{-3}$ $(\mathrm{km\ s^{-1}})^{-1}$. Furthermore studying $P_{\mathrm{1D},\ell}(k)$ on scales smaller than those shown on the graph reveal that on scales $k > 0.2$ $(\mathrm{km\ s^{-1}})^{-1}$ this power spectrum tends to a constant. This constant is larger than the recovered signal on large scales and so

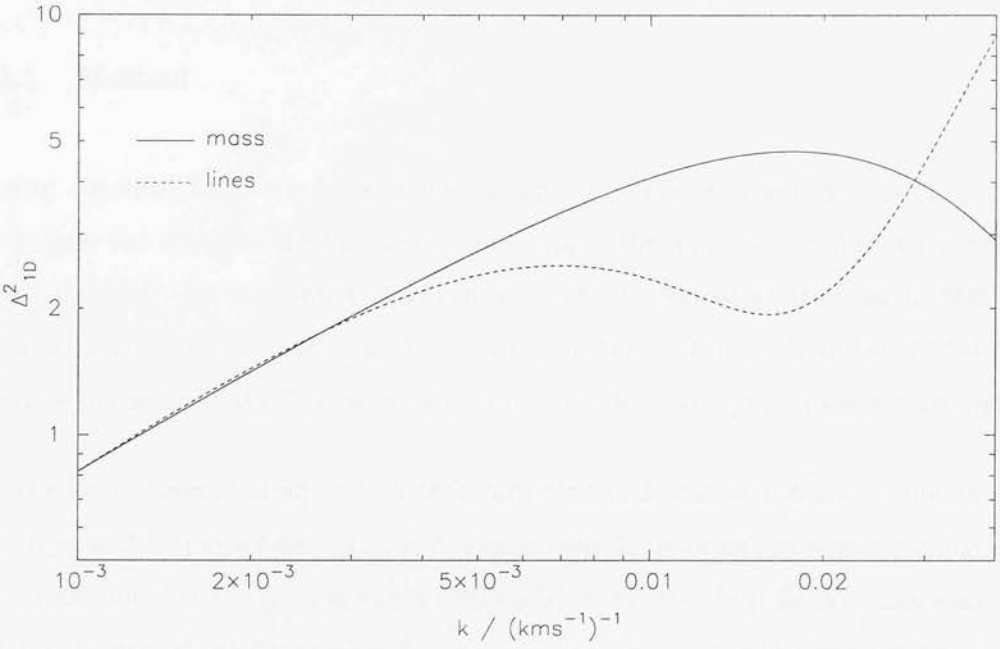


**Figure 3.10:** The dimensionless power spectrum of the underlying mass density field and the recovered line power spectrum which has been renormalized to match the mass power spectrum at $k = 1.0 \times 10^{-3}$ $(\mathrm{km\ s^{-1}})^{-1}$. Although on very large scales these quantities are closely matched a large discrepancy becomes apparent at $k \sim 3.0 \times 10^{-3}$ $(\mathrm{km\ s^{-1}})^{-1}$.

even the results from larger scales must be considered dubious.

The reason for the failure of this method to recover the mass power spectrum is thought to be due to the effects of Doppler broadening of the lines. If the flux in each pixel was dependent only on the density at that point, rather than the general surrounding region, then marking the points of greatest absorption would give a much better indicator of the underlying mass. Even though the Doppler broadening is a relatively local effect its distortion of the link between underlying density and optical depth at any given point is enough to render this method highly unreliable.

## 3.5 Flux power spectrum of scrambled line positions

### 3.5.1 Method

Having obtained line lists for a sample of spectra, this information can be used to investigate the effect on the flux power spectrum of the different profiles and positions of the lines. If the flux power spectrum seen can be better understood in terms of the absorption systems that cause it, then this in turn will lead to an insight into the relationship between the flux power spectrum and the underlying mass density field.

The lines corresponding to each spectrum were maintained in pairs of column density, $N_{\rm HI}$, and Doppler parameter, $b$, but their velocity parameters were scrambled. So for a given line list the centres of the absorption lines were kept in the same positions but the profiles of the absorption features were placed randomly amongst these positions. A spectrum was then generated from this new absorber configuration, and its flux power spectrum calculated. The flux power spectra from each randomization were averaged to find a sample mean. This process was then repeated 100 times and the mean flux power spectrum for all the realizations was found.

### 3.5.2 Results and discussion

Figure 3.11 shows the original flux power spectrum and that obtained from the scrambled line profiles for a sample of 500 spectra. Scrambling the spatial positions of the absorption profiles suppresses power on large scales, while boosting it by a factor by up to 2 on intermediate scales. The power spectra are well matched in the regime $0.1\ (\mathrm{km\ s^{-1}})^{-1} < k < 0.3\ (\mathrm{km\ s^{-1}})^{-1}$, while for $k$ values above this small scale power is boosted by the scrambling process.

If the Croft *et al.* method of recovering the mass power spectra is valid, then a change in the flux power spectrum when a line list is scrambled is required. Investigating the nature of this change, however, should give a thorough understanding of the relationship between the individual absorption systems and the flux power spectrum.

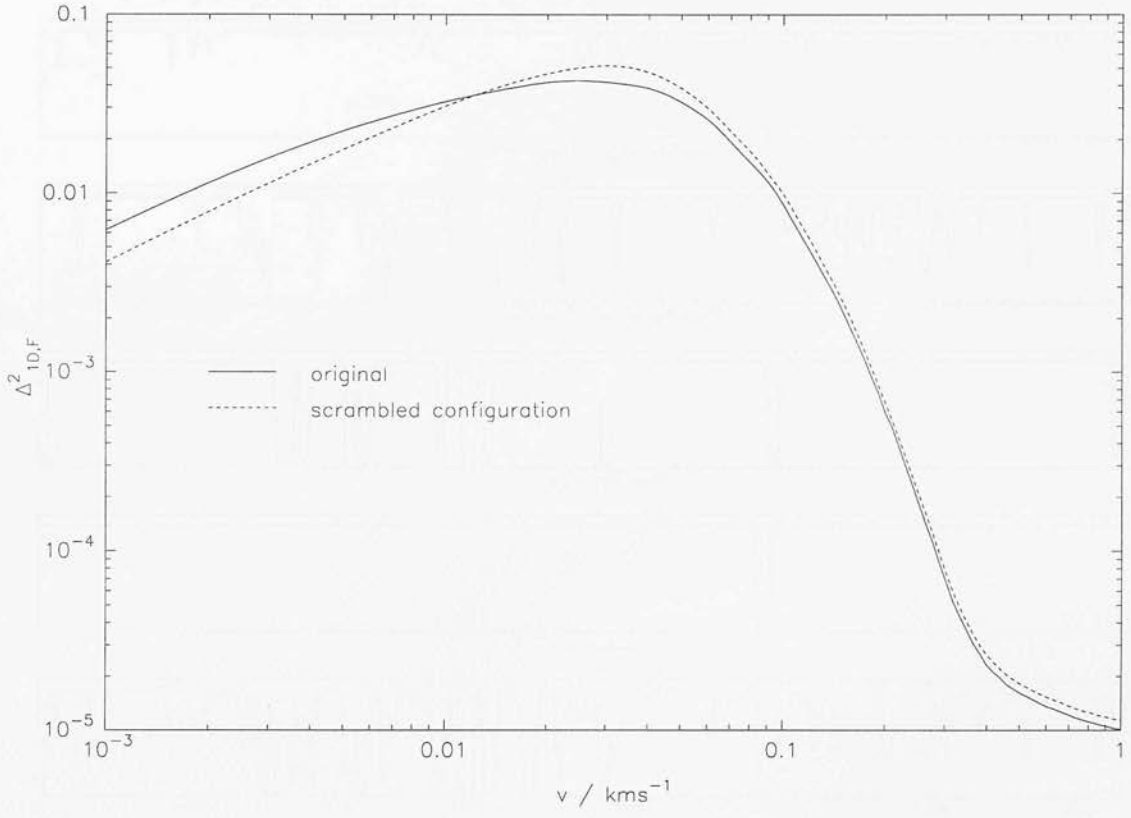


**Figure 3.11:** The dimensionless flux power spectra for the scrambled and original line positions. The scrambling process causes a change of power on many scales. Large scale power is suppressed while an increase in power is seen on small and some intermediate scales. The power spectra are very similar for $k$ values between $0.1 - 0.3$ $(\mathrm{km\,s^{-1}})^{-1}$.

Slightly different scrambling processes were also performed in order to study the cause of the change in the flux power spectrum further. The difference is that this time only certain subsets of lines had their profiles scrambled. Each line list was split into subsets depending on the column densities of the absorbers. Figure 3.12 shows a sample spectrum broken down into lines of different ranges of column density. The absorption features from four different regimes of $N_{\mathrm{HI}}$ are shown, while the bottom plot shows the spectrum obtained when all the features are included. As would be expected the lower column density lines are more numerous, yet their contributions to the spectrum are not as dramatic as those seen from the high column density systems.

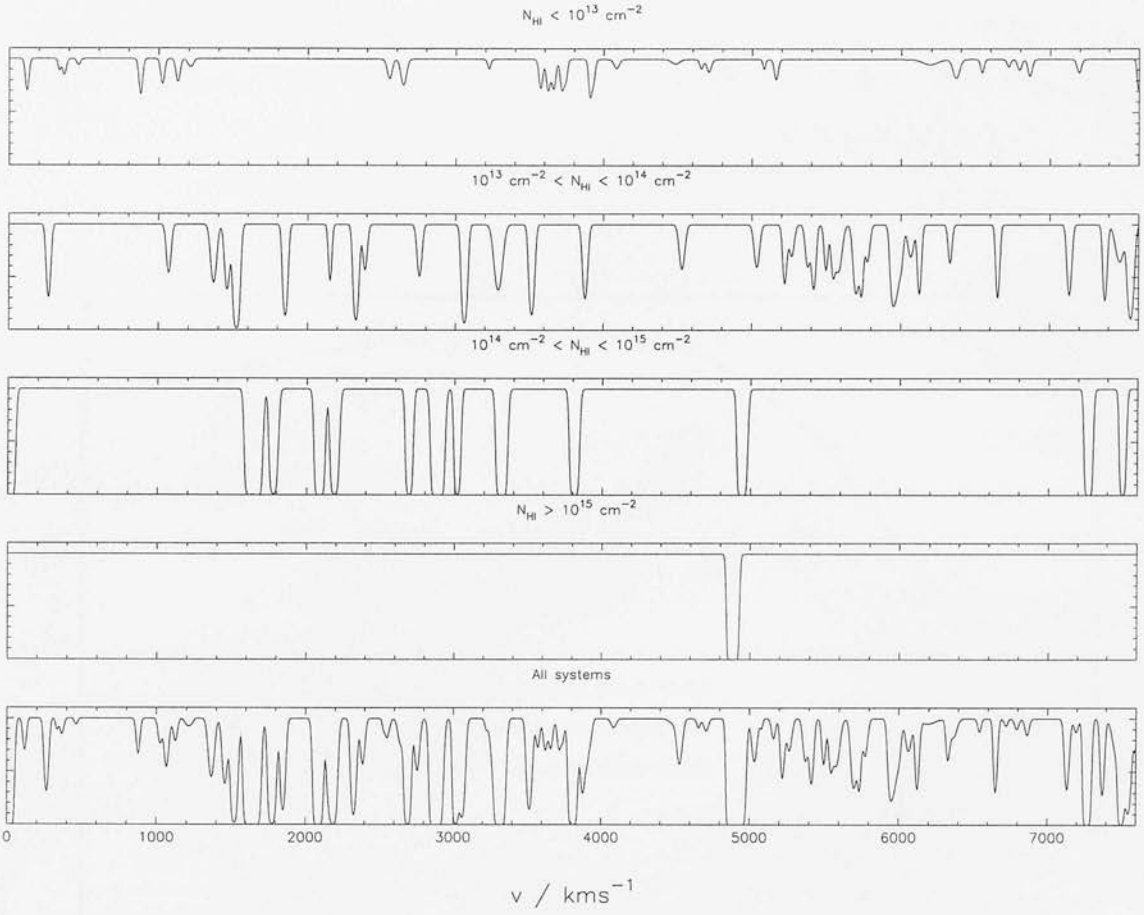


**Figure 3.12:** The absorption profiles of a spectrum broken down into different regimes of column density. The top four plots show the profiles of features from given ranges of column density while the bottom plot shows the spectrum obtained when all the features are included.

These column density subsets were then scrambled, while the configuration of the remaining absorption systems was maintained. Figure 3.13 shows the resulting flux power spectra as a ratio of the original flux power spectrum. This is done since the lines are very hard to distinguish when plotted on the normal log-log plot. The results of scrambling each decade of column densities results in flux power spectra which are the same, to within about 5%, of the original in every case.

