

Copyright  
by  
Richard Hsueh-Yee Leu  
2011

The Dissertation committee for Richard Hsueh-Yee Leu

Certifies that this is the approved version of the following dissertation:

## Inferencing Neutrino Mass Hierarchy from Cosmology

Committee:

---

Duane Dicus, Supervisor

---

Arno Bohm

---

Charles Chiu

---

Eiichiro Komatsu

---

Greg Sitz

**Inferencing Neutrino Mass Hierarchy from Cosmology**

by

**Richard Hsueh-Yee Leu, B.S.As.E, B.S.Phy., M.S.Stat.**

**DISSERTATION**

Presented to the Faculty of the Graduate School of  
The University of Texas at Austin  
in Partial Fulfillment  
of the Requirements  
for the Degree of

**DOCTOR OF PHILOSOPHY**

THE UNIVERSITY OF TEXAS AT AUSTIN

May 2011

Dedicated to my wife Ivonne.

## Acknowledgments

First and foremost, I wish to thank my doctoral advisor: Duane Dicus. The completion of this work would not be possible without him. His constant support and patience were instrumental throughout this process. I feel extremely fortunate to have him as my advisor.

Secondly, I would like to thank the many people and resources that helped me in a technical way. The computational calculations were done with the Texas Advanced Computing Center resources. Some of the statistical analysis was done on the resources provided by the Division of Statistics and Scientific Computing. In addition, I would like to thank the dissertation committee for their suggestions. I would also like to acknowledge the help of graduate coordinator Matt Ervin.

Lastly, I would like to thank my friends and family for their support. In particular, I would like to thank my father for his encouragement and advise on the PhD process.

# Inferencing Neutrino Mass Hierarchy from Cosmology

Publication No. \_\_\_\_\_

Richard Hsueh-Yee Leu, Ph.D.  
The University of Texas at Austin, 2011

Supervisor: Duane Dicus

The observation of solar and atmospheric neutrino oscillations place bounds on the mass differences. However, these probes are insensitive to the absolute mass. To date, cosmology has provided the best bounds on the total neutrino mass. These bounds are based on a degenerate mass model. With the increasing precision of cosmological data, we investigate the effect of the neutrino mass hierarchy. The precision of the parameter estimates stems from precise observations of the cosmic microwave background. However, the effect of neutrino mass hierarchy on this observation is smaller than the cosmic variance. Therefore, we rely on the measurement of the matter power spectrum for hierarchy effects. We propose the use of importance sampling rather than the commonly used Markov chain Monte Carlo. Importance sampling takes advantage of the microwave background's statistical insensitivity to hierarchy. We present forecasted bounds due to Planck and the proposed CMBPol. We also discuss the needed precision for future galaxy surveys in detecting the effect of neutrino mass hierarchy.

# Table of Contents

<b>Acknowledgments</b>	<b>v</b>
<b>Abstract</b>	<b>vi</b>
<b>List of Tables</b>	<b>x</b>
<b>List of Figures</b>	<b>xi</b>
<b>Chapter 1 Introduction</b>	<b>1</b>
<b>Chapter 2 Cosmic Equations</b>	<b>5</b>
2.1 The FRW Universe . . . . .	5
2.2 Perturbations to the Metric . . . . .	13
2.3 Perturbations to the Boltzmann Equations . . . . .	17
2.3.1 Massless Case . . . . .	17
2.3.2 Massive Case . . . . .	21
2.3.3 Radiation Hierarchy . . . . .	25
2.4 Summary . . . . .	28
<b>Chapter 3 Observables</b>	<b>31</b>
3.1 Geometric Probes . . . . .	32
3.1.1 Standard Candles . . . . .	33
3.1.2 Supernovae Type Ia . . . . .	37
3.1.3 Cepheids . . . . .	39
3.1.4 Determining the Hubble Constant . . . . .	40
3.2 Cosmic Microwave Background . . . . .	43
3.2.1 Theoretical Angular Power Spectrum . . . . .	46
3.2.2 Estimation of the Angular Power Spectrum . . . . .	50
3.2.3 'Theoretical' Pixelated Sky . . . . .	52

3.2.4	Observed Anisotropy and Noise . . . . .	53
3.2.5	Error Analysis and Forecasting . . . . .	55
3.3	Large Scale Structure . . . . .	57
3.3.1	Theoretical Matter Power Spectrum . . . . .	59
3.3.2	Galaxy Surveys . . . . .	61
3.3.3	Halo Power Spectrum . . . . .	63
3.3.4	Forecasting . . . . .	64
<b>Chapter 4</b>	<b>Massive Neutrinos in Cosmology</b>	<b>66</b>
4.1	Effect on the FRW universe and perturbations . . . . .	67
4.2	Effect on Observables . . . . .	69
4.3	Numerical Calculations . . . . .	73
4.4	Effects from Hierarchy . . . . .	76
<b>Chapter 5</b>	<b>Numerical and Computational Aspects</b>	<b>80</b>
5.1	General Accuracy . . . . .	80
5.1.1	Integration Scheme . . . . .	81
5.1.2	Numerical Radiation Hierarchy . . . . .	84
5.1.3	Calculation of Observables . . . . .	85
5.2	Testing Results . . . . .	86
<b>Chapter 6</b>	<b>Inference with Importance Sampling</b>	<b>92</b>
6.1	General Framework of Importance Sampling . . . . .	93
6.2	Application to Cosmological Inference . . . . .	98
6.3	Obtaining Upper Bounds on Mass . . . . .	100
6.3.1	Summary of the Truncated Normal . . . . .	101
6.3.2	Estimation of the Truncated Normal . . . . .	103
6.3.3	Making Bounds on the Neutrino Mass . . . . .	104
6.4	Testing Importance Sampling . . . . .	104
6.4.1	Testing the Truncated Normal Fitting . . . . .	105
6.4.2	Testing Importance Sampling . . . . .	106

<b>Chapter 7</b>	<b>Current and Forecasted Bounds on Neutrino Mass</b>	<b>110</b>
7.1	Bounds from Current Data . . . . .	112
7.2	Effect of Future CMB Measurements . . . . .	113
7.3	Effect of Hubble Constant . . . . .	117
7.4	Future Galaxy Surveys . . . . .	120
<b>Chapter 8</b>	<b>Summary</b>	<b>125</b>
<b>Appendices</b>		<b>129</b>
<b>Appendix A</b>	<b>Statistical Properties</b>	<b>130</b>
<b>Appendix B</b>	<b>Derivation of Mean and Variance of Truncated Normal</b>	<b>132</b>
B.1	Mean . . . . .	133
B.2	Variance . . . . .	135
<b>Appendix C</b>	<b>Numerical Techniques</b>	<b>138</b>
C.1	Numerical Integration of functions . . . . .	138
C.2	Numerical Integration of ODES . . . . .	140
C.3	Optimization and Nonlinear Solvers . . . . .	141
<b>Bibliography</b>		<b>143</b>