This in itself is interesting since although the column densities being scrambled in each case are similar, each subset of absorbers contains a range of Doppler parameters. The curve of growth, shown in figure 1.3, shows that many of the lines in a given

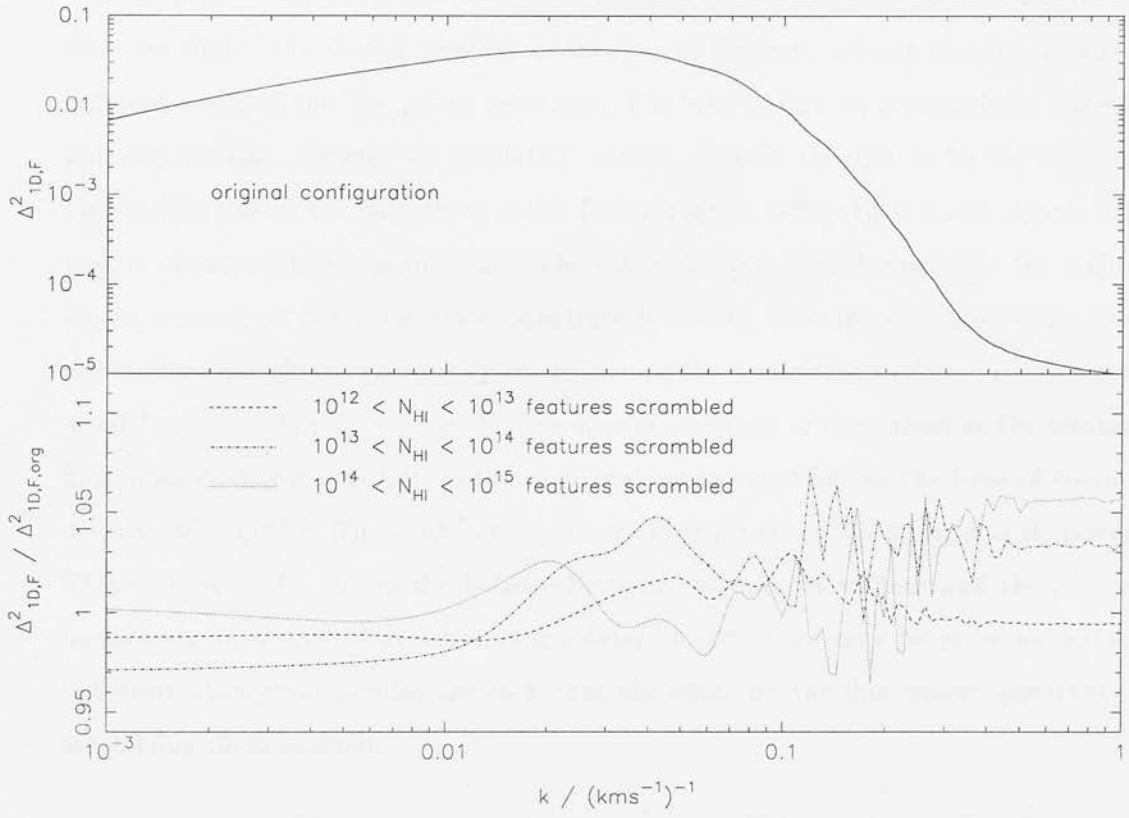


**Figure 3.13:** The flux power spectra obtained when subsets of the absorption features are scrambled. The top plot shows the flux power spectrum of the original configuration while the bottom plot shows the ratio of the new flux power spectra and the original. Although all the ratios are close to one it can clearly be seen that power on different scales is affected differently by scrambling different subsets of the absorption features. The wiggles seen on intermediate scales are thought to be artifacts of the spline line fitting.

subset will have significant differences in their equivalent widths. However it is seen that scrambling these subsets of features is relatively unimportant to the resulting flux power spectrum. This suggests that the Doppler broadening of lines, which may have been problematic (as discussed in chapter 2), does not have a dramatic effect when the Croft *et al.* inversion method is applied.

Another interesting feature to be noted in figure 3.13 is that, although the differences seen are slight, the spatial position of features of different column density do affect different scales of the flux power spectrum. The information on intermediate scales is not very reliable, showing an oscillating nature. This is thought to be the result of systematics due to the smoothing spline fitting routine rather than a real signal. The regime of most interest is the scales where $k < 0.1$ (km s$^{-1}$)$^{-1}$ since this the regime where recovery of the mass power spectrum is usually attempted. These large scale correlations are almost completely insensitive to the scrambling of lines in the regime of $10^{14}$ cm$^{-2} < N_{\rm HI} < 10^{15}$ cm$^{-2}$. The spatial positions of the subset of the weakest lines does contribute slightly to the large scale structure whereas the lines of column density $10^{13}$ cm$^{-2} < N_{\rm HI} < 10^{14}$ cm$^{-2}$ has the greatest effect on the large scale power. This is likely to be due to the balance between the number of lines and the relative importance of a single feature. Lower column density lines may be more numerous, but their absorption profiles are such that the effect on the flux power spectrum of scrambling them is small.

To study the difference in the subsets of absorption features further, the two-point correlation function of the lines in the different column density regimes was calculated. To evaluate the two-point correlation function, $\xi(v)$, the estimator advocated by Hamilton (1993) was used. Here two catalogues are used, the data catalogue and a larger random catalogue. The correlation function was then estimated using the expression

$$\xi(v) = \frac{DD(v) - 2DR(v) + RR(v)}{RR(v)}, \tag{3.8}$$

where $DD(v)$ and $RR(v)$ represents the average number of pairs found within a given distance, $v$, of each other, divided by the number of pairs in the entire sample, for the data and random catalogues respectively. $DR(v)$ is the same quantity for the cross pairs across the catalogues. Other estimators of $\xi(v)$ were also tested and found

to be in good agreement with the expression given in equation 3.8.

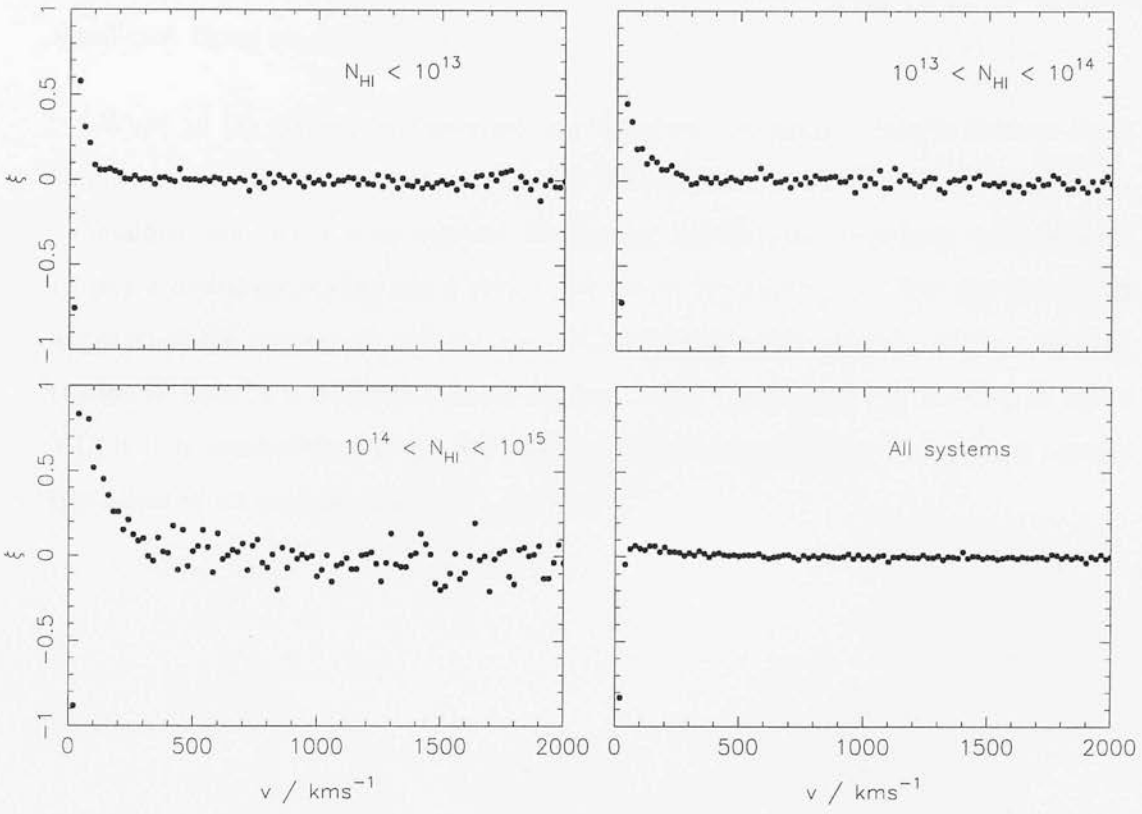


**Figure 3.14:** The two-point correlation function for different column density regimes. The plot shows that although the amplitude of $\xi(v)$ is small when all the features are considered, different column density regimes do show differing degrees of clustering on small scales. The noise increases as the number of lines decreases and the result for $10^{14}$ cm$^{-2} < N_{\rm HI} < 10^{15}$ cm$^{-2}$ shows considerable scatter. There is a strong anti-correlation on very small scales since line centres are rarely extremely close.

Figure 3.14 shows the recovered values of $\xi(v)$, for 100 bins of width $v = 20$ km s$^{-1}$. The first thing to note is the strong anti-correlation seen at 20 km s$^{-1}$, this is since the lines themselves have widths comparable to this scale and so the only place lines are in such close proximity are in heavily blended systems. In all cases the value of $\xi(v)$ on scales of over several hundred km s$^{-1}$ is consistent with zero.

The effect of the number of features in each subset can be clearly seen with the

increase of scatter in $\xi(v)$ as the column densities increase (and thus the number of objects decreases). The lines showing the strongest clustering on small scales correspond to the highest decade of column densities considered. Each subset, however, shows a significant signal on small scales.

When all the systems are included the two-point correlation tends to zero on much smaller scales than when only a subset of lines is considered. This appears to be an anomalous result since three subsets, all showing considerable clustering, are combined to give a distribution that has a very weak two-point correlation. However this result arises since the regions of voids for one column density regime correspond to a concentration of lines in a different column density range. This effect can be seen in figure 3.12 if it is remembered that each individual absorption feature is weighted equally regardless of its column density.

## 3.6 Peculiar velocities

A complication to the analysis that has not yet been addressed in this thesis is that of peculiar velocities. The recovered mass power spectrum is required in real space whereas the flux power spectrum of the absorption spectra is calculated in velocity space. The presence of a peculiar velocity field leads to a distorting effect between real and velocity space and induces anisotropies in the 3 dimensional flux power spectrum.

The effects of the peculiar velocity field cannot be properly considered in 1 dimension alone, since the magnitude of the field along directions other than the line of sight can have an effect. For example a density fluctuation can occur where there is no velocity fluctuation along the line of sight if the gas is infalling perpendicularly. Conversely a point noted by McGill (1990) is that peculiar velocities can give rise to apparent absorption features in velocity space even in the absence of density fluctuations. In this case material can be converging along the line of sight, giving rise to an absorption feature, whilst perpendicular to this infall the gas is diverging such that the density contrast is zero. Therefore the effects of a peculiar velocity field are best studied by 3 dimensional $N$-body simulations. Failing this more simplistic approaches can be applied in 3 dimensions, for example the Zel'dovich (1970) approximation. Essentially this is first order perturbation theory expressed in Lagrangian space whose results show that in a triaxial system collapse will occur along the shortest axis. An excellent approximation in highly asymmetric systems, this approach works least well when considering situations of complete spherical symmetry.

A peculiar velocity field has several effects on the absorption spectra produced as has been shown by $N$-body simulations (e.g. Zhang *et al.* 1998). Firstly, and most obviously, systematic offsets occur in the positions of an absorption feature in velocity space, due to the density fluctuations responsible for the absorption features following the bulk flow motions. At a redshift of 3.0 the amplitude of systematic peculiar velocities can be as high as several hundred km $s^{-1}$. An additional, more subtle, effect is that the bulk motions within gas clouds can lead to a substantial alteration of the Doppler parameter, $b$. In the absence of peculiar velocities the Doppler parameter is determined solely by the temperature of the gas cloud (see equation 1.50), but this is boosted by the non-

thermal motion arising from the peculiar velocities.

Modelling peculiar velocities in 1 dimensional simulations cannot be tackled exactly, but several different approximations exist in the literature. One approach is to neglect the effect of non-linearities and work solely in the linear regime. This allows the 3 dimensional density and velocity power spectra to be linked via the continuity equation. The usual transformation can then be made to 1 dimension power spectra and Gaussian realizations can then give correlated 1 dimensional density and velocity fields (e.g. Bi 1993). This procedure introduces different line blending and clustering effects as the position of the absorption features are seen by an observer to be shifted. The effects of peculiar velocities arising in the non-linear regime can be crudely modelled by filtering the spectrum in velocity space with a Gaussian, with the effect that a cut off in the power spectrum is introduced at high $k$ (e.g. Kaiser & Peacock 1991).

The related problem of taking 1 dimensional data and inferring the peculiar velocity field along the line of sight (thus allowing the effects to be corrected for) was tackled by Nusser & Hehnelt (1999). The method used is an iterative approach where a density field is inferred from the spectra and a velocity field calculated, the density field is then recalculated with the new peculiar velocity components and the process repeated until convergence is reached. Reconstructing the velocity field from the density field is performed using the same 1 dimensional approximation of the 3 dimensional fluid equations as Bi (1993) outlined above. This approach gives a probable 1 dimensional velocity field consistent with the line of sight density field but is not rigorous due to the 3 dimensional effects discussed above being neglected.

As previously mentioned Hui (1999) showed that neglecting the effects of peculiar velocities in the initial method of Croft *et al.* (1998) leads to underestimation of the steepness of the recovered mass power spectrum. However this problem should not exist with the most recent technique where the mass power spectrum is found using the bias parameter, $b(k)$. Since the bias parameter is calculated from $N$-body simulations where peculiar velocities are well modelled any error arising from these effects should be slight, especially compared to the errors introduced by, for example, uncertainties in the temperature, mean optical depth and reionization history of the IGM.

# Chapter 4

# The statistics of the flux power spectrum

In this chapter the statistical properties of the simulations from chapter 3 will be discussed. Firstly analysis of the correlation matrix of the inferred 3 dimensional flux power spectrum will be presented. This is followed by a comparison of two different methods of inferring the 3 dimensional flux power spectrum from its 1 dimensional counterpart. Analysis is then presented showing the relative likelihoods, between cosmological models, when using measurements of $P_{\mathrm{1D,F}}(k)$ directly, rather than any inversion techniques. Finally the relationship between the mass and flux variances found in individual realizations is briefly discussed.

## 4.1 Correlations in the flux power spectrum

In general to test models against observable data it is necessary to quantify the likelihood that a set of data would arise for a given model. Assuming a set of $N$ data points, $y_i$ $(i = 1, N)$, and a set of models which predict values for these data, $y_{i,\mathrm{mod}}$, then the question that needs to be asked for each model is what is the probability that this data set could have occurred? Then the likelihood of the parameters given the data is

identified as being proportional to the probability of the data given the parameters.

If each data point is assumed to be measured independently, with a Gaussian measurement error, then the probability of the data is the product of the probabilities of each point, leading to an expression for likelihood given by

$$\mathcal{L} = \exp(-\chi^2/2) = \prod_{i=1}^{N} \left\{ \exp\left[ -\frac{1}{2} \left( \frac{y_i - y_{i,\mathrm{mod}}}{\sigma_i} \right)^2 \right] \right\}, \tag{4.1}$$

or alternatively

$$\ln \mathcal{L} = -\chi^2/2 = -\sum_{i=1}^{N} (y_i - y_{i,\mathrm{mod}})^2 / 2\sigma_i^2. \tag{4.2}$$

The parameter $\chi^2$ has been introduced in the above expressions and should be minimized to find the maximum likelihood estimates of the model parameters.

The assumption of Gaussian distributed errors is usually justified by appealing to the central limit theorem. This fundamental theorem in statistics states that, regardless of the underlying probability distribution function, if a large number of points are sampled then the distribution of these sampled points tends towards a Gaussian. It should be noted that there are distributions for which this theorem does not hold true, however that it does for so many distributions is the reason that the Gaussian distribution is so important in statistics.

However another assumption in the derivation of equation 4.2 is that the individual data points, $y_i$, are uncorrelated. If this is not the case a more rigorous analysis is required where each data point is not simply considered in isolation, but with respect to the other points. A correlation matrix $C_{ij}$ can be defined as

$$C_{ij} \equiv \left\langle \frac{(y_i - y_{i,\mathrm{mod}})}{\sigma_i} \frac{(y_j - y_{j,\mathrm{mod}})}{\sigma_j} \right\rangle = \begin{pmatrix} 1 & r_{12} & \cdots & r_{1N} \\ r_{12} & 1 & \cdots & r_{2N} \\ \vdots & \vdots & & \\ r_{1N} & r_{2N} & & 1 \end{pmatrix}. \tag{4.3}$$

Where $r_{ij}$ is known as the correlation coefficient between the distributions of $i$ and $j$. The values for this coefficient range between 1 and $-1$ corresponding to perfect

correlation ($y_i - y_{i,\mathrm{mod}} = y_j - y_{j,\mathrm{mod}}$) and anti-correlation ($y_i - y_{i,\mathrm{mod}} = -(y_j - y_{j,\mathrm{mod}})$) respectively between $y_i$ and $y_j$. Then, in the general case, the relative likelihood that a set of data is consistent with a model is given by

$$\ln \mathcal{L} = -\frac{1}{2} \sum_{ij} (y_i - y_{i,\mathrm{mod}}) C_{ij}^{-1} (y_j - y_{j,\mathrm{mod}}). \tag{4.4}$$

In practice it is often convenient to use an approximate analysis where the correlation matrix is simply equal to the identity matrix (in which case the above expression is equivalent to that shown in equation 4.2). However this can only be justified if it is shown that $r_{ij} \simeq 0$ for all combinations of $i \neq j$. Although even in the case where a set of data are truly independent, the correlation coefficients for $N$ realizations, will have a standard deviation given by $1/\sqrt{N}$, so will in general, be scattered about zero rather than being equal to zero, although the scatter will decrease as the number of realizations is increased.

The case of interest with respect to the simulations presented in chapter 3 is to what level the flux power spectrum measurements on different $k$ scales are independent of each other. If there are significant correlations between the flux power spectrum on different scales then any likelihood analysis performed which neglects these correlations will be inaccurate.

Croft *et al.* (2000) show a correlation matrix calculated from a subset of their data set using the jackknife method. This technique provides good error estimates for many data sets, though is not infallible. Given a data set one point is removed from the sample at a time, explicitly the data set remaining after deletion can be represented by

$$\mathbf{x}_{(i)} = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n). \tag{4.5}$$

Now letting $s_{(i)} = s_{(\mathbf{x}_{(i)})}$ represent the value of a given statistic for the data corresponding to the deleted point, then the jackknife estimate for the standard error of $s(\mathbf{x})$ is given by

$$\hat{\mathrm{se}}_{\mathrm{jack}}\{s\} = \left[ \frac{n-1}{n} \sum_{i=1}^{n} (s_{(i)} - s_{()})^2 \right]^{1/2}, \tag{4.6}$$

where

$$s_{()} = \sum_{i=1}^{n} \frac{s_{(i)}}{n}. \tag{4.7}$$

The correlations in the flux power spectrum between given wave numbers in a sample of simulations can be computed exactly. Explicitly the correlation coefficients for the flux power spectrum between two wave numbers $k_i$ and $k_j$ is

$$r_{ij} = \left\langle \frac{[P_{\rm F}(k_i) - P_{\rm F}^{\rm mod}(k_i)]}{\sigma_i} \frac{[P_{\rm F}(k_j) - P_{\rm F}^{\rm mod}(k_j)]}{\sigma_j} \right\rangle, \tag{4.8}$$

where $\sigma_i^2$ is simply the variance of $P_{\rm F}(k_i)$. However in this case the model value of $P_{\rm F}(k_i)$ is simply the mean value of this parameter over many realizations. Therefore substituting this into the above expression and writing $\sigma_i$ in terms of the flux power spectrum gives

$$r_{ij} = \frac{\langle P_{\rm F}(k_i) P_{\rm F}(k_j) \rangle - \langle P_{\rm F}(k_i) \rangle \langle P_{\rm F}(k_j) \rangle}{[\langle P_{\rm F}^2(k_i) \rangle - \langle P_{\rm F}(k_i) \rangle^2]^{1/2} [\langle P_{\rm F}^2(k_j) \rangle - \langle P_{\rm F}(k_j) \rangle^2]^{1/2}}. \tag{4.9}$$

So for a given number of realizations there is a straightforward way to calculate the correlation coefficients $r_{ij}$ between wave scales $k_i$ and $k_j$.

This calculation was performed for spectra generated using a similar prescription to that given in chapter 3. However in order to compare the results with those in Croft *et al.* some parameters were altered to match their data set. Therefore the redshift range was chosen to be $z = 2.5 - 2.9$. Despite the redshift range being considerable the evolution of the mass power spectrum over this range has been ignored and the normalization redshift was set at $z = 2.7$. The constant of proportionality in equation 3.2 was determined using the constraint $\bar{\tau}(z = 2.7) \simeq 0.3$ in accordance with Meiksin, Bryan & Machacek (2001). Additionally, to replicate the resolution effects of the Keck HIRES sample, the spectra were convolved with a Gaussian with a full width at half maximum of 8 km s$^{-1}$. Finally the spectra were rebinned to a pixel size of 2.1 km s$^{-1}$.

The main complication arises since the analysis of the correlation between wave numbers must be performed using the 3 dimensional flux power spectrum. Following the technique of Croft *et al.* a quantity $P_{\rm F}(k)$ is defined as

$$P_{\rm F}(k) \equiv -\frac{2\pi}{k} \frac{d}{dk} P_{\rm 1D,F}(k). \tag{4.10}$$

In practice performing this differentiation is non-trivial. Typically the 1 dimensional flux power spectrum is a noisy function, which in all probability will not be monotonically decreasing. This leads to the possibility of calculating a 3 dimensional power which is negative. Reducing the noise inherent in the measured 1 dimensional flux power spectrum is therefore of key importance when this inversion is applied.

Evaluating the necessary gradient at a given point can be done in a number of different ways. The technique applied by Croft *et al.* is to estimate the derivative at the point $(x_i, y_i)$ by averaging the gradients between this point and its two neighbouring points, $(x_{i-1}, y_{i-1})$ and $(x_{i+1}, y_{i+1})$ (Croft, private communication). This can result in a mean positive gradient and so the magnitude of the power spectrum quoted for a given wavenumber is calculated by averaging the results about this value in $k$ bins of fixed logarithmic length. However even after this process is applied a 3 dimensional flux power spectrum which is negative on some scales is often calculated when individual spectra are analysed, especially at low wave numbers where only a small number of modes are averaged.

Although the occurrence of negative power is unphysical, it need not effect the correlation analysis. Bearing in mind that what is calculated is an estimator of the true flux power spectrum it is unimportant that this is sometimes less than zero, provided that the mean for a large sample is positive. If this is not the case then using this estimator cannot be justified.

The correlation coefficients were calculated for wave numbers which were chosen to match the values used in Croft *et al.* (2000). Table 4.1 gives the values of $k$ corresponding to each index, $i$, of the correlation matrix presented.

Figure 4.1 shows the correlation matrices calculated for 50 realizations, and by Croft *et al.* from their data set using the jackknife estimator. The top plot also shows the area of circles that would correspond to scatter of $1\sigma$, $2\sigma$ and $3\sigma$ in uncorrelated data, for 50 realizations. There are considerable differences between these two matrices, firstly the positive correlations they find at high wave numbers is absent. The authors note that if the correlation seen is statistically significant it is probably due to the differencing

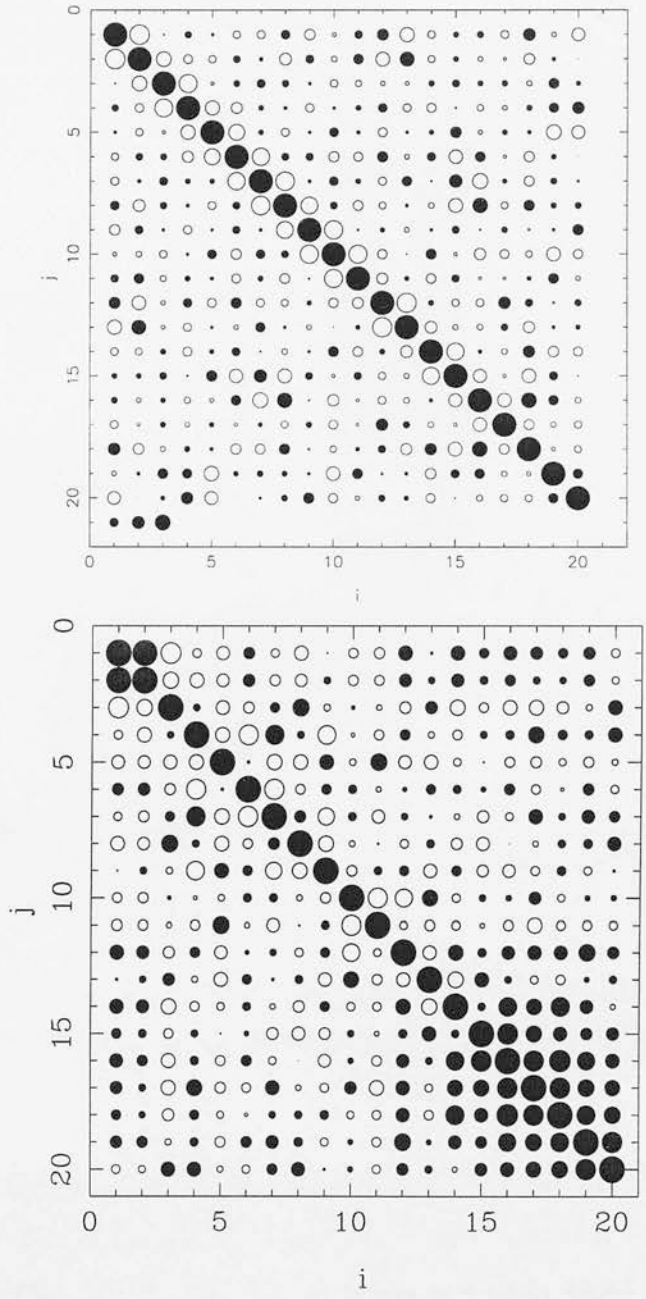


**Figure 4.1:** The correlation matrices calculated for a sample of simulated spectra (top plot) and for the fiducial Keck HIRES data of Croft *et al.* using the jackknife technique (bottom plot taken from their paper). In each case the amplitude of the coefficient is proportional to the area of the circle. The filled circles represent positive coefficients, while negative coefficients are represented with hollow circles. The bottom row in the top plot shows the size of expected $1\sigma$, $2\sigma$ and $3\sigma$ variations arising from 50 realizations of uncorrelated data. The wave scales corresponding to each integer value can be found in table 4.1.

| Number | $k$ $(\mathrm{km\,s^{-1}})^{-1}$ |
|---|---|
| 1 | $1.99 \times 10^{-3}$ |
| 2 | $2.59 \times 10^{-3}$ |
| 3 | $3.37 \times 10^{-3}$ |
| 4 | $4.37 \times 10^{-3}$ |
| 5 | $5.68 \times 10^{-3}$ |
| 6 | $7.38 \times 10^{-3}$ |
| 7 | $9.58 \times 10^{-3}$ |
| 8 | $1.24 \times 10^{-2}$ |
| 9 | $1.62 \times 10^{-2}$ |
| 10 | $2.10 \times 10^{-2}$ |
| 11 | $2.72 \times 10^{-2}$ |
| 12 | $3.55 \times 10^{-2}$ |
| 13 | $4.61 \times 10^{-2}$ |
| 14 | $5.98 \times 10^{-2}$ |
| 15 | $7.77 \times 10^{-2}$ |
| 16 | 0.101 |
| 17 | 0.131 |
| 18 | 0.170 |
| 19 | 0.221 |
| 20 | 0.287 |

**Table 4.1:** The wave scales for the correlation matrices calculated by Croft *et al.* and for a sample of simulated spectra generated using the lognormal approximation.

needed to infer the flux power spectrum. The second major difference seen is that the calculation on the simulations shows anti-correlation between neighbouring wave numbers.

A possible explanation for this may arise from the way the flux power spectrum is estimated. If the flux power spectrum found at a particular wavenumber is very large for a given realization, this corresponds to region in $k$ space where the 1 dimensional flux power spectrum has shown a sharper fall off than normal. This means the 1 dimensional flux power spectrum is at a lower value on some given scale than would be expected. Therefore, since this quantity is always positive, the gradient in the higher $k$ regimes is likely to be lower. This will result in anti-correlations between neighbouring $k$ being likely.