## List of Tables

4.1	Testing Values for Non-Neutrino Parameters. $H_0$ in units of (km/s)/Mpc . . . . .	76
4.2	Testing Values for Neutrino Masses . . . . .	76
6.1	Estimation of $\mu$ , $\sigma$ and $\sum m_\nu$ based on MCMC chain . . . . .	106
6.2	Importance Sampling of WMAP7 by $H_0$ for $\Lambda$ CDM Model . . . . .	107
6.3	Importance Sampling of WMAP7 by LRG for $\Lambda$ CDM Model . . . . .	108
6.4	Importance Sampling of WMAP7 by $H_0$ & LRG for $\Lambda$ CDM + $\sum m_\nu$ Model . . . . .	109
7.1	Importance Sampling of WMAP7 by $H_0$ & LRG for $\Lambda$ CDM + $\sum m_\nu$ Model . . . . .	112
7.2	Experimental Set-ups[1] . . . . .	115
7.3	Forecasted Errors on Parameters from Planck and CMBPol[1] . . . . .	115
7.4	Importance Sampling of Planck forecasts by LRG . . . . .	116
7.5	Importance Sampling of CMBPol forecasts by LRG . . . . .	119
7.6	Importance Sampling of Planck forecasts by $H_0$ Forecast & LRG . . . . .	119

## List of Figures

4.1	Estimated Power Spectrum from [2] . . . . .	71
4.2	Effect of Massive Neutrinos on Power Spectrum[3] . . . . .	72
4.3	Interpolation error from 8-fold increase in calculation of $\rho_\nu$ . Integral calculated at points from 0.01 eV to 600 eV. Numbers on horizontal axis are the array indexes for points between this range. . . . .	75
4.4	Percent difference between Mass hierarchies and Degeneracy . . . . .	77
4.5	Percent difference between Normal and Inverted Hierarchy . . . . .	78
4.6	Percent Difference in the CMB between Neutrino Hierarchy and Degeneracy . . . . .	79
5.1	Accuracy of Power Spectrum for <code>boost = 1</code> . . . . .	88
5.2	Accuracy of Power Spectrum for <code>lboost = 1</code> . . . . .	89
5.3	Accuracy of Power Spectrum for <code>boost = 2</code> . . . . .	90
5.4	Accuracy of Power Spectrum for <code>lboost = 2</code> . . . . .	91
6.1	Estimated distribution of Neutrino mass for Various data sets[4] . . . . .	100
7.1	Probability Density of $\sum m_\nu$ for WMAP + SDSS LRG + $H_0$ . . . . .	113
7.2	Probability Density of $m_1$ for WMAP + SDSS LRG + $H_0$ . . . . .	114
7.3	Probability Density of $\sum m_\nu$ for Planck + LRG . . . . .	117
7.4	Probability Density of $m_1$ for Planck + LRG . . . . .	118
7.5	Probability Density of $\sum m_\nu$ for CMBPol + LRG . . . . .	120
7.6	Probability Density of $m_1$ for CMBPol + LRG . . . . .	121
7.7	Probability Density of $\sum m_\nu$ for Planck + $H_0$ + LRG . . . . .	122
7.8	Probability Density of $m_1$ for Planck + $H_0$ + LRG . . . . .	123
7.9	Forecasted Errors of Matter Power Spectrum . . . . .	124

# Chapter 1

## Introduction

Neutrinos have been of high interest in particle physics for the past few decades. The experimental evidence that neutrinos are massive provides a window to physics beyond the standard model. Measurements of the mixing angles have been done through observations of solar and atmospheric oscillations as well as reactor and accelerator experiments. These experiments are sensitive to mass squared differences only. To date, there have been no particle experiments sensitive enough to measure the absolute mass.

Cosmology offers an alternative route. The past couple of decades have brought cosmology from a qualitative field to a quantitative one. The precision of the data has allowed cosmologists to place bounds, sometimes stringent, on their model parameters. The collection of data has led to a standard model of cosmology where only 20% of the total matter is Baryonic and the majority of universal expansion is due to the mysterious Dark Energy. In these models, neutrinos were treated as massless and whose behavior was completely tied to photons. When neutrinos are treated massive, they act as hot dark matter (to differentiate from cold dark matter).

Currently, cosmology offers better bounds on the absolute mass. The

WMAP measurement of the cosmic microwave background (CMB) has a bound of  $\sum m_\nu < 1.3$  eV [5][6]. Including other linear data, the WMAP team found a bound of  $\sum m_\nu < 0.58$  eV [6]. A bound of  $\sum m_\nu < 0.28$  was found by combining this data with the Mega-z [4]. These upper bounds change according which data sets are chosen. Seemingly as more and more data is included, the smaller the inferred upper bounds become. The above examples demonstrate this. This statement is important if we understand how neutrinos are treated in cosmology models. Massive neutrinos are treated as degenerate. By degenerate, we mean that all the mass eigenstates have the same mass. Therefore, there is no neutrino hierarchy. This assumption is valid as long as the total mass is large enough where the known mass differences are small. However, as cosmologists push this limit downwards, it may be time to revisit this assumption.

Another reason for the assumption of degeneracy is the computational cost. As stated before, massive neutrinos would act like hot dark matter. Where as the behavior of cold dark matter and baryons can be approximated well by fluid equations, hot dark matter can not. Hot dark matter acts closer to radiation in the earlier universe. The majority of the computational cost in current cosmological codes involve solving the photon Boltzmann equation. Each additional radiation species would roughly double the time. In addition, the effect of neutrino hierarchy are small. The accuracy of current cosmology Boltzmann solvers may need to be improved.

The current practice of cosmic inferencing is to employ a Markov chain

Monte Carlo (MCMC) in a Bayesian framework. In Bayesian statistics, the parameters of a model are treated as random variables rather than having some true value. MCMC is a method to sample from the probability distribution of the parameters. MCMC has some advantages over traditional grid methods such as better scaling with respect to the number of parameters and ease in including new data. A practical issue of MCMC is the number of samples. The algorithm is shown to converge to the correct distribution. However, no rigorous methods have shown how fast this convergence is. Researchers have developed several “convergence tests” to determine if enough samples have been made. For a flat- $\Lambda$ CDM model, the WMAP team used 500,000 iterations. This number is expected to increase as more complex models are tested. In the case of massive neutrinos, this could mean a large increase in time needed for statistical inference.

In this work, we focus on the effects of neutrino hierarchy. We begin by reviewing the theoretical framework of cosmology in chapter 2. This review pays particular attention to the source of the radiation hierarchy. In chapter 3 we discuss three cosmological observables: Hubble constant, CMB, and large scale structure (LSS). Chapter 3 connects theory with observation and reviews the error analysis of cosmic observables. In chapter 4 we review the qualitative effects of massive neutrinos on these observables. We check the quantitative effects of the neutrino hierarchy on the CMB and LSS. These effects are small that in chapter 5 we examined the needed precision. In both chapter 4 and 5, we used the CAMB[7] code to integrate the perturbation equations. The

effects of neutrino hierarchy are so small that the CMB can not statistically differentiate between degenerate neutrinos or hierarchy neutrinos. In chapter 6 we discuss the importance sampling technique which takes advantage of this property. This technique is used instead of an MCMC brute-force approach. Chapter 7 applies this technique on current and forecasted data. Chapter 8 summarizes our work. Factors of  $\hbar$  and  $c$  are set to 1 unless specifically stated. We adopt the  $\{-, +, +, +\}$  metric convention.