This is demonstrated in figure 4.2 for the simple case where the 1 dimensional flux power spectrum is a well defined exponential. The solid line represents gradient

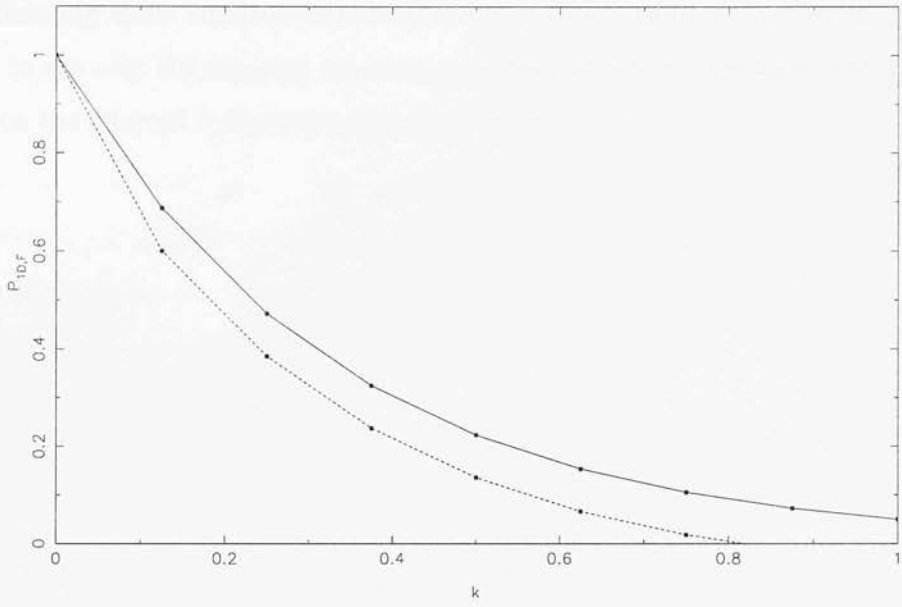


**Figure 4.2:** A demonstration that a high value of the inferred 3 dimensional flux power spectrum in one $k$ bin cannot occur without affecting measurements on other scales.

estimates if a realization exactly followed this form. The dashed line shows the case if one tries to enforce a high power spectrum measurement in one isolated regime of $k$ ($k = 0.1$) while maintaining the correct values of inferred 3 dimensional power for all other values of $k$. The requirement that the gradients match in all regions, except $k = 0.0 - 0.1$, means the 1 dimensional flux power spectrum must fall below zero. Therefore this scenario cannot happen and the high measurement in one $k$ bin requires a low measurement in one or more of the other bins. This argument assumes that the 1 dimensional flux power spectrum always has a similar value at very low $k$ values. Although this should be true statistically it is quite reasonable to expect to find isolated cases where high values of the flux power spectrum are calculated.

Outside of neighbouring $k$ bins, and the positive correlation found by Croft *et al.* in the high $k$ regime both the matrices presented have off diagonal elements which are consistent with zero, although the scatter in the Croft *et al.* results is more pronounced than for the correlations computed directly from the simulations. As previously discussed some scatter around zero is inevitable and these elements may not be statistically

significant. However even if this is the case then the significant correlations between neighbouring wave numbers implies the assumption made by Croft *et al.*, that it is valid to use only the diagonal elements, may lead to errors when calculating the error bars on the inferred 3 dimensional mass power spectrum.

## 4.2 Calculating the 3 dimensional flux power spectrum

As previously mentioned a number of techniques can be used to infer the 3 dimensional flux power spectrum from its 1 dimensional counterpart. This is intimately linked with the question of how large a data sample needs to be before it converges reliably. In this section an alternative to the direct calculation of the gradient is considered. With this technique the derivative is found using a spline fit of the 1 dimensional flux power spectrum. This smooths the data in its original form before the derivative is found.

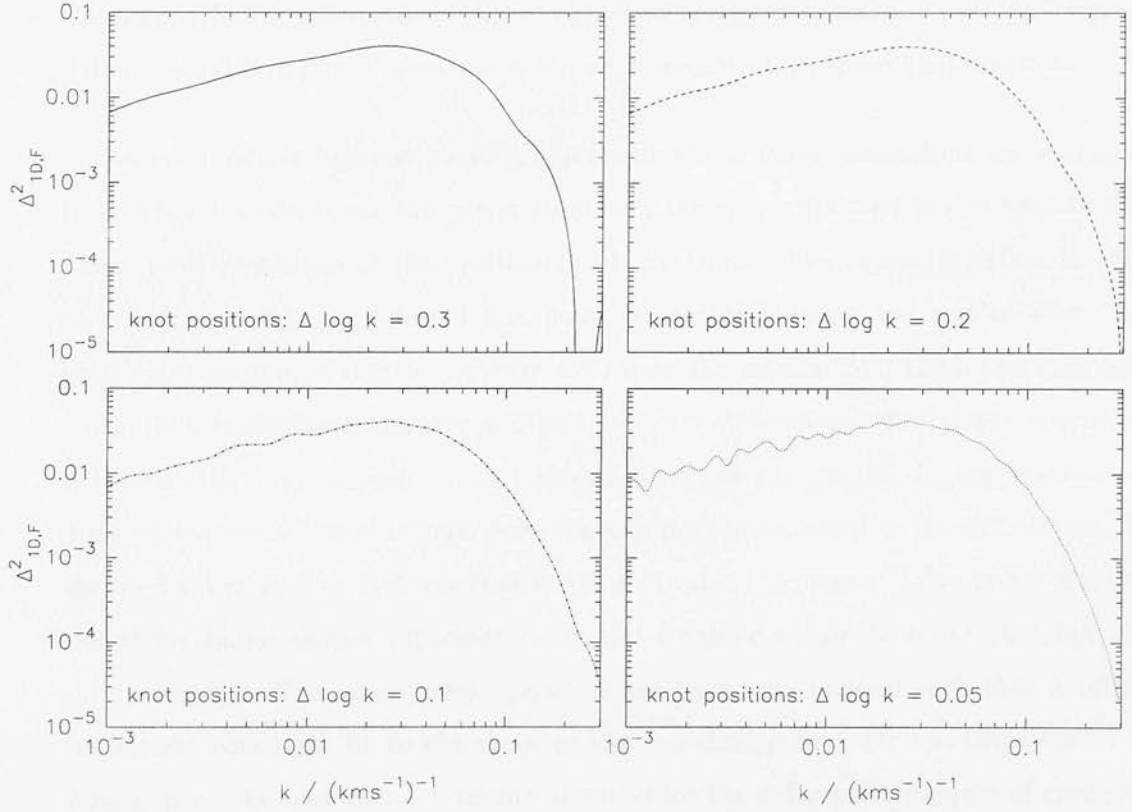


**Figure 4.3:** The effect of knot positions on the spline fit. If the knots are spaced in close proximity then an oscillating function can result, shown in the bottom right hand plot. At the other extreme too few knots also result in an inaccurate fit as seen in the high $k$ regime of the top left hand plot.

Using this method raises the crucial question of how the data is best fitted. The key when using this spline fit method is to attempt to accurately represent the data points

whilst still offering a function which is suitably smooth. It is important to balance these two factors carefully and this is controlled by the positioning of internal knots in the data set.

A routine is used that fits the data with a function which is usually continuous in its first three derivatives. At certain points, the knot positions, the third derivative is allowed to be discontinuous. The positioning of these knots is crucial in determining the final form of the function. This effect is shown in figure 4.3, the same 1 dimensional flux power spectrum (averaged from 50 realizations) is fitted with four different sets of knot positions, chosen such that they are separated by intervals of constant logarithmic $k$. This figure clearly demonstrates that the gradient at a given point, and thus the inferred 3 dimensional flux power spectrum is strongly dependent on these knot positions.

As the function becomes smoother, for example as more realizations are averaged to find the 1 dimensional flux power spectrum, the spline fits tend to converge to the same result regardless of the positioning of the knots. The crucial question in this case is how closely the inferred flux power spectrum matches the results when the gradient is computed directly. Figure 4.4 shows the results from these two methods when 1000 realizations are averaged and so the 1 dimensional flux power spectrum is intrinsically very smooth. Also included, to guide the eye, is the renormalised 3 dimensional linear baryonic mass power spectrum corresponding to the simulations. If the mechanism applied by Croft *et al.* is well motivated the mass and flux power spectra should be similar shapes, especially in the low $k$ regime where the bias factor, $b(k)$, is slowly varying. The mass power spectrum has been renormalised such that it offers the closest consistent fit to the mean of the two data points, for the three lowest $k$ values. It can be seen that the results obtained for the different techniques of gradient calculation are only very well matched on certain scales.

This is likely to be due to a weakness in the spline fitting technique. Smoothing the data in this way inevitably leads to information being discarded. This result does however highlight the sensitivity of the inferred 3 dimensional flux power spectrum on the calculated gradient.

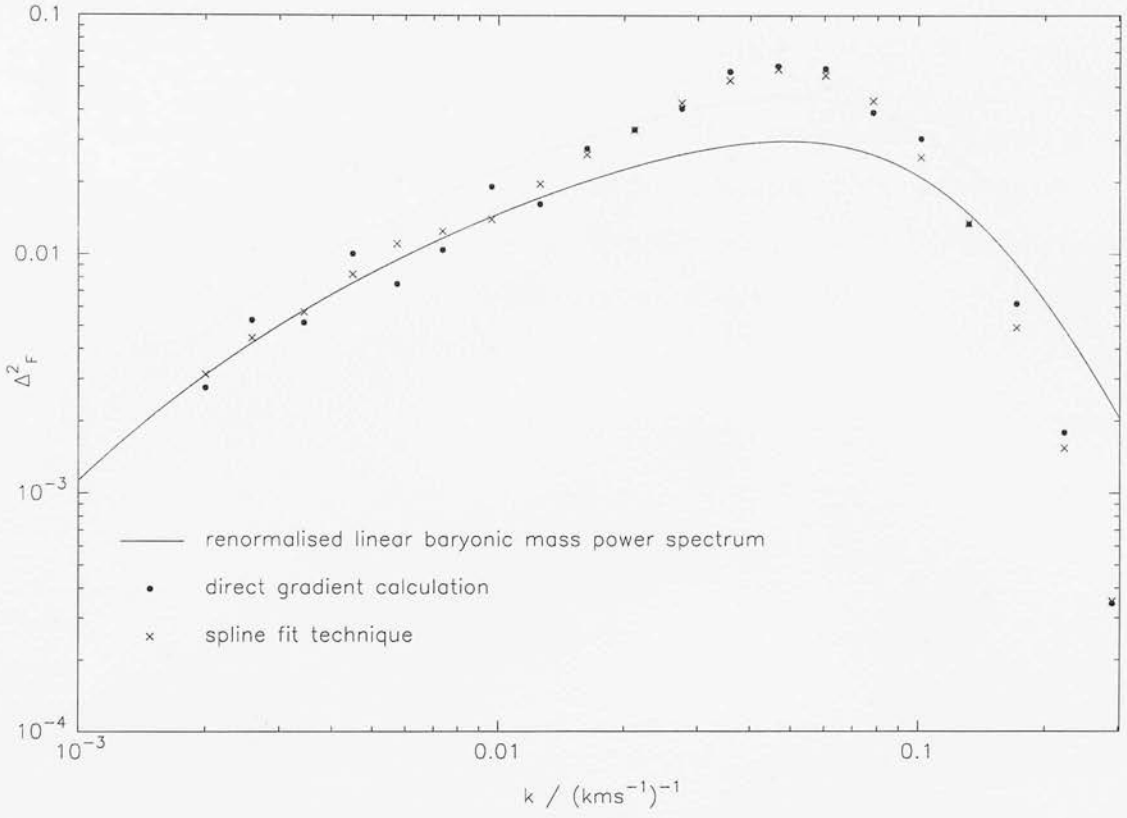


**Figure 4.4:** Comparison of the 3 dimensional flux power spectrum where the calculation is performed for two different techniques. The inferred flux power spectrum can be seen to highly dependent on the method used to calculate the gradient, especially at large scales.

In the case when a more conservative number of realizations are used to calculate the mean 1 dimensional flux power spectrum, the discrepancy between the inferred 3 dimensional results for these two methods is even more pronounced, and more significantly both methods show considerable deviation from the results found when a large number of realizations are used. This is shown in figure 4.5 which shows the inferred 3 dimensional flux power spectrum for a sample of 50 realizations. The spline fit performed had knot positions separated by an interval in logarithmic $k$ of 0.2, since it was found that this fit most closely matched the result when many realizations were averaged.

This highlights the problems inherent in differentiating a noisy function accurately, irrespective of the method used. Although the 1 dimensional flux power spectrum

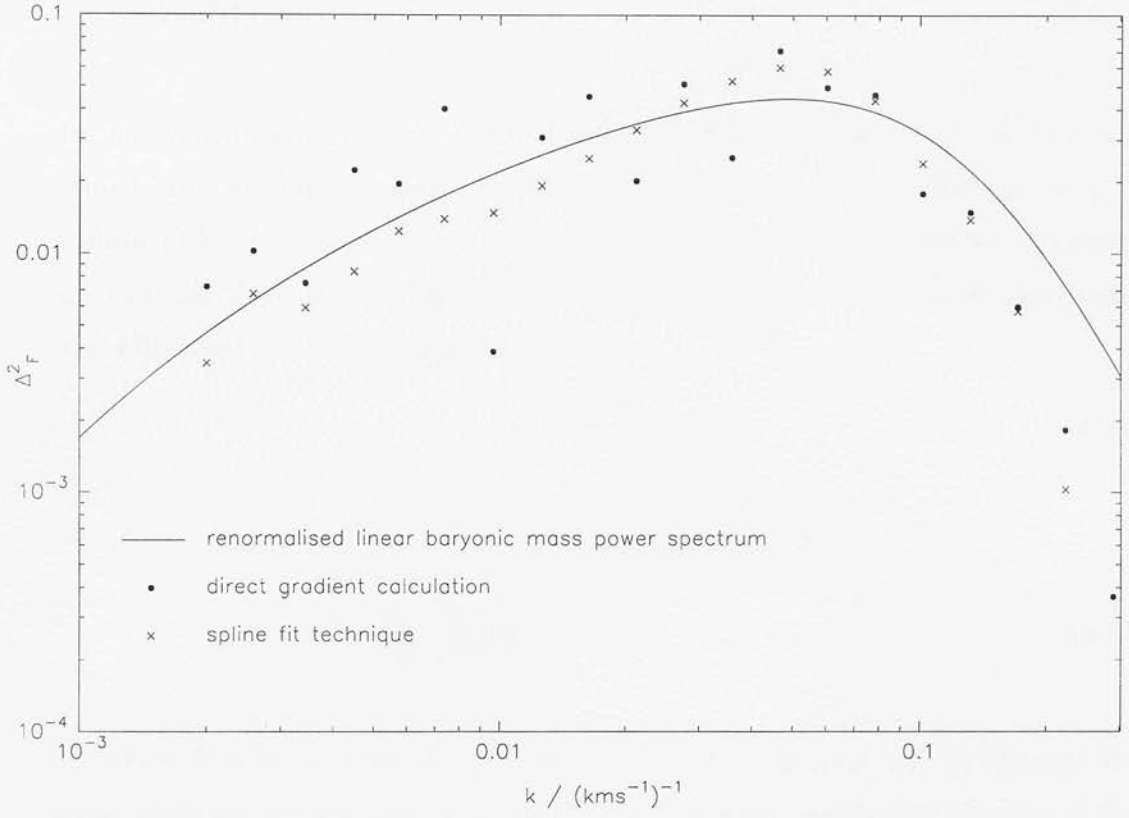


**Figure 4.5:** Comparison of the 3 dimensional flux power spectrum where the calculation is performed on a sample size of 50. The inferred flux power spectrum can be seen to differ considerably from that where a very large number realizations are used (see figure 4.4).

converges well for a sample size of 50 realizations the scatter in each individual point still causes considerable difficulties when differentiation is performed. The uncertainties arising from this step are shown to be considerable, even when a sample size comparable to the largest observational data sets is used.