# Chapter 2

## Cosmic Equations

In this section we review some of the basic principles of cosmology. This discussion begins with the cosmological principle and the Friedman-Robertson-Walker (FRW) metric. In this base model or zero order model, the universe is treated as homogeneous and isotropic. The universe depends on only one parameter: time. Some fundamental observables such as light element abundance and supernova luminosity distances only require knowledge of the zero order model. However, the universe is not perfectly homogeneous. Therefore, perturbations to this FRW universe need to be studied. We review the first order perturbations and their connection to observables such as the cosmic microwave background (CMB) and the large scale structure (LSS).

### 2.1 The FRW Universe

We begin with the description of space-time geometry through the metric. Note that the choice of the metric is independent of any gravitational law. In order to choose an appropriate metric we need to decide on some basic principles about our universe. This is termed the Copernican Principle by Bondi[8]. It states that no position in the universe is special.

**Copernican Principle** - The Earth is not at a special point in the universe.

The cosmological principle is a more testable version of the Copernican principle.

**Cosmological Principle** - Our universe is homogeneous and isotropic

Therefore, a metric should describe a universe that is homogeneous and isotropic. This principle is not based solely on simplicity but also on observation. At large scales, the structure of galaxies appears homogeneous[9]. We want a metric that describes this observation as a starting point. Friedman[10][11] formulated this metric as a solution to the Einstein equations for an expanding universe. Lemaitre[12] independently came to the same conclusion. Robertson[13] [14] [15] and Walker [16] showed that this metric was the *only* solution to the cosmological principle.

The FRW metric is defined within the line element as

$$ds^2 = -dt^2 + a^2(t)\gamma_{ij}dx^i dx^j$$

in rectangular coordinates where  $\gamma_{ij}$  is the spatial metric and  $dx^i = \{dx, dy, dz\}$ , or in polar coordinates

$$ds^2 = -dt^2 + a^2(t) \left( \frac{dr^2}{1 - kr^2} + r^2 d\theta^2 + r^2 \sin^2 \theta d\phi^2 \right).$$

For this paper, we restrict our discussion to a flat universe:  $k = 0$ . This choice is not only convenient but also matches observations of the CMB[17][18][19].

The only dynamical variable in the metric is the scale factor:  $a(t)$ . The scale factor describes the expansion or size of the universe. In rectangular coordinates, a flat universe leads to  $\gamma_{ij} \rightarrow \delta_{ij}$  where  $\delta_{ij}$  is the Kronecker delta.

We can now apply this metric to the Einstein equations. In particular, we can solve for the Ricci tensor and scalar.

$$R = R^\alpha_\alpha = g^{\alpha\beta} R_{\alpha\beta}$$

$$R_{\mu\nu} = R^\alpha_{\mu\alpha\nu}$$

where  $R^\alpha_{\beta\gamma\delta}$  is the Riemann tensor defined as

$$R^\alpha_{\beta\gamma\delta} = \partial_\gamma \Gamma^\alpha_{\delta\beta} - \partial_\delta \Gamma^\alpha_{\gamma\beta} + \Gamma^\alpha_{\gamma\mu} \Gamma^\mu_{\delta\beta} - \Gamma^\alpha_{\delta\mu} \Gamma^\mu_{\gamma\beta}$$

The Christoffel symbol,  $\Gamma_{\alpha\beta\gamma}$ , is defined as

$$\Gamma_{\alpha\beta\gamma} = \frac{1}{2}(\partial_\gamma g_{\alpha\beta} + \partial_\beta g_{\alpha\gamma} - \partial_\alpha g_{\beta\gamma}).$$

For the FRW metric, it is easier to work in rectangular coordinates. The non-zero  $\Gamma^\alpha_{\beta\gamma}$ 's are

$$\Gamma^0_{ij} = \delta_{ij} a \dot{a}$$

$$\Gamma^i_{0j} = \delta_{ij} \frac{\dot{a}}{a}$$

where the over-dot denotes the time derivative. This notation differs from the convention found in literature[20] where the over-dot is the derivative with respect to conformal time:  $\tau = \int a(t) dt$ . The Ricci tensor components are

$$R_{00} = 0 - 3 \left( \frac{\ddot{a}}{a} - H^2 \right) + 0 - 3(H^2) = -3 \frac{\ddot{a}}{a}$$

$$R_{ij} = (a\ddot{a} + 2\dot{a}^2) \delta_{ij},$$

and the Ricci scalar is

$$R = g^{\mu\nu} R_{\mu\nu} = (-1) \left( -3 \frac{\ddot{a}}{a} \right) + 3 \left( \frac{1}{a^2} \right) (a\ddot{a} + 2\dot{a}^2) \delta_{ij} = 6 \left( H^2 + \frac{\ddot{a}}{a} \right).$$

The Einstein equation from general relativity is

$$R_{\mu\nu} + \frac{1}{2} g_{\mu\nu} R = 8\pi G T_{\mu\nu}$$

where  $T_{\mu\nu}$  is the stress energy tensor. For now, we assume that the stress-energy tensor has the perfect fluid form

$$T_{\nu}^{\mu} = \begin{pmatrix} -\rho & 0 & 0 & 0 \\ 0 & P & 0 & 0 \\ 0 & 0 & P & 0 \\ 0 & 0 & 0 & P \end{pmatrix}$$

where  $\rho$  is the energy density and  $P$  is the pressure. Applying the FRW metric to the 00 component of the Einstein equation results in the Friedmann equation, namely

$$3H^2 = 8\pi G\rho \tag{2.1}$$

where  $H = \dot{a}/a$  is the Hubble rate. The spatial components of the left-hand side of the Einstein equation yield

$$(a\ddot{a} + 2\dot{a}^2) - \frac{1}{2} a^2 \left( 6 \left( H^2 + \frac{\ddot{a}}{a} \right) \right) = -2 \left( \frac{\dot{a}^2}{2} + a\ddot{a} \right)$$

The resulting equation is

$$\frac{\ddot{a}}{a} + \frac{1}{2} H^2 = -4\pi G P \tag{2.2}$$

In the FRW universe, equations 2.1 and 2.2 govern the dynamics of the scale factor. This dynamic depends on the energy density and pressure of matter.

Now we need to understand the dynamics of the energy density and pressure. To this end, we relate the stress energy tensor to the phase space distribution,  $f$ . The phase space distribution is the number density at a given time, location, and momentum. When integrated over the entire phase-space, the result is the number of particles.

$$N = \int d^3x d^3p f(x, p, t)$$

In cosmology, we are more interested in densities such as the number density.

$$n = \int d^3p f(x, p, t) \quad (2.3)$$

In general, the components of the stress energy tensor are related to the phase space distribution by

$$T_{\nu}^{\mu} = g \int \frac{d^3P}{(2\pi)^3} \frac{1}{\sqrt{-\det g}} \frac{P^{\mu} P_{\nu}}{P^0} f(x, p, t) \quad (2.4)$$

where  $g$  is the number of degeneracies (e.g. number of spin states),  $P^{\mu} = \frac{dx^{\mu}}{d\lambda}$ ,  $d^3P = dP_1 dP_2 dP_3$ , and  $\lambda$  is an affine parameter. We call  $P^{\mu}$  the comoving momentum. However, it is more convenient to work with  $p^2 = g^{ij} P_i P_j$  where  $p$  is termed the physical momentum. With the relation  $dP^i = a dp^i$ , equation 2.4 becomes

$$T_{\nu}^{\mu} = g \int \frac{d^3p}{(2\pi)^3} \frac{g_{\nu\alpha} P^{\mu} P^{\alpha}}{E} f(x, p, t) \quad (2.5)$$

where the energy  $E = \frac{dt}{d\lambda}$ .