One complication, that may help reduce the scatter seen, is that if the 1 dimensional flux power spectrum is calculated using the Lomb periodogram technique (as Croft *et al.* do in their analysis) rather than an FFT then this quantity is found for values of $k$ which are not necessarily equally spaced. Therefore the number of modes averaged over to determine each data point at a given value of $k$ will differ from the analysis presented here. If more modes are averaged over then the scatter will be reduced.

## 4.3 Likelihood analysis of the 1 dimensional flux power spectrum

In this section analysis working with the directly observable quantity $P_{\rm 1D,F}(k)$ is presented. This has the advantage that the difficulties associated with inferring the 3 dimensional flux power spectrum from its 1 dimensional counterpart, highlighted earlier, are avoided. To do this the likelihood techniques outlined in §4.1 are used, specifically $\mathcal{L}$ is calculated using the expression

$$\mathcal{L} = \exp(-\chi^2/2), \tag{4.11}$$

where

$$\chi^2 = \sum_N \left[ \sum_{ij} (y_i - y_{i,\rm mod}) C_{ij}^{-1} (y_j - y_{j,\rm mod}) \right], \tag{4.12}$$

where $N$ is the number of simulations contained in the data set. In this case the data points are bin averaged values of $P_{\rm 1D,F}(k)$. In order to keep the inversion of the correlation matrix practical the data is handled in 20 wave bins, similar to those used in the Croft *et al.* analysis but slightly different due to using a different pixel size, the values for these wave bins are shown in table 4.2. The simulations used span a redshift range of $z = 2.5-2.9$, have a pixel size of 2 km s$^{-1}$, and a mean optical depth constraint of $\bar{\tau} \simeq 0.3$. Model values for each wave bin and correlation matrices (see equation 4.9) are calculated using 1000 realizations for each given set of cosmological parameters.

Values of $\chi^2$ are calculated for a given set of data for all 7 cosmological models. Some of these values of $\chi^2$ can be large (especially when $N$ is large), leading to vanishingly small values of $\mathcal{L}$, therefore, for convenience, the quantity considered is

$$R_{\rm m} = \log_{10}\left(\frac{\mathcal{L}_{\rm m}}{\mathcal{L}_{\rm true}}\right) = \log_{10}(e)\frac{\chi^2_{\rm true} - \chi^2_{\rm m}}{2} \qquad {\rm m} = 1, ..., 7. \tag{4.13}$$

So, for example, when this quantity equals $-1$, for a given set of data, these data are 10% as likely to have come from the cosmological model in question than the actual model from which they were generated. However interpretation of the significance of

| Number | $k$ $(\mathrm{km\ s^{-1}})^{-1}$ |
|---|---|
| 1 | $2.00 \times 10^{-3}$ |
| 2 | $2.58 \times 10^{-3}$ |
| 3 | $3.43 \times 10^{-3}$ |
| 4 | $4.48 \times 10^{-3}$ |
| 5 | $5.72 \times 10^{-3}$ |
| 6 | $7.44 \times 10^{-3}$ |
| 7 | $9.73 \times 10^{-3}$ |
| 8 | $1.25 \times 10^{-2}$ |
| 9 | $1.62 \times 10^{-2}$ |
| 10 | $2.12 \times 10^{-2}$ |
| 11 | $2.75 \times 10^{-2}$ |
| 12 | $3.57 \times 10^{-2}$ |
| 13 | $4.65 \times 10^{-2}$ |
| 14 | $6.03 \times 10^{-2}$ |
| 15 | $7.82 \times 10^{-2}$ |
| 16 | 0.102 |
| 17 | 0.132 |
| 18 | 0.171 |
| 19 | 0.222 |
| 20 | 0.288 |

**Table 4.2:** The wave scales used in the likelihood analysis.

specific values of $R_{\mathrm{m}}$ should be done with consideration to the number of realizations in each data set. The reason for this is that this quantity is proportional to the difference in the values of $\chi^2$ between models. For a system with $M$ degrees of freedom one typically expects (e.g. Lupton 1993)

$$\chi^2_{\mathrm{true}} \sim M \pm \sqrt{2M}. \tag{4.14}$$

In the case considered here $M$ is equal to the number of realizations multiplied by the number of wave bins, so $M = 20N$. Thus the $1\sigma$ error bar on the expected value $\chi^2_{\mathrm{true}}$ increases as the square root of the number of realizations. This implies two data sets of 50 realizations ($M = 1000$) generated from the same cosmological model could easily give rise to values of $\chi^2$ which differed by about 50 ($R_{\mathrm{m}} \simeq -10.86$). Therefore low values of $R_{\mathrm{m}}$ do not necessarily imply that a model is confidently ruled out, even though at first glance the ratio of the likelihoods is tiny.

### 4.3.1 Results and Discussion

Figures 4.6 and 4.7 show the quantity $R_{\rm m}$ calculated for a single realization and 50 realizations (the approximate size of the largest data set currently available) respectively. It can be seen that when using a single realization the most likely cosmological model is sometimes found to be a different model than that from which the data was actually generated ($R_{\rm m} > 0.0$). The reason for this is that the individual realizations have a large scatter and so differentiating between different cosmological models can be difficult. Even in the case when the correct cosmological model is identified as the most likely, other models often have similar values of likelihood and ruling out models with any degree confidence is difficult. The exception to this is the Einstein-de Sitter model (cosmological model number 7), any likelihood ratio between this model and any of the ΛCDM cases are vanishingly small. This is not unexpected since the Einstein-de Sitter model is radically different from all the other cosmological models considered.

The results are more conclusive when a greater number of realizations are used, with the true model invariably being identified as the most likely. Even in this case however only some of the comparison models are confidently ruled out. Perhaps the most interesting result arises when the data is generated using model 4, where the technique finds that the data are almost equally likely to have come from model 1 or model 4 (the difference between these two being the value of the spectral index $n$). A bigger likelihood ratio between these two models is seen when model 1 is used to generate the data, but model 4 is not confidently ruled out. This indicates that the technique may not be very powerful when distinguishing between different values of certain parameters. However any constraints that could be imposed would be useful and combining results with those from other sources should help set stringent constraints on cosmological parameters.

Although these results are an encouraging indication for the validity of using likelihood analysis on the 1 dimensional data, the method used here is somewhat rudimentary. Ideally, rather than considering a small number of models, a large parameter space would be studied. Then maximum likelihood regions could be found and probability contours used to show the degree of confidence in each result. Also if this technique

were applied to real data then the model parameters should ideally be taken from full hydro-PM codes. The major problem is that each $N$-body simulation is expensive in computing time and memory, and a huge number of these simulations would be required to span a large parameter space. However, more economical codes, where the gas is assumed to trace the dark matter through simple scaling relations have been shown to be in good, though not perfect, agreement (White & Meiksin 2000) with the more sophisticated codes. Aside from issues of computer efficiency, there is no reason why direct likelihood analysis of the 1 dimensional data should not provide a complementary or alternative method to determining the primordial mass density power spectrum to the inversion techniques currently being advocated in the literature.

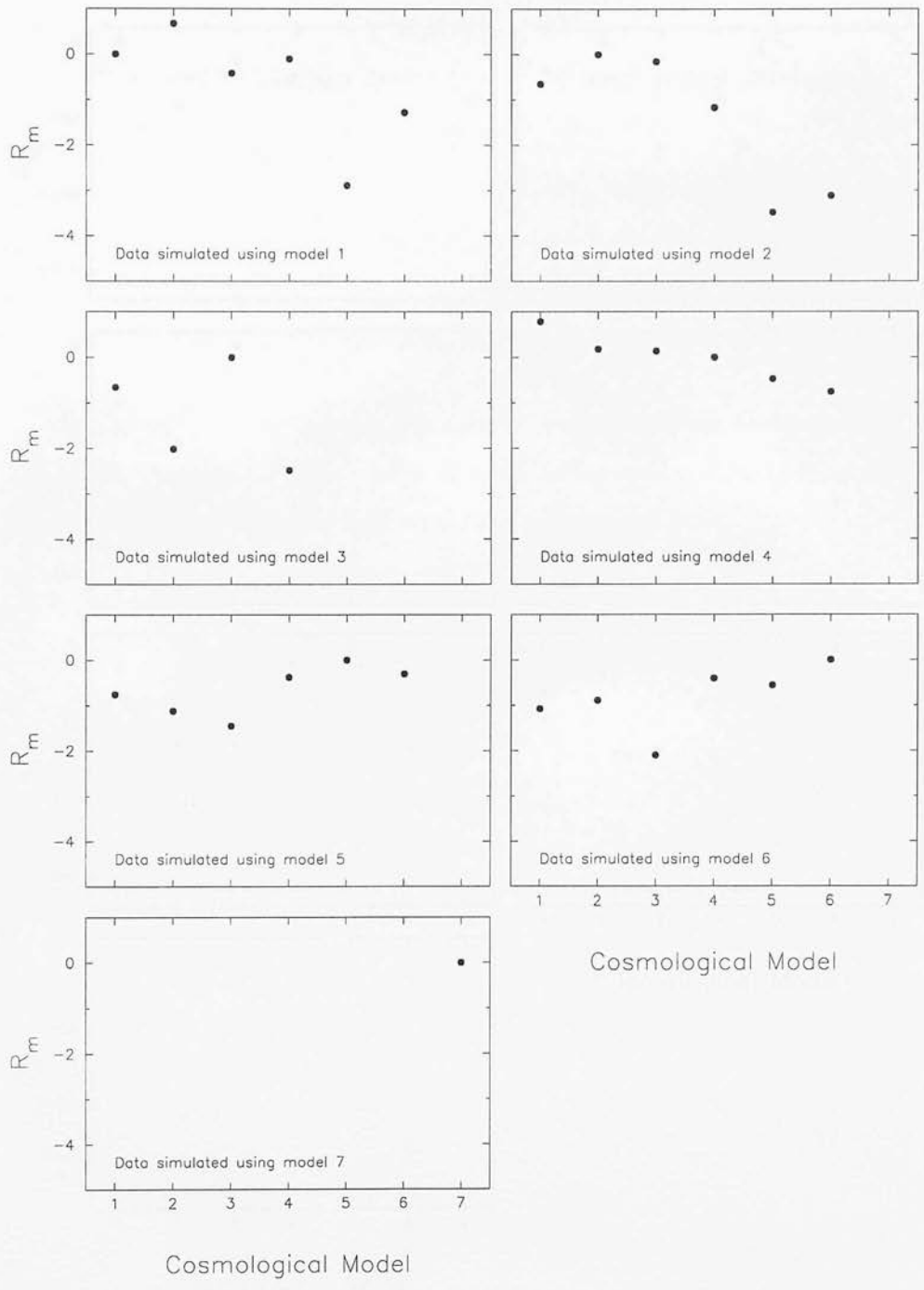


**Figure 4.6:** The logarithm of likelihood ratios for single realizations of each model. The true cosmological model is found to be the most likely in only 5 of the 7 cases, due to the large scatter encountered on individual realizations. Despite this the difference between the Einstein-de Sitter model and ΛCDM cases are sufficiently large that any likelihood ratios involving the Einstein-de Sitter model are vanishingly small.

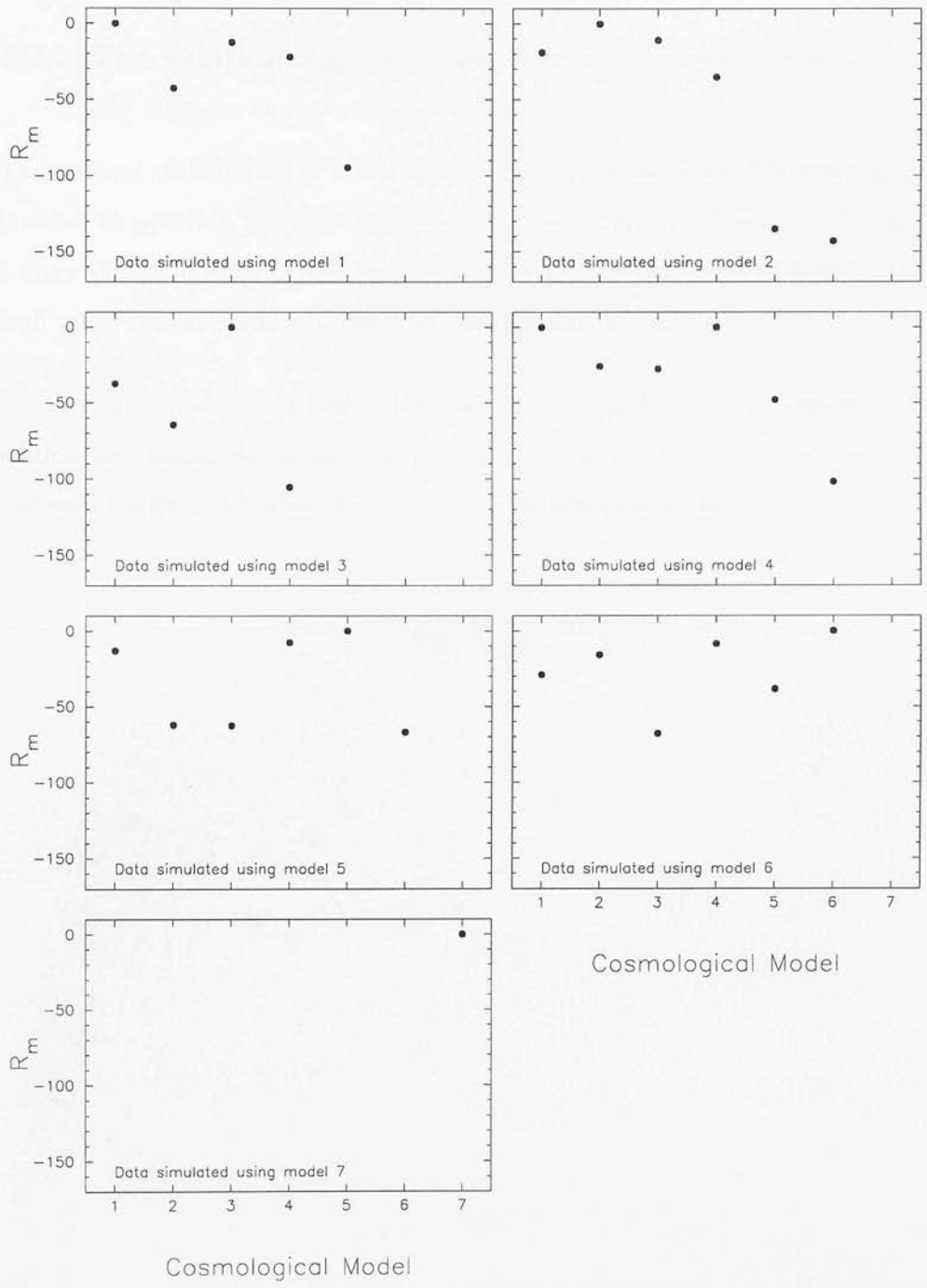


**Figure 4.7:** The logarithm of likelihood ratios for 50 realizations of each model. With sufficient numbers of simulations the true model is always found to be the most likely, although in one case (for data simulated using model 4) only by a negligibly small amount. The likelihood ratios corresponding to incorrect models are also often smaller than for a single realization.

## 4.4 The relationship between flux and mass variances

The unusual distribution of mass variances arising when using the lognormal approximation to generate the mass field was discussed briefly in chapter 2. In this section a more thorough analysis will be presented to analyse the effect that realizations with high mass variances have on the mean flux spectra of a sample.

The key question is whether the spectra generated using the lognormal approximation are reasonable for any value of mass variance. An obvious investigation is to compare the flux and mass variances for individual realizations.

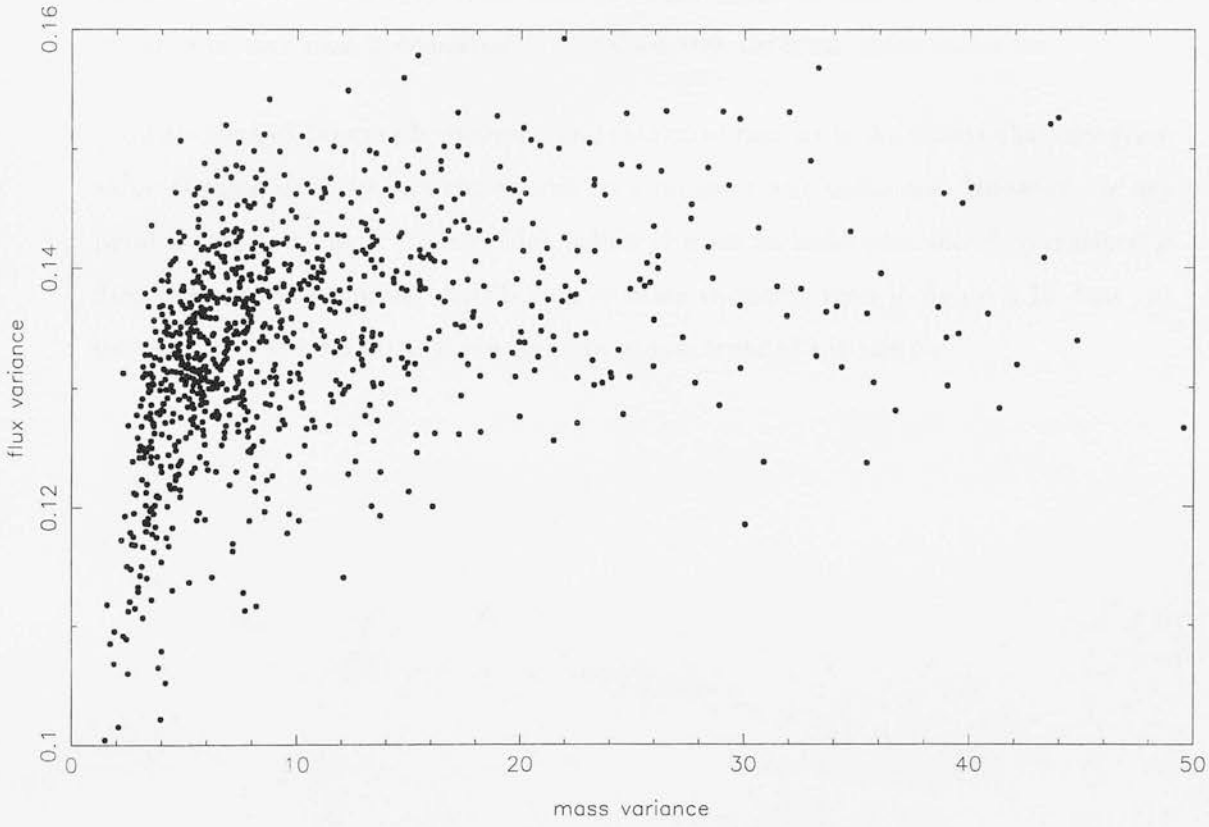


**Figure 4.8:** The relationship between the variance in the mass and flux for a sample of simulated spectra. The result is close to a scatter plot, except at very low mass variances which correspond to low flux variances.

Figure 4.8 shows the flux and mass variances for a sample of 1000 simulated spectra.

The realizations with low flux variances invariably correspond with low mass variances ($\sigma_m^2 < 5.0$). However there is considerable scatter for any given range of mass variances. In the regime of $\sigma_m^2 > 5.0$ the figure is consistent with a scatter diagram with no overall trend seen between the mass and flux variances. The reason for this is the relationship between the flux and the mass at a given point is, neglecting Doppler broadening, given by $f \propto \exp(-\alpha\rho_b^\beta)$. This leads to a situation where the mass levels above some threshold correspond to a saturated region. Therefore high values of variance in the mass field do not transfer to the flux variance as the saturated features reduce the impact of the high mass points.

In this sample studied there are a small number (26) of realizations with mass variances greater than 50.0 which have not been shown on the plot. However the flux variance in each case is consistent with values seen for lower mass variances.

The effect of Doppler broadening and saturated regions is the reason that any given value of mass variance can correspond to a range of flux variances. However the key point is that since even the very high values of mass variance correspond to reasonable flux variances the unusual distribution of mass variances seen in figure 2.12 does not unduly affect the resulting mean flux power spectrum of the sample.

# Chapter 5

# Conclusions

In this chapter I summarise the main results and conclusions resulting from the work presented.

**Monte Carlo Simulations**

In chapter 2 absorption spectra generated from a simple model were analysed. The intrinsic properties of the absorption features were set by various parameters, and cases were studied for both a Poisson distribution and where the line centres were clustered. This analysis has several interesting results which are important with respect to recovery of the mass density power spectrum.

The flux power spectrum is non zero even for a Poisson line distribution, since Ly-$\alpha$ absorption features are not point like objects.

The extent of clustering in an absorption line system is seen to affect the resulting flux power spectrum. This result is required if the techniques used to determine the underlying mass distribution are valid.

The intrinsic statistical properties of individual absorption features is seen to affect the flux power spectrum. Although the change in the flux power spectrum seen from altering the distribution of the Doppler parameters is small, there is a considerable

difference as the column density distribution parameter is altered. In terms of the physical picture, these parameters are dependent on the scaling relations that exist between the dark matter and the baryonic gas.

It is thought that in observed data both the positions and profiles of the absorption features will be determined by the underlying mass field. The differing effects that these lines have on the flux power spectrum may lead to degeneracies between some cosmologies. For example the change to the flux power spectrum resulting from a variation in the large scale mass distribution may be counteracted by a shift in the mass behaviour on small scales, via an alteration in the intrinsic line profiles.

### Realistic Spectra

In this chapter spectra were generated using the lognormal approximation. The absorption features are calculated by summing the optical depth contributions from each pixel along the line of sight. Analysis of these simulations illustrates that inferring a correctly normalized mass power spectrum from the flux power spectrum is non trivial.

Across the range of realistic values of the mean optical depth the shape of the flux power spectrum is not altered significantly. Therefore assuming the inversion method is valid, only the normalization of the mass power spectrum, not its form, will be subject to errors from the uncertainties about the value of this parameter.

The results presented clearly show that the amplitude of the ratio between the flux and mass power spectra is highly dependent on cosmology. Although different cosmological parameters do affect the shape of the flux power spectrum, the effect seen on the linear mass power spectrum is far more pronounced. Therefore, it is concluded that, considerable errors may be present in a linear mass field recovered using the Croft *et al.* inversion method. Rather than calculate a bias parameter from a set of simulations assuming one cosmological model it may be better to investigate the relative likelihood that an observed flux power spectrum can arise from any one of several different cosmologies.

The simulated absorption spectra can be fitted with Voigt profiles which approxi-

mates the properties of the original spectra in both real and Fourier space very well. Therefore deconvolution of spectra into Voigt profiles remains a valuable method of analysis even in the cosmic web paradigm of the Ly-$\alpha$ forest. It is shown that the positions of line centres are very poor indicators of the underlying mass density in realistic cases. This is attributed to the effects of Doppler broadening of the baryonic features, which distorts the relationship between the optical depth at a given point and the mass density at that point.

Further to this the relative importance of the positions and profiles of the absorption features is demonstrated by the line scrambling process. That a different flux power spectrum results after this scrambling process is evidence that the Croft *et al.* method is well motivated. Additionally it is seen that when lines within small column density regimes are scrambled the change of the flux power spectrum is small. Therefore it is concluded that the Doppler broadening of lines is of negligible importance to the accuracy of the Croft *et al.* method.

**The statistics of the flux power spectrum**

In this chapter some of the important statistical properties of the simulated data sets were studied. The correlation coefficients calculated directly from a set of simulations differ considerably from those calculated by Croft *et al.* using the jackknife estimator. The considerable anti-correlation seen between almost all neighbouring wave bins suggests that analysis ignoring the off diagonal elements may not be rigorous enough to accurately determine the errors.

Determining the gradient of a noisy function, such as the 1 dimensional power spectrum, is non-trivial. Two techniques are compared for a large number of realizations and the inferred 3 dimensional flux power spectrum is shown to be dependent on the method of calculating the gradient. Both techniques require a large sample of data for convergence of the inferred 3 dimensional flux power spectrum. For the size of data sets currently available, calculation of the gradient may introduce considerable uncertainties to the analysis.

Determination of the relative likelihoods is shown to successfully identify the un-

derlying cosmological model when large numbers of realizations are used. Although the analysis is fairly basic it is suggestive that working directly with the 1 dimensional data in this way would be an illuminating complementary approach to the techniques already being applied.

### Overall summary

The Ly-$\alpha$ forest currently presents the best opportunity to probe the mass distribution at high redshifts. However the present results must be treated with some caution.

One main weakness of this technique is that the data available is by its nature 1 dimensional. Although the 3 dimensional power spectrum is related to its 1 dimensional counterpart by a known expression, this applies to fields of infinite extent. The noise inherent in any measurements therefore ensures that any conversion is non-trivial.

Since the directly observable quantity is $\Delta^2_{\mathrm{1D,F}}(k)$, I would advocate the development of techniques that study directly how its form is affected by changes to the shape and amplitude of the 3 dimensional mass power spectrum. In this way it should be possible via likelihood analysis and statistical studies to constrain the mass power spectrum without changing the form of the data.

Another concern arises from the lack of knowledge about the precise values of temperature and mean optical depth of the IGM and the reionization history of the Universe. At the current time these uncertainties are potentially a source of substantial errors in the recovered mass power spectrum. However, as the quantity of data increases, parameter uncertainties are reduced, and the analysis techniques are refined, the Ly-$\alpha$ forest should provide strong constraints on both the shape and the amplitude of the mass density power spectrum.

# Appendix A

# Atomic processes

This appendix contains the background theory needed to calculate the properties of atomic transitions. The starting point is to apply quantum mechanics to a single atom, these results can then be extended to the classical regime. In an isolated atom with a set of allowed energy levels, any transition by an electron to a different energy level must be compensated by a shift in energy elsewhere in the system. For example, a transition to a lower energy level is accompanied by the emission of a photon. Similarly a photon of suitable energy could be absorbed to allow a transition to a higher energy level. These effects can be formalised and thus the effect on light of passing through an absorber can be calculated for a given system. These techniques are then applied to the Ly-$\alpha$ transition to obtain parameters used in the main body of this thesis.