To proceed, we need to understand the relationship between the energy,  $E$ , and the comoving momentum  $P$ . This relationship is simply  $P_\mu P^\mu = -E^2 + p^2 = -m^2$ . For a very non-relativistic species,  $E \approx m$  while for a very relativistic species,  $E \approx p$ . In general the trace of equation 2.5 becomes

$$-\rho + 3P = g \int \frac{d^3p}{(2\pi)^3} \frac{-E^2 + p^2}{E} f(x, p, t)$$

or

$$= g \int \frac{d^3p}{(2\pi)^3} \frac{-m^2}{E} f(x, p, t).$$

For a very relativistic species,  $m \rightarrow 0$  so the trace of the stress energy tensor approaches 0 as well. Thus,

$$P_r \rightarrow \frac{1}{3}\rho_r.$$

For a non-relativistic particle,  $p \rightarrow 0$  and so

$$P_m \rightarrow 0.$$

In general, the energy density and pressure are related to energy and momentum by

$$\rho = g \int \frac{d^3p}{(2\pi)^3} E f(x, p, t) \tag{2.6}$$

and

$$P = g \int \frac{d^3p}{(2\pi)^3} \frac{p^2}{3E} f(x, p, t).$$

So far, we have dynamical equations for the scale factor. Also, we know the relationship between the phase space density and  $\rho$  &  $P$ . We have yet to discuss the evolutionary behavior of  $\rho$  and  $P$ . To this end, we examine the

Boltzmann equation which is the dynamical equation for  $f$ . We denote the FRW phase space density as  $f_0$ . The general Boltzmann equation is

$$\frac{df}{dt} = C[f]$$

where  $C[f]$  is called the collision term and describes all the non-gravity interactions. This term is particular important for the coupling of Baryons and photons in the early times of the universe. Ultimately, we are only interested in neutrinos which are weakly interacting. Therefore, we ignore this term in this paper. The full time derivative can be expanded into partial derivatives:

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \frac{\partial f}{\partial x^i} \frac{dx^i}{dt} + \frac{\partial f}{\partial p^\mu} \frac{dp^\mu}{dt}. \quad (2.7)$$

At this point it is useful to separate the momentum terms into magnitude and direction

$$\frac{\partial f}{\partial p^\mu} \frac{dp^\mu}{dt} \rightarrow \frac{\partial f}{\partial p} \frac{dp}{dt} + \frac{\partial f}{\partial \hat{p}^i} \frac{d\hat{p}^i}{dt}$$

where  $p^i = p\hat{p}^i$ .

To proceed further, we need to discuss the form of  $f_0$ . Rather than depend on a specific form such as the Fermi-Dirac or Bose-Einstein distribution, we rely only the cosmological principle. In general, the phase space distribution depends on time, space, and momentum. For a homogeneous universe, we would not expect  $f_0$  to depend on position. An isotropic universe also should not depend on direction. Therefore,  $f_0 = f_0(t, p)$ , and the Boltzmann equation in an FRW universe should be

$$\frac{df_0}{dt} = \frac{\partial f_0}{\partial t} + \frac{\partial f_0}{\partial p} \frac{dp}{dt} = 0.$$

To solve for  $dp/dt$ , we resort to the geodesic equation. The geodesic equation describes the path of a particle through a curved space with a given metric.

$$\frac{d^2x^\mu}{d\lambda^2} = -\Gamma_{\alpha\beta}^\mu \frac{dx^\alpha}{d\lambda} \frac{dx^\beta}{d\lambda}$$

or

$$\frac{dP^\mu}{d\lambda} = -\Gamma_{\alpha\beta}^\mu P^\alpha P^\beta$$

First, from the definition of the mass invariance, we have

$$E \frac{dE}{dt} = p \frac{dp}{dt}.$$

Second, the geodesic leads to

$$E \frac{dE}{dt} = -p^2 H. \tag{2.8}$$

Therefore, the Boltzmann equation becomes

$$\frac{\partial f_0}{\partial t} = p H \frac{\partial f_0}{\partial p}$$

Based on equation 2.6, we can integrate the above equation by  $g \int \frac{d^3p}{(2\pi)^3} E$ . This step will result in a dynamical equation for  $\rho$ .

$$\dot{\rho} = H g \int \frac{d^3p}{(2\pi)^3} p E \frac{\partial f_0}{\partial p}$$

The right hand side can be solved utilizing integration by parts and noting that  $f_0$  approaches zero when  $p$  goes to zero or infinity. This step is often repeated in this chapter. The final result is

$$\dot{\rho}_r = -4H\rho_r \rightarrow \rho_r \propto a^{-4}$$

$$\dot{\rho}_m = -3H\rho_m \rightarrow \rho_m \propto a^{-3}.$$

The dynamical behavior of  $\rho$  is linked to the dynamics of the scale factor.

In summary, we have taken the FRW metric and applied it to Einstein's equation. The result is a set of two dynamical equations for the scale factor. The evolution of  $a$  depends on the type of matter-energy content, namely  $\rho$  and  $P$ . The next step is to understand the dynamics of  $\rho$  and  $P$ . To this end, we use the relationship of  $\rho$  and  $P$  to the phase space density  $f$ . The Boltzmann equation describes the evolution of  $f$ . Applying the metric, we are able to determine the behavior of  $\rho$  with respect to  $a$ . Note that  $\rho$  and  $P$  in equations 2.1 and 2.2 are the sum of the energies and pressures for all the species in the universe, respectively. In addition, each species has its own Boltzmann equation to follow.

## 2.2 Perturbations to the Metric

The previous section developed the dynamics for a smooth, homogeneous universe. As stated before, this metric is applicable at very large scales. At small scales, the universe is clumpy and inhomogeneous. To describe the inhomogeneous universe, we perturb around the FRW metric. We also require perturbations to the phase-space distributions of each particle species. For our purposes, we only discuss scalar perturbations; both tensor and vector perturbations are possible. There also exists a gauge freedom on how to perturb around the metric. Throughout this paper, we work in the Newtonian gauge[21]. See [22] for a complete treatment including vector and tensor per-

turbations and other gauge choices. The perturbations in this gauge have the form

$$g_{00} = -(1 + 2\psi(x, t))$$

$$g_{ij} = a^2(1 + 2\phi(x, t))\delta_{ij}.$$

For linear perturbation theory to be valid, we assume  $\psi$  and  $\phi$  are *small*. This statement may be troubling since galaxies and clusters of galaxies exist. Do small perturbations have anything to do with the actual universe? Observations of the CMB show that very small perturbations did exist during earlier periods of the universe [23]. The evolutionary model for structure formation is that the perturbations were small in the early periods and grew over time to form the galaxies we see today. At large scales, the perturbations should still be small enough for linear theory to be accurate.