## A.1 Atomic transitions

### A.1.1 The Fermi Golden Rule

The easiest way to treat matter in a radiation field is to think of the interaction Hamiltonian as a perturbation. This is possible since the coupling constant (the fine structure constant $\alpha = 1/137$) is small. With this assumption quantum theory can provide a solution, since the unperturbed system of uncoupled atoms and radiation can be fully

described. Applying perturbation theory to an isolated atom in a radiation field it is found that the probability of a transition from an initial to a final state is

$$p_{\rm f}(t) = \frac{|H'_{\rm fi}|^2 f(t,\omega_{\rm fi})}{\hbar^2}, \tag{A.1}$$

where $|H'_{\rm fi}|$ is the matrix element of the Hamiltonian of the transition from the initial to the final state, $\omega_{\rm fi}$ is the angular frequency corresponding to the change in energy $E_f - E_i$ and the function is

$$f(t,\omega) = \frac{\sin^2\left(\frac{\omega t}{2}\right)}{\left(\frac{\omega}{2}\right)^2}. \tag{A.2}$$

Equation A.1 can be made more general, by considering any number of transitions to the final states grouped within some small energy range. The density of states $g(E)$ can be introduced so that $g(E_{\rm f})dE_{\rm f}$ is the total number of final states available. Then the probability of the system being in any one of these states at time $t$ is

$$p_{\rm G}(t) = \frac{1}{\hbar^2}\int_{E_{\rm i}-\Delta E}^{E_{\rm i}+\Delta E} |H'_{\rm fi}|^2 f(t,\omega_{\rm fi}) g(E_{\rm f})\; dE_{\rm f}. \tag{A.3}$$

Now this expression can be simplified if $t$ is large (meaning in this context that $t \gg 2\pi/\omega_{\rm fi}$). In this regime the dominant contribution of the function $f(t,\omega)$ is a peak centred on $\omega = 0$ of height $t^2$ and width $4\pi/t$. Assuming all contributions from the integral that fall outside of the integration limits are negligible and assuming the density of states and Hamiltonian matrix element to be constants in this region, allows the integral to be performed analytically. Finally differentiating the equation with respect to time gives the transition rate. This results in the equation known as the Fermi Golden Rule

$$\frac{dp_{\rm G}(t)}{dt} = \frac{2\pi}{\hbar}\left[|H'_{\rm fi}|^2 g(E_{\rm f})\right]_{E_{\rm f}=E_{\rm i}}. \tag{A.4}$$

So to find the transition rate, the form of the interaction Hamiltonian is required (e.g. Landau & Lifshitz 1982)

$$\hat{H}' = q\int j^\mu A_\mu\; dV, \tag{A.5}$$

where $q$ is electric charge, $j^{\mu}$ is the 4-current and $A_{\mu}$ is the 4-vector potential of the electromagnetic field. The matrix element of $\hat{H}'$ between the initial and final states of the atomic system and radiation field is

$$\langle \psi_{\rm i}|\langle N'_{\mathbf{k},\alpha}|q\int j^{\mu}A_{\mu}\ dV|\psi_{\rm f}\rangle|N_{\mathbf{k},\alpha}\rangle, \tag{A.6}$$

where $\psi$ is the wave function of the electron involved and $N_{\mathbf{k},\alpha}$ is the number of photons in the state with wave vector $\mathbf{k}$ and polarization state $\alpha$. With a suitable choice of frame the zeroth component of the 4-scalar product is zero and if the atom is non-relativistic then $\mathbf{j} = \psi^{*}\mathbf{v}\psi$. The key point now is that the plane wave description of $\mathbf{A}_{\mathbf{k},\alpha}$ can be written in terms of a combination of ladder operators for the radiation field. This is important since to obtain a non-zero matrix element requires an operator that acts on $|N_{\mathbf{k},\alpha}\rangle$ to give a multiple of $|N'_{\mathbf{k},\alpha}\rangle$. This happens if and only if $N' = N \pm 1$. For example $N' = N + 1$ relates to the absorption of a photon and in this case the matrix element becomes

$$\sqrt{\frac{\hbar}{2\epsilon_0\omega V}}\sqrt{N_{\mathbf{k},\alpha}}q\int \psi_{\rm f}^{*}\mathbf{v}\psi_{\rm i}.\mathbf{e}_{\mathbf{k},\alpha}e^{i\mathbf{k}.\mathbf{r}}\ dV. \tag{A.7}$$

The constants and imaginary exponential in the above expression follow from the definition of $\mathbf{A}_{\mathbf{k},\alpha}$ and the constants of proportionality obtained from the ladder operators (e.g. Landau & Lifshitz 1982). The integral in the above equation can now be solved if the dipole approximation is applied, which is justified if the wavelength of the radiation considered is much greater than the atomic dimensions. This is usually the case when the velocity of any particles are much less than the speed of light. The approximation then has the effect of neglecting the photon momentum. In this case $\mathbf{v} = -i\omega\mathbf{r}$ and $e^{i\mathbf{k}.\mathbf{r}} = 1$. Then all that is needed is the total number of allowed final states which is just the number of available atomic states multiplied by the number of allowed final photon states. Treating the radiation field as being plane waves of a large box gives

$$g_{\rm phot} = \frac{V}{(2\pi)^3}\frac{k^2}{\hbar c}\ d\Omega. \tag{A.8}$$

So finally averaging over the dipole direction and assuming both polarization direc-

tions are equally populated gives

$$R_{\text{abs}}(\Omega)d\Omega = \frac{d\Omega}{4\pi}\frac{4\omega^3}{3\hbar c^3}N_{\mathbf{k},\alpha}\frac{e^2}{4\pi\epsilon_0}|\mathbf{r}_{\text{fi}}|^2 g_{\text{u}}, \tag{A.9}$$

where $R_{\text{abs}}$ is the rate of absorption. To calculate the rate for emission the same calculation is performed but with $g_{\text{l}}$ and $N_{\mathbf{k},\alpha}+1$ replacing $g_{\text{u}}$ and $N_{\mathbf{k},\alpha}$

$$R_{\text{em}}(\Omega)d\Omega = \frac{d\Omega}{4\pi}\frac{4\omega^3}{3\hbar c^3}(N_{\mathbf{k},\alpha}+1)\frac{e^2}{4\pi\epsilon_0}|\mathbf{r}_{\text{fi}}|^2 g_{\text{l}}. \tag{A.10}$$

So comparing equations A.9 and A.10 it can be seen that $g_{\text{u}}R_{\text{em}} = g_{\text{l}}R_{\text{abs}} + g_{\text{u}}R_{\text{spon}}$, where $R_{\text{spon}}$ is the rate of spontaneous emission and is given by

$$R_{\text{spon}} = \frac{4\omega^3}{3\hbar c^3}\frac{e^2}{4\pi\epsilon_0}|\mathbf{r}_{\text{fi}}|^2 g_{\text{l}}. \tag{A.11}$$

### A.1.2 Blackbody radiation

It be can seen from equations A.9 and A.10 that

$$\frac{g_{\text{u}}}{g_{\text{l}}}\frac{R_{\text{em}}}{R_{\text{abs}}} = \frac{N_{\mathbf{k},\alpha}+1}{N_{\mathbf{k},\alpha}}. \tag{A.12}$$

If a system is in thermodynamic equilibrium then on average there are as many transitions up as down, therefore if the rate of spontaneous emission is negligible then it follows that $n_{\text{l}}R_{\text{abs}} = n_{\text{u}}R_{\text{em}}$. Where $n_{\text{u}}$ and $n_{\text{l}}$ are the numbers of electrons in the upper and lower atomic states respectively. In addition to this condition the occupancy of the energy levels can be determined by the Boltzmann distribution such that $n_{\text{l}}/g_{\text{l}} = n_{\text{u}}/g_{\text{u}}e^x$, where $x = (E_{\text{u}} - E_{\text{l}})/kT$. Substituting these expressions into equation A.12, gives an expression for $N_{\mathbf{k},\alpha}$, so multiplying by $h\nu$ gives the energy in each radiation state. Finally multiplying by the density of states gives the energy density as

$$U(\nu)d\nu = \frac{8\pi\nu^2}{c^3}\frac{h\nu}{e^x - 1}d\nu. \tag{A.13}$$

Classically it is usual to quantify a radiation field by its specific intensity, $I_\nu$, such that $I_\nu d\nu d\Omega$ is the energy per unit area per unit time of the photons whose frequency is within $d\nu$ about $\nu$. Since radiation is isotropic the specific intensity for a blackbody

$B_\nu(T)$ is just the above expression multiplied by $c/4\pi$. Therefore a blackbody at temperature $T$ has specific intensity which is given by

$$B_\nu(T) = \frac{2\nu^2}{c^2}\frac{h\nu}{e^x - 1}. \tag{A.14}$$

### A.1.3 Einstein coefficients

All that has gone before can now be related to the Einstein coefficients. The coefficient $A_{\rm ul}$ is the probability that a transition from the upper to lower energy will spontaneously occur, and so by definition is equal to $R_{\rm spon}$. Clearly there is no such coefficient for the reverse jump, as the energy level can never rise spontaneously. The two other coefficients are $B_{\rm ul}$ and $B_{\rm lu}$ which when multiplied by the specific intensity of the surrounding radiation field give the rates of transitions which are stimulated in an atom by this field. To find how these coefficients relate it is easiest to consider a system in thermodynamic equilibrium, which must fulfill the condition

$$n_{\rm l} B_{\rm lu} I_\nu = n_{\rm u}(B_{\rm ul} I_\nu + A_{\rm ul}). \tag{A.15}$$

Or re-arranging and using the Boltzmann condition for the occupancy levels

$$I_\nu = \frac{\frac{A_{\rm ul}}{B_{\rm ul}}}{\left(\frac{g_{\rm l}}{g_{\rm u}}\frac{B_{\rm lu}}{B_{\rm ul}}e^x - 1\right)}, \tag{A.16}$$

and now applying the condition that $I_\nu = B_\nu(T)$ as given in equation A.14 it can be seen immediately that the Einstein coefficients must obey the following rules

$$g_{\rm l} B_{\rm lu} = g_{\rm u} B_{\rm ul}, \tag{A.17}$$

$$A_{\rm ul} = \frac{2h\nu^3}{c^2} B_{\rm ul} = \frac{2h\nu^3}{c^2}\frac{g_{\rm l}}{g_{\rm u}} B_{\rm lu}. \tag{A.18}$$

Rather than referring directly to the Einstein $B_{\rm lu}$ coefficient it is often easier to think in terms of the cross section per particle $\sigma$. This can be thought of in classical terms such that a large cross section corresponds to a high probability that an interaction will occur

$$\sigma = \int \sigma_\nu \, d\nu = \frac{h\nu}{4\pi} B_{\rm lu}. \tag{A.19}$$

The results obtained so far can now be applied to the Ly-$\alpha$ transition. It is shown in equation A.11 that to calculate $A_{ul}$ requires an expression for the matrix element $|\mathbf{r}_{fi}|^2$. This can be found for the hydrogen atom since it is an exactly soluble system whose wave function is $\psi_{n\ell m} = R_{n\ell}(r)Y_{\ell m}(\theta,\phi)$. In this equation $n$, $\ell$ and $m$ represent the energy level, angular moment and $z$-component of the angular momentum respectively. $R_{n\ell}(r)$ are the radial functions and $Y_{\ell m}(\theta,\phi)$ are the spherical harmonic functions, both of which are known (e.g. Rae 1992). This system has energy levels given by

$$E_n = -\frac{m_e}{2\hbar^2}\left(\frac{e^2}{4\pi\epsilon_0}\right)^2 \frac{1}{n^2}. \tag{A.20}$$

The upper state in the Ly-$\alpha$ transition has values $n = 2$, $\ell = 1$ and $m = -1, 0, 1$ and the ground state is $n = 1$ and $\ell = m = 0$. Therefore the corresponding energy change $(E_2 - E_1)$ is 10.2 eV and the wavelength of radiation involved is 1216 Å. Now writing the initial wave function of the electron explicitly gives

$$\begin{aligned} |\psi_{21m}\rangle &= R_{2,1}(r)Y_{1,m}(\theta,\phi) \\ &= \frac{r}{\sqrt{24a_o^5}} e^{-\frac{r}{2a_o}}\, Y_{1,m}(\theta,\phi), \end{aligned} \tag{A.21}$$

where

$$a_0 = \frac{(4\pi\epsilon_0)\hbar^2}{m_e e^2} \simeq 0.53 \text{ Å}, \tag{A.22}$$

and the corresponding spherical harmonics are

$$Y_{1,\pm1} = \mp\sqrt{\frac{3}{8\pi}}\sin\theta\; e^{\pm i\phi}, \qquad Y_{1,0} = \sqrt{\frac{3}{4\pi}}\cos\theta.$$

Now the final state of the wave function of the election is given by

$$\begin{aligned} |\psi_{100}\rangle &= R_{1,0}(r)Y_{0,0}(\theta,\phi) \\ &= \frac{2}{\sqrt{a_0^3}} e^{-\frac{r}{a_o}}\; \frac{1}{\sqrt{4\pi}}. \end{aligned} \tag{A.23}$$

Now the operator $\hat{r} = \mathbf{r} = x\mathbf{e_x} + y\mathbf{e_y} + z\mathbf{e_z}$, can be expressed in polar coordinates as

$$x = r\sin\theta\cos\phi = r\sqrt{\tfrac{2\pi}{3}}\ (Y^*_{1,-1} - Y^*_{1,1}).$$

$$y = r\sin\theta\sin\phi = -ir\sqrt{\tfrac{2\pi}{3}}\ (Y^*_{1,-1} + Y^*_{1,1}).$$

$$z = r\cos\theta = r\sqrt{\tfrac{4\pi}{3}}\ Y^*_{1,0}.$$

So for example the $x$ component of the integral $|\mathbf{r}_{\mathrm{fi}}|$ is

$$\langle\psi_{100}|x|\psi_{21m}\rangle = \int\int r^3 R^*_{1,0}(r)R_{2,1}(r)Y^*_{0,0}(\theta,\phi)\frac{x}{r}Y_{1,m}(\theta,\phi)\ dr\ d\Omega,$$

which can be split into radial and angular integrals such that

$$\langle\psi_{100}|x|\psi_{21m}\rangle = I\sqrt{\frac{2\pi}{3}}\int Y^*_{0,0}(Y^*_{1,-1} - Y^*_{1,1})Y_{1,m}\ d\Omega$$

$$= I\sqrt{\frac{2\pi}{3}}\frac{1}{\sqrt{4\pi}}\int Y^*_{1,-1}Y_{1,m} - Y^*_{1,1}Y_{1,m}\ d\Omega.$$