In the FRW universe, we only had to worry about derivatives with respect to time. Here our perturbations  $\psi$  and  $\phi$  depend on space too. It is more convenient to Fourier transform the equations to remove the spatial derivatives. The Fourier transform is

$$\tilde{f}(\vec{q}) \propto \int d\vec{x} f(x) e^{-i\vec{q}\cdot\vec{x}}$$

with the inverse transform

$$f(\vec{x}) \propto \int d\vec{q} \tilde{f}(\vec{q}) e^{i\vec{q}\cdot\vec{x}}.$$

The derivative with respect to  $x$  is

$$\frac{\partial f(\vec{x})}{\partial x^i} \propto \int d\vec{q} i q_i \tilde{f}(\vec{q}) e^{i\vec{q}\cdot\vec{x}}.$$

Therefore, spatial derivatives in the Fourier space disappear.

$$\frac{\partial f(\vec{x})}{\partial x_i} \rightarrow \imath q_i \tilde{f}(\vec{q})$$

For this paper, we remove the tilde for the Fourier transformed functions.

As was done in section 2.1, we begin with the Christofel terms up to first order. The non-zero terms are:

$$\begin{aligned}\Gamma_{00}^0 &= \dot{\psi} \\ \Gamma_{0i}^0 &= \imath q_i \psi \\ \Gamma_{ij}^0 &= \delta_{ij} a^2 \left( H + 2H(\phi - \psi) + \dot{\phi} \right) \\ \Gamma_{00}^i &= \frac{\imath q_i}{a^2} \psi \\ \Gamma_{0j}^i &= \delta_{ij} \left( H + \dot{\phi} \right) \\ \Gamma_{jk}^i &= \delta_{ij} \imath q_k \phi + \delta_{ik} \imath q_j \phi - \delta_{jk} \imath q_i \phi\end{aligned}$$

The Ricci tensor to first order becomes

$$\begin{aligned}R_{00} &= -3\frac{\ddot{a}}{a} - \frac{q^2}{a^2} \psi - 3\ddot{\phi} + 3H(\dot{\psi} - 2\dot{\phi}) \\ R_{ij} &= q_i q_j (\psi + \phi) + \delta_{ij} \left[ q^2 \phi + a^2 \left( -H\dot{\psi} + 6H\dot{\phi} \right) + a^2 \ddot{\phi} \right. \\ &\quad \left. + 2a^2 H + \ddot{a} a + 4H^2 a^2 (\phi - \psi) + 2\ddot{a} a (\phi - \psi) \right]\end{aligned}$$

and the Ricci scalar becomes

$$R = 6 \left( \frac{\ddot{a}}{a} + H^2 \right) + \frac{2}{a^2} q^2 (\psi + 2\phi) + 6\ddot{\phi} - 6H(\dot{\psi} - 4\dot{\phi}) - 12\psi \ddot{a}/a - 12H^2 \psi$$

The Einstein equations result in 4 equations. For the stress-energy tensor, we assume perturbations around a perfect fluid. Let  $\rho \rightarrow \rho_0(1 + \delta)$ . Then the 00 component of the Einstein equations results in

$$\frac{q^2}{a^2}\phi + 3H(\dot{\phi} + H\psi) = -4\pi G\rho_0\delta \quad (2.9)$$

for the first order perturbations. The spatial component of the Einstein equation results in

$$q_i q_j \frac{\psi + \phi}{a^2} + \delta_{ij} \left[ -\frac{q^2}{a^2}(\psi + \phi)2H(\dot{\psi} - 3\dot{\phi}) - 2\ddot{\phi} + 2H^2(2\phi + \psi) + \frac{2\ddot{a}}{a}(\phi + 2\psi) - \frac{q^2}{a^2} \right] = 8\pi G\delta T_j^i$$

The above equation contains two parts: a diagonal part and an off-diagonal part. Therefore, we can break up the stress energy tensor perturbation into  $\delta T_j^i = \delta P\delta_{ij} + \Sigma_j^i$  where  $\Sigma_i^i = 0$ .  $\delta P$  is the perturbation to the zero order pressure while  $\Sigma_j^i$  is the anisotropic perturbations. The off-diagonal part is somewhat easier to deal with. First, we apply the traceless projection operator  $\frac{1}{q^2}q^i q^j - \frac{1}{3}\delta_{ij}$  to the above equation. Then, we define the anisotropic stress to be

$$\sigma = -\frac{1}{\rho_0 + P_0} \left( \frac{1}{q^2}q^i q^j - \frac{1}{3}\delta_{ij} \right) \delta T_j^i$$

The result is

$$\frac{1}{a^2}q^2(\psi + \phi) = -12\pi G(\rho_0 + P_0)\sigma \quad (2.10)$$

The other two equations can be derived from the trace component of  $T_j^i$  and the perturbed part of  $T_i^0$ . However, only two of the four equations are needed to evolve  $\psi$  and  $\phi$ .

## 2.3 Perturbations to the Boltzmann Equations

In section 2.1,  $f_0$  only depended on the magnitude of the momentum and time. Now, with the introduction of the perturbations,  $f = f(t, x, p)$ . We examine equation 2.7 and only keep those in first order. Note that both  $\frac{\partial f}{\partial \vec{p}^i}$  and  $\frac{d\vec{p}^i}{dt}$  are first order terms. Therefore, these two terms multiplied form a second order term; we should then drop this term altogether. After a Fourier transformation, we are left with

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + v\vec{q} \cdot \frac{d\vec{x}}{dt} f + \frac{\partial f}{\partial p} \frac{dp}{dt} = 0.$$

Note that the second term is first order only.  $f$  in this part has no zero order component. We first start with the massless case and then move onto the massive case. The massive case is done both for a non-relativistic particle and in a general way. This discussion leads to the radiation hierarchy.

### 2.3.1 Massless Case

We examine the effect of the perturbed metric for the massless case in this subsection. The massless case provides some simplification which allows some of the algebraic steps to be a little more clean as compared to the massive case. First, by definition,  $\frac{P^i}{P^0} = \frac{dx^i}{dt}$ . For a massless particle,  $g_{\mu\nu}P^\mu P^\nu = 0$ . We can expand this out such that

$$-(1 + 2\psi) (P^0)^2 + a^2(1 + 2\phi)\delta_{ij}P^i P^j = -E^2 + p^2 = 0.$$

Thus,

$$P^0 = \frac{p}{\sqrt{1 + 2\psi}} \approx p(1 - \psi)$$

where the approximation comes from Taylor expansion of the square root. We see that energy is lost as the particle pulls out of a potential well. An overdensity is a negative  $\psi$  value so  $E$  becomes smaller as the overdensity becomes less. Similarly,  $P^i$  can be written in terms of  $p$ .

$$P^i = \frac{p}{a} \frac{1}{\sqrt{1+2\phi}} \hat{p}^i \approx \frac{p}{a} (1-\phi) \hat{p}^i$$

Returning to the Boltzmann equation, we can write out the  $\frac{dx^i}{dt}$  term.

$$\frac{dx^i}{dt} = \frac{P^i}{P^0} \approx \frac{p}{a} \frac{1}{p} (1-\phi)(1+\psi) \hat{p}^i$$

To first order we are left with

$$\frac{dx^i}{dt} \approx \frac{1}{a} (1-\phi+\psi) \hat{p}^i$$

Now we see that this term multiplies  $\imath f \vec{q}$  which is a first order term itself. Therefore, we keep only the zero order of  $\frac{dx^i}{dt}$ . The resulting Boltzmann equation is

$$\frac{\partial f}{\partial t} + \imath \frac{1}{a} \vec{q} \cdot \hat{p} f + \frac{\partial f}{\partial p} \frac{dp}{dt} = 0.$$

Next, we work on  $\frac{dp}{dt}$ .

$$\frac{dp}{dt} = \frac{1}{P^0} \frac{dp}{d\lambda} = \frac{1+\psi}{P^0} \frac{dP^0}{d\lambda} + P^0 \frac{d\psi}{dt} \quad (2.11)$$

The  $\frac{d\psi}{dt}$  term can be expanded by partial derivatives.

$$\frac{d\psi}{dt} = \dot{\psi} + \imath \psi \frac{1}{a} \vec{q} \cdot \hat{p}. \quad (2.12)$$

For  $\frac{dP^0}{d\lambda}$ , the geodesic equation is used.

$$\frac{dP^0}{d\lambda} = -\Gamma_{\alpha\beta}^0 P^\alpha P^\beta = -p^2 \left[ \dot{\psi} + 2\imath \frac{\psi}{a} \vec{q} \cdot \vec{\hat{p}} + H + \dot{\phi} - 2H\psi \right] \quad (2.13)$$

Combining equations 2.12 and 2.13 into equation 2.11 yields

$$\frac{dp}{dt} = -p \left( i \frac{1}{a} \psi \vec{q} \cdot \hat{p} + H + \dot{\phi} \right). \quad (2.14)$$

Our Boltzmann equation now looks like

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + i \frac{1}{a} \vec{q} \cdot \hat{p} f - p \frac{\partial f}{\partial p} \left[ i \frac{1}{a} \vec{q} \cdot \hat{p} \psi + H + \dot{\phi} \right].$$

This equation is the collisionless Boltzmann equation for a massless particle. The first two terms are analogous to the continuity and Euler equations. The third term takes into account the expansion of the universe along with the gravitational perturbations.

We are now ready to start looking at the perturbation of the phase space distribution  $f$ . For this paper, we perturb  $f$  in two ways. The first way is to perturb the temperature  $T(t, x, p) = T(1 + \Delta)$  where  $T$  is the average or zero order temperature.  $T$  only depends on time while  $\Delta = \Delta(t, x, \hat{p})$ . This convention is useful when we are comparing the observed CMB anisotropies to theoretical predictions.  $\Delta$  can be observed directly. The phase-space distribution is

$$f(t, x, p) = \frac{1}{e^{p/T(1+\Delta)} \pm 1} \approx f_0 - p \frac{\partial f_0}{\partial p} \Delta$$

where the approximation is derived by Taylor expansion and  $f_0$  is the zero order phase space distribution as in section 2.1. By plugging this form into the Boltzmann equation, we can separate the zero order terms and the first order perturbation terms. The zero order terms for the collisionless Boltzmann yield  $\frac{df}{dt} = \frac{\partial f_0}{\partial t} - p H \frac{\partial f_0}{\partial p} = 0$ . This is the same result as in section 2.1.

We now move onto the first order terms. Collecting these terms results in

$$-p \frac{\partial f_0}{\partial p} \left( \dot{\Delta} + \iota \frac{1}{a} \vec{q} \cdot \hat{p} \Delta + \iota \frac{1}{a} \vec{q} \cdot \hat{p} \psi + \dot{\phi} \right) - p \Delta \frac{\partial^2 f_0}{\partial p \partial t} + p \Delta H \frac{\partial}{\partial p} \left( p \frac{\partial f_0}{\partial p} \right) = 0$$

The last two terms cancel since  $\frac{\partial f_0}{\partial t} \approx p H \frac{\partial f_0}{\partial p}$ . The resultant dynamic equation of the phase space perturbation  $\Delta$  is

$$\dot{\Delta} + \iota \frac{1}{a} \vec{q} \cdot \hat{p} \Delta = -\dot{\phi} - \iota \frac{1}{a} \vec{q} \cdot \hat{p} \psi \quad (2.15)$$

The second way to perturb  $f$  is to set

$$f(t, x, p) = f_0(1 + \Psi) \quad (2.16)$$

Again, the zero order Boltzmann yields  $\frac{\partial f_0}{\partial t} = p H \frac{\partial f_0}{\partial p}$  which is exactly the same result as before. To yield an observable result, we can multiply this equation with  $g \int \frac{d^3 p}{(2\pi)^3}$ . The resulting equation is

$$\dot{n}_0 - H g \int \frac{d^3 p}{(2\pi)^3} p \frac{\partial f_0}{\partial p} = 0$$

where  $n_0 \equiv g \int \frac{d^3 p}{(2\pi)^3} f_0$ . The second term can be rewritten in terms of  $n_0$  using integration by parts.

$$g \int \frac{dp}{(2\pi^2)} p^3 \frac{\partial f_0}{\partial p} = 0 - 3g \int \frac{dp}{(2\pi^2)} p^2 f_0 = -3n_0.$$

This results in

$$\dot{n}_0 + 3H n_0 = 0$$

or  $n_0 \propto a^{-3}$ . This same result can be found in the FRW universe. It simply states that, in the absence of interactions that change the number of particles,

the number density decreases as space increases. In other words  $N$  remains constant as the volume increases. These steps are repeated often in the next section.

The first order equation for  $\Psi$  is

$$\Psi \frac{\partial f_0}{\partial t} + f_0 \dot{\Psi} + \imath \frac{1}{a} \vec{q} \cdot \hat{p} f_0 \Psi - p \frac{\partial f_0}{\partial p} \left( \imath \frac{1}{a} \vec{q} \cdot \hat{p} \psi + \dot{\phi} \right) - p H \frac{\partial}{\partial p} (f_0 \Psi) = 0.$$

The first and last terms cancel each other since  $\frac{\partial f_0}{\partial t} = p H \frac{\partial f_0}{\partial p}$  as before and noting that  $\Psi$  does not depend on  $p$  as long as interactions are elastic. This is the same assumption made when  $\Delta$  depended on  $\hat{p}$  but not  $p$ . The end result is

$$\dot{\Psi} + \imath \frac{1}{a} \vec{q} \cdot \hat{p} \Psi = \frac{p}{f_0} \frac{\partial f_0}{\partial p} \left( \imath \frac{1}{a} \vec{q} \cdot \hat{p} \psi + \dot{\phi} \right) \quad (2.17)$$

The two perturbation schemes are linked by equation 83 in Ma and Bertschinger[20].

The first method is convenient for calculation of the photons and the CMB.

The second method is more general and easier to relate to observations of matter perturbations as seen in the next subsection.

### 2.3.2 Massive Case

For the massive, non-relativistic case, we can approach the Boltzmann equation in two ways. The first way is to assume a fluid approximation. We expect massive particles to behave like pressure-less perfect fluids. In this case we would only need two equations: continuity equation and Euler's equation. As long as a particle has a very small velocity then the pressure and anisotropic stress can be ignored.