Now invoking the orthogonality relations of the spherical harmonics

$$\int Y^*_{\ell',m'}Y_{\ell,m}\ d\Omega = \delta^D_{\ell\ell'}\delta^D_{mm'}, \tag{A.24}$$

gives

$$\langle\psi_{100}|x|\psi_{21m}\rangle = \frac{I}{\sqrt{6}}(\delta^D_{-1,m} - \delta^D_{1,m}), \tag{A.25}$$

where $I$ is the radial integral. Similar calculations now give

$$\langle\psi_{100}|y|\psi_{21m}\rangle = \frac{-iI}{\sqrt{6}}(\delta^D_{-1,m} + \delta^D_{1,m}), \tag{A.26}$$

$$\langle\psi_{100}|z|\psi_{21m}\rangle = \frac{I}{\sqrt{3}}\delta^D_{0,m}, \tag{A.27}$$

and so squaring all the contributions and adding gives a result which is independent of $m$

$$|\langle\psi_{100}|\mathbf{r}|\psi_{21m}\rangle|^2 = \frac{I^2}{3}. \tag{A.28}$$

The radial integral can then be calculated

$$I = \int_0^\infty r^3 R_{1,0}^* R_{2,1}\, dr$$

$$= \int_0^\infty r^3 \frac{2}{\sqrt{a_o^3}} e^{-\frac{r}{a_o}} \frac{r}{\sqrt{24 a_o^5}} e^{-\frac{r}{2a_o}}\, dr$$

$$= \frac{1}{a_o^4\sqrt{6}} \int_0^\infty r^4 e^{-\frac{3r}{2a_o}}\, dr$$

$$= \frac{256 a_o}{81\sqrt{6}}, \tag{A.29}$$

which upon substitution into A.28 gives

$$|\langle \psi_{100} | \mathbf{r} | \psi_{21m} \rangle|^2 = 0.555 a_o^2. \tag{A.30}$$

Thus the Einstein coefficient is given by

$$A_{2,1\to 1,0} = \frac{4\omega^3}{3\hbar c^3} \frac{e^2}{4\pi\epsilon_o} |\langle \psi_{100} | \mathbf{r} | \psi_{21m} \rangle|^2 = 6.3 \times 10^8 \text{ s}^{-1}. \tag{A.31}$$

This in turn can be substituted into equations A.18 and A.19 to give the cross section for the Ly-$\alpha$ transition

$$\sigma = \frac{h\nu}{4\pi} \frac{g_u}{g_l} \frac{c^2}{2h\nu^3} A_{ul} = 1.11 \times 10^{-2} \text{ cm}^2\text{s}^{-1}. \tag{A.32}$$

# Appendix B

# Properties of Gaussian statistics

This appendix contains the proof that for a probability distribution which is Gaussian, with a mean of zero, $\langle e^x \rangle$ is equal to $e^{\frac{\langle x^2 \rangle}{2}}$.

For a Gaussian with a mean of zero and a standard deviation of $\sigma$ the probability distribution function is given by

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}}. \tag{B.1}$$

The mean of $e^x$ is given by

$$\langle e^x \rangle = \int_{-\infty}^{\infty} p(x) e^x \, dx. \tag{B.2}$$

Substituting in equation (B.1) gives

$$\langle e^x \rangle = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2\sigma^2}+x} \, dx. \tag{B.3}$$

Completing the square of the term in the exponential gives

$$-\frac{x^2}{2\sigma^2} + x = -\left(\frac{x}{\sqrt{2}\sigma} - \frac{\sigma}{\sqrt{2}}\right)^2 + \frac{\sigma^2}{2}. \tag{B.4}$$

Thus making the substitution

$$u = \frac{x}{\sqrt{2}\sigma} - \frac{\sigma}{\sqrt{2}},$$

$$du = \frac{dx}{\sqrt{2}\sigma}, \tag{B.5}$$

gives

$$\langle e^x \rangle = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{\sigma^2}{2}} \int_{-\infty}^{\infty} e^{-u^2} \sqrt{2}\sigma \; du \tag{B.6}$$

$$= \frac{1}{\sqrt{\pi}} e^{\frac{\sigma^2}{2}} \int_{-\infty}^{\infty} e^{-u^2} \; du \tag{B.7}$$

$$= e^{\frac{\sigma^2}{2}}. \tag{B.8}$$

Now comparing this result with the one for $\langle x^2 \rangle$,

$$\langle x^2 \rangle = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} x^2 e^{-\frac{x^2}{2\sigma^2}} \; dx. \tag{B.9}$$

Noting the function being integrated is even allows a change on the integrating limits such that

$$\langle x^2 \rangle = \sqrt{\frac{2}{\pi}} \frac{1}{\sigma} \int_{0}^{\infty} x^2 e^{-\frac{x^2}{2\sigma^2}} \; dx. \tag{B.10}$$

Now using the standard result for the integral which is (e.g. Gradshteyn & Ryzhik 1994)

$$\int_0^{\infty} x^{2n} e^{-ax^2} \; dx = \frac{1.3....(2n-1)}{2^{n+1} a^n} \sqrt{\frac{\pi}{a}}. \tag{B.11}$$

So setting $n = 1$ and $a = 1/2\sigma^2$ gives

$$\langle x^2 \rangle = \sqrt{\frac{2}{\pi}} \frac{1}{\sigma} \frac{2\sigma^2}{4} \sqrt{2\sigma^2 \pi} \tag{B.12}$$

$$= \sigma^2. \tag{B.13}$$

Thus proving the result that for a Gaussian distribution with $\bar{x} = 0$

$$\langle e^x \rangle = e^{\frac{\langle x^2 \rangle}{2}}. \tag{B.14}$$

# Appendix C

# References


Abel, T., Anninos, P., Zhang, Y., Norman, M.L., 1997, New Astron, **2**, 181

Atwood, B., Baldwin, R.A., Carswell, R.F., 1985, ApJ, **292**, 58

Balbi, A., 2000, ApJ, **545**, L1

Bardeen, J.M., Bond, J.R., Kaiser, N., Szalay, A.S., 1986, ApJ, **304**, 15

Bi, H., 1993, ApJ, **405**, 479

Bi, H., Börner, G., Chu, Y., 1992, A&A, **226**, 1B

Bi, H., Davidsen, A.F., 1997, ApJ, **479**, 523

Bi, H., Ge, J., Fang, L.Z., 1995, ApJ, **452**, 90

Berestetskii, V.B., Lifshitz, E.M., Pitaevskii, L.P., 1982, *Quantum Electrodynamics, Landau & Lifshitz, Course for Theoretical Physics*, Volume 4, 2nd Edition, Pergonion Press

Bond, J.R., Kofman, L., Pogosyan, D., 1996, Nature, **380**, 603

Carroll, S.M., Press, W.H., Turner, E.L., 1992, ARA&A, **30**, 499

Carswell, R.F., Morton, D.C., Smith, M.G., Stockton, A.N., Turnshek, D.A., Weymann, R.J., 1984, ApJ, **278**, 486

Carswell, R.F. Lanzetta, K.M., Parnell, H.C., Webb, J.K., 1991, ApJ, **371** 36

Cen, R., Miralda-Escudé, J., Ostriker, J.P., Rauch, M., 1994, ApJ, **437**, L9

Coles, P., Jones, B., 1991, MNRAS, **248**, 1

Colombi, S., 2000, http://www.cfht.hawaii.edu/News/Lensing/#IC

Cooke, A.J., Espey, B., Carswell, R.F., 1997, MNRAS, **284**, 552

Croft, R.A.C., Weinberg, D.H., Katz, N., Hernquist L., 1998, ApJ, **495**, 44

Croft, R.A.C., Weinberg, D.H., Pettini, M., Hernquist L., 1998, Katz, N., 1999, ApJ, **520**, 1

Croft, R.A.C., Weinberg, D.H., Bolte, M., Burles, S., Hernquist, L., Katz, N., Kirkman, D., Tytler, 2001, astro-ph/0012324

Davé, R., Hernquist, L., Weinberg, D.H., Katz, N., 1997, ApJ, **477**, 21

de Bernardis, P. *et al.*, 2000, Nature, **404**, 955

Fang, L.Z., Bi, H., Xiang, S., Börner, G., 1993, ApJ, **413**, 477

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

Feng, L.L., Fang, L.Z., 2000, ApJ, **535**, 519

Fischler, W., Ratra, B., Susskind, L., 1985, *nuc. phys B*, **259**, 730

Freedman, W. L., *et al.*, 2001, ApJ, **553**, 47

Giallongo, E., Cristiani, S., Fontana, A., Trevese, D., 1993, ApJ, **416**, 137

Gnedin, Y.G., Hamilton, A.J.S., 2002, MNRAS, **334**, 107

Gnedin, Y.G., Hui, L., 1996, ApJ, **472**, L73

Gradshteyn, I.S., Ryzhik, I.M., 1994, *Table of Integrals, Series, and Products*, Academic Press Ltd, Editor Jeffery, A.

Gunn, J.E., Peterson, B.A., 1965, ApJ, **142**, 1633

Guth, A.H., 1981, Phys.Rev.D, **23**, 347

Hamilton, A.J.S., 1993, ApJ, **417**, 19

Hamilton, A.J.S., Kumar, P., Lu, E., Matthews, A., 1991, ApJ, **374**, L1

Heath, D.J., 1977, MNRAS, **179**, 351

Hernquist, L., Katz, N., Weinberg, D., Miralda-Escudé, J., 1996, ApJ, **457**, L51

Hu, E.M., Kim, T-S., Cowie, L.L., Songaila, A., Rauch, M., 1995, AJ, **110**, 1526

Hui, L., 1999, ApJ, **516**, 519

Hui, L., Gnedin, N., 1997, MNRAS, **292**, 27

Hui, L., Gnedin, N., Zhang, Y., 1997, ApJ, **486**, 599

Impey, C.D., Petry, C.E., Malkan, M.A., Webb, W., 1996, ApJ, **463**, 473

Jamkhedkar, P., Hongguang, B., Fang, L.Z., 2001, ApJ, **561**, 94

Jha, S., 2001, *In New Cosmological Data and the Values of the Fundamental Parameters*, eds Lasenby, A., Wilkinson, A. astro-ph/0101521

Kaiser, N., Peacock, J.A., 1991, ApJ, **379**, 482

Kirkman, D., Tytler, D., 1997, ApJ, **484**, 672

Lahav, O., *et al.*, 2002, MNRAS, **333**, 961

Lumsden, S.L., Heavens, A.F., Peacock, J.A., 1989, MNRAS, **238**, 293

Lupton, R., 1993, *Statistics in Theory and Practice*, Princeton University Press

Lynds, C.R., 1971, ApJ, **164**, L73

Ma, C.P., 1996, ApJ, **471**, 13

McDonald, P., Miralda-Escudé, J., Rauch, M., Sargent, W.L.W., Barlow, T.A., Cen, R., Ostriker, J.P., 2000, ApJ, **543**, 1

McDonald, P., 2001, astro-ph/0108064

McGill, C., 1990, MNRAS, **242**, 544

Meiksin, A., Madau, P., 1993, ApJ, **412**, 34

Meiksin, A., 1994, ApJ, **431**, 109

Meiksin, A., White, M., 2001, MNRAS, **324**, 141

Meiksin, A., Bryan, G., Macheck, M., 2001, MNRAS, **327**, 296

Muecket, J.P., Petitjean, P., Kates, R.E., Riediger, R., 1996, A&A, **308**, 17

Nusser, A., Haehnelt, M., MNRAS, **303**,179

Osterbrock, D.E., 1974, *Astrophysics of Gaseous Nebulae, Freeman and Co*, San Francisco

Osterbrock, D.E., 1989, *Astrophysics of Gaseous Nebulae and Active Galactic Nuclei*, University Science Books, Mill Valley/California

Outram, P.J., Carswell, R.F., Theuns, T., 2000, ApJ, **529**, L73

Peacock, J.A., 1999, *Cosmological Physics*, Cambridge University Press

Peacock, J.A., Dodds, S.J., 1996, MNRAS, **280**, L19

Peebles, P.J.E., 1993, *Principles of Physical Cosmology*, Princeton University Press

Percival, W.J., *et al.*, 2001, MNRAS, **327**, 1297

Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P., 1992, *Numerical Recipes in Fortran*, 2nd Edition, Cambridge University Press

Press, W.H., Rybicki, G.B., Schneider, D.P., 1993, ApJ, **414**, 64

Rae, A.I.M, 1992, *Quantum Mechanics 3rd Edition*, Institute of Physics Publishers

Rauch, M., 1998, ARA&A, **36**, 267

Rauch, M., Carswell, R.F., Chaffee, F.H., Foltz, C.B., Webb, J.K., Weymann, R.J., Bechtold, J., Green, R.F., 1992, ApJ, **390**, 387

Sachs, R.K., Wolfe, A.M., 1967, ApJ, **147**, 73

Scoccimarro, R., Freman, J.A., 1996, ApJ, **473**, 620

Seljak, U., Zaldarriaga, M., 1996, ApJ, **469**, 437

Smoot, G.F., *et al.*, 1982, ApJ, **396**, L1

Spitzer, L., 1978, *Physical Processes In The Interstellar Medium*, Wiley-Interscience

Steidel, C.C., Sargent, W.L.W., 1987, ApJ, **313**, 171

Theuns, T., Leonard, A., Efstathiou, G., Pearce, F.R., Thomas, P.A., 1998, MNRAS, **301**, 478

Tytler, D., 1987, ApJ, **321**, 49

Tytler, D., O'Meara, J.M., Suzuki, N., Lubin, D., 2000, Physics Scripta, Volume T, **85**, 12

Weinberg, D.H., 1992, MNRAS, **254**, 315

Williger, G.M., Baldwin, J.A., Carswell, R.F., Cooke, A.J., Hazard, C., Irwin, M.J., McMahon, R.G., Storrie-Lombardi, L.J., 1994, ApJ, **428**, 574

Verde, L., *et al.*, 2001, astro-ph/0112161

Viel, M., Matarrese, S., Mo, H.J., Haehnelt, M.G., Theuns, T., 2002, MNRAS, **329**, 848

Zaldarriaga, M., Hui, L., Tegmark, M., 2001, ApJ, **557**, 519

Zaldarriaga, M., Scoccimarro, R., Hui, L., 2001 astro-ph/0111230

Zel'dovich, Y.B., 1970, A&A, **5**, 84

Zhang, Y., Anninos, P., Norman, M.L., 1995, ApJ, **453**, L57

Zhang, Y., Meiksin, A., Anninos, P., Norman, M.L., 1998, ApJ, **485**, 496

Zuo, L., Lu, L., 1993, ApJ, **418**, 601