Alternatively, one can find this approximation with the Boltzmann equation. We proceed in this way in order to understand the fluid approximation. Again, we begin with the collisionless Boltzmann seen in equation 2.7. Of course the main difference between the massive case and the massless case is

$$P^\mu P_\mu = -m^2 = -E^2 + p^2$$

where the above equation defines  $E$  and  $p$ . Unlike the massless case, we can not interchange  $E$  with  $p$ .  $P^\mu$  can be written in terms of  $E$  and  $p$  such that

$$P^\mu \approx \left\{ E(1 - \psi), \frac{p}{a}(1 - \phi)\hat{p} \right\}.$$

The algebraic work to write equation 2.7 in terms of  $E$  and  $p$  is similar to the massless case. Working through the same steps as the massless case, we obtain

$$\begin{aligned} \frac{dx^i}{dt} &= \frac{p}{aE}\hat{p} \\ \frac{dE}{dt} &= - \left[ v\psi \frac{p}{a}\vec{q} \cdot \hat{p} + \frac{p^2}{E} (H + \dot{\phi}) \right]. \end{aligned}$$

The resulting Boltzmann equation is

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \frac{p}{aE}\vec{q} \cdot \hat{p}f - \left[ v\psi \frac{p}{a}\vec{q} \cdot p + \frac{p^2}{E} (H + \dot{\phi}) \right] \frac{\partial f}{\partial E}. \quad (2.18)$$

Setting the perturbation the same as equation 2.16 and working through the same algebraic steps results in

$$\dot{\Psi} + v \frac{1}{a} \frac{p}{E} \vec{q} \cdot \hat{p} \Psi = - \frac{\partial \log f_0}{\partial \log p} \left( v \frac{1}{a} \frac{E}{p} \vec{q} \cdot \hat{p} \psi + \dot{\phi} \right) \quad (2.19)$$

where  $\frac{\partial \log f_0}{\partial \log p} = \frac{p}{f_0} \frac{\partial f_0}{\partial p} = \frac{p^2}{E} \frac{1}{f_0}$ . The form is identical to equation 2.17 except that factors of  $\frac{p}{E}$  appear. Note that we have yet to say whether the particle is non-relativistic or relativistic. Equation 2.19 can be applied to any type of massive particle.

We would now like to connect equation 2.19 to the observed large scale structure. Unlike the CMB, we do not observe the temperature perturbations but rather density perturbations. In other words we are more concerned with quantities like number density. Therefore, we proceed as in the massless case and find the dynamic equation for  $n$ . First, we go back to the relationship between  $f$  and  $T_\nu^\mu$ . In section 2.2, we defined the perturbations to the stress energy tensor such that

$$\begin{aligned}\delta T_0^0 &= g \int \frac{d^3 p}{(2\pi)^3} E f_0 \Psi = \rho_0 \delta \\ \delta T_j^i &= g \int \frac{d^3 p}{(2\pi)^3} \frac{p^2 \hat{p}^i \hat{p}^j}{E} f_0 \Psi = \delta P \delta_{ij} + \Sigma_j^i\end{aligned}$$

In addition, we define  $\vec{v}$  such that

$$\delta T_i^0 = g \int \frac{d^3 p}{(2\pi)^3} p \hat{p}^i f_0 \Psi = (\rho_0 + P_0) v^i. \quad (2.20)$$

Note that  $v^i$  is a perturbation with no analogous term in an FRW universe. It is related to the energy flux or momentum density.

The observed large scale structure is the power spectrum of  $\delta$ . To find a dynamical equation for  $\delta = \delta\rho/\rho_0$ , we work with

$$\dot{\delta} = \frac{1}{\rho_0} \left( \dot{\delta}\rho - \dot{\rho}_0 \delta \right). \quad (2.21)$$

For the second term, section 2.1 showed that  $\dot{\rho}_0 = -4H\rho_0$  for the massless case and  $\dot{\rho}_0 = -3H\rho_0$  for the non-relativistic case. For the first term,

$$\dot{\delta\rho} = g \int \frac{d^3p}{(2\pi)^3} E \left( \dot{f}_0 \Psi + f_0 \dot{\Psi} \right).$$

Since  $\dot{f}_0 = pH \frac{\partial f_0}{\partial p}$ , the first part of  $\dot{\delta\rho}$  cancels with  $\dot{\rho}_0 \delta$ , and we are left with

$$\dot{\delta} = \frac{1}{\rho_0} \left( g \int \frac{d^3p}{(2\pi)^3} E f_0 \dot{\Psi} \right)$$

If we now apply equation 2.19 for  $\dot{\Psi}$ , the resulting equation is

$$\dot{\delta} + \iota \frac{1}{a} \vec{q} \cdot \vec{v} = -3\dot{\phi}$$

The evolution of  $\delta$  depends on the metric perturbations and  $\vec{v}$ . Therefore, we also need to find the time derivative of  $\vec{v}$ .

The time derivative of equation 2.20 yields

$$\frac{\partial}{\partial t} [\vec{v}(\rho_0 + P_0)] = g \int \frac{d^3p}{(2\pi)^3} p \hat{p} \left( \dot{f}_0 \Psi + f_0 \dot{\Psi} \right)$$

First, let us assume we are dealing with non-relativistic particles:  $P_0 = 0$ . We have already encountered  $\dot{\rho}_0$  and  $\dot{f}_0$  in dealing with  $\dot{\delta}$ . These two terms combine into  $H\vec{v}$  following the same steps as before. Also, we can plug in the Boltzmann equation for  $\dot{\Psi}$  as before which yields

$$g \int \frac{d^3p}{(2\pi)^3} p \hat{p} f_0 \dot{\Psi} = g \int \frac{d^3p}{(2\pi)^3} p \hat{p} f_0 \left[ -\iota \frac{1}{a} \frac{p}{E} \vec{q} \cdot \hat{p} \Psi + \frac{p}{f_0} \frac{\partial f_0}{\partial p} \left( \iota \frac{1}{a} \frac{E}{p} \vec{q} \cdot \hat{p} \psi + \dot{\phi} \right) \right].$$

The first term on the right hand side can be ignored for the non-relativistic case since  $p \ll E$ . The last term with  $\dot{\phi}$  disappears since integrals of  $\hat{p}$  over

the entire phase space are zero. The remaining term, using integration by parts, becomes  $-\imath\frac{1}{a}\vec{q}\rho_0\psi$ . The resulting dynamical equation for  $\vec{v}$  is

$$\dot{\vec{v}} + H\vec{v} = -\imath\frac{1}{a}\vec{q}\psi.$$

Rather than deal with a vector, we define  $\theta = \imath\vec{q} \cdot \vec{v}$ . Finally, the equations describing the evolution of non-relativistic particles are

$$\begin{aligned}\dot{\delta} + \frac{1}{a}\theta &= -3\dot{\phi} \\ \dot{\theta} + H\theta &= \frac{1}{a}q^2\psi\end{aligned}$$

These are the fluid equations for non-relativistic collisionless particles.

### 2.3.3 Radiation Hierarchy

Section 2.3.2 connected the Boltzmann equation for  $\Psi$  to the observable  $\delta$ . This step is also important since the evolution equations of  $\phi$  and  $\psi$  depend on  $\delta$ ,  $\theta$ , and  $\sigma$ . In section 2.3.1, this step was avoided since for photons we can observe  $\Delta$ . However, in the case of massless neutrinos or any other type of non-standard radiation, it is unlikely that we can observe  $\Delta$  directly. In addition, we still need to connect  $\Delta$  to equations 2.9 and 2.10 through  $\delta$  and  $\sigma$ .

To this end, we apply the steps of section 2.3.2 to the massless case. For  $\delta$  we begin with equation 2.21.

$$\dot{\delta} = \frac{1}{\rho_0} \left( \dot{\delta}\rho - \rho_0\dot{\delta} \right)$$

The steps taken here are identical to the massive case except that  $\dot{\rho}_0 = -4H\rho_0$  and  $E = p$ . This results in a similar equation for  $\delta$ .

$$\dot{\delta} + \imath \frac{4}{3} \frac{1}{a} \theta = -4\dot{\phi}$$

The only difference here between the massive and massless case are the coefficients. This result should not be too surprising since this equation is the analog of the continuity equation in fluid mechanics.

The dynamical equation for  $\theta$  follows the same steps as in the massive case but with a few important differences. First, in section 2.3.2 one term was ignored assuming  $p \ll E$ ; this term can not be ignored here. Second, the  $H\theta$  term is canceled since  $\dot{\rho}_0 = -4H\rho_0$  instead of  $-3H\rho_0$ . Lastly, we use the relation  $\delta P = \frac{1}{3}\delta\rho$ . The dynamical equation for  $\theta$  is

$$\dot{\theta} = \frac{1}{a} q^2 \left( \frac{1}{4} \delta - \sigma + \psi \right).$$

Instead of  $\dot{\theta}$  depending on  $\theta$ , it depends on  $\delta$  and  $\sigma$ . While we already have a dynamical equation for  $\delta$ , we do not have one for  $\sigma$ . It turns out, these steps are infinitely repeated. For example,  $\dot{\sigma}$  depends on  $\theta$  and a higher moment. This higher moment's time derivative depends on  $\sigma$  and the next higher moment. This infinite set of equations is termed the radiation hierarchy.

The dynamical equation for each moment depends on the neighboring moments. This behavior can be seen by taking the Legendre moments of equation 2.17. Let  $\Psi_\ell = (\imath)^\ell \int_{-1}^1 (d\mu/2) P_\ell(\mu) \Psi$  where  $\mu = (1/q)\vec{q} \cdot \hat{p}$  and  $P_\ell$  denotes the Legendre polynomials. Two properties of Legendre polynomials

are important. One, they are orthogonal:  $\frac{1}{2} \int d\mu P_\ell P_{\ell'} = \delta_{\ell\ell'}/(2\ell + 1)$ . Two, there exists a recursion relation:  $(2\ell + 1)\mu P_\ell = (\ell + 1)P_{\ell+1} - \ell P_{\ell-1}$ . The Legendre moment of equation 2.17 is

$$\dot{\Psi}_\ell + \imath \frac{1}{a} (-\imath)^\ell \int \frac{d\mu}{2} P_\ell(\mu) \mu \Psi = (-\imath)^\ell \int \frac{d\mu}{2} P_\ell(\mu) (\imath \frac{1}{a} \mu \psi + \dot{\phi}).$$

For the right hand side, the  $\dot{\phi}$  term is only non-zero when  $\ell = 0$  since  $P_0 = 1$ . For the same reason, the  $\psi$  term is only non-zero when  $\ell = 1$ . For  $\ell \geq 2$ , the right hand side vanishes. The second term on the left hand side contains  $\mu P_\ell(\mu)$ . This term can be expanded based on the recursion relation and is the source of the neighboring moments characteristic. In the end

$$\dot{\Psi}_\ell = -\frac{1}{a} \frac{q}{2\ell + 1} [(\ell + 1)\Psi_{\ell+1} - \ell\Psi_{\ell-1}] \quad \ell \geq 2 \quad (2.22)$$

For  $\ell = 0, 1$ , additional terms appear from the metric perturbations.

For the massive case, the fact that  $p \ll E$  led to a truncation after the first moment. However, for the massless case, a full solution would be an infinite amount of equations where each moment's dynamic equation depends on the neighboring moments. Of course for practicality, this infinite set is truncated. Rather than an abrupt cutoff, Ma & Bertschinger[20] proposed

$$\Psi_{\ell-\max} \rightarrow \frac{2\ell - \max + 1}{q\tau} \Psi_{\ell-\max} - \Psi_{\ell-\max-1}$$

where details on the reasoning behind this cutoff can be found.

## 2.4 Summary

Without thorough derivation we present the Boltzmann equations and metric perturbations. For neutrinos and CDM, we only require the collisionless Boltzmann equation which we have derived. The extra terms in the Baryon and photon equations come from the collision term based on Compton scattering. As stated before the metric perturbations evolve based on 4 equations but only 2 are necessary. Equations 2.9 and 2.10 were derived in section 2.2.

$$\begin{aligned}\frac{q^2}{a^2}\phi + 3H(\dot{\phi} + H\psi) &= -4\pi G\rho_0\delta \\ \frac{1}{a^2}q^2(\psi + \phi) &= -12\pi G(\rho_0 + P_0)\sigma\end{aligned}$$

One alternative is the  $\delta T_i^0$  component. This terms results in

$$q^2\left(\frac{1}{a^2}\dot{\phi} + H\psi\right) = 4\pi G(\rho_0 + P_0)\theta.$$

The equation for  $\sigma$  resulted from examining the off-diagonal terms of the stress energy tensor. The last alternative can be derived from the diagonal components and relating them to  $\delta P$ . However, the resultant equation is much more complicated than the other three. Also,  $\rho_0$ ,  $P_0$ ,  $\delta$ ,  $\sigma$ , and  $\theta$  are summed over all the species. CDM follows the fluid equations:

$$\begin{aligned}\dot{\delta}_c + \frac{1}{a}\theta_c &= -3\dot{\phi} \\ \dot{\theta}_c + H\theta_c &= \frac{1}{a}q^2\psi.\end{aligned}$$

The Baryons follow the fluid equations also but contain corrections for Compton scattering. We do not explicitly state these terms since our focus is on

neutrinos. The details of the collision terms can be found in [20] with a textbook explanation in [24]. We denote the collision terms simply as  $C$ .

$$\begin{aligned}\dot{\delta}_b + \frac{1}{a}\theta_b &= -3\dot{\phi} \\ \dot{\theta}_b + H\theta_b &= \frac{1}{a}q^2\psi + C\end{aligned}$$

Photons follow the radiation hierarchy with additional collision terms. The first two moments are:

$$\begin{aligned}\dot{\delta}_\gamma + \frac{4}{3}\frac{1}{a}\theta_\gamma &= -4\dot{\phi} \\ \dot{\theta}_\gamma &= \frac{1}{a}q^2\left(\frac{1}{4}\delta_\gamma - \sigma + \psi\right) + C.\end{aligned}$$

The higher moments can be found with the recursion relation from section 2.3.3.

$$\dot{\Psi}_\ell = -\frac{1}{a}\frac{q}{2\ell+1}\left[(\ell+1)\Psi_{\ell+1} - \ell\Psi_{\ell-1}\right]$$

For  $\dot{\sigma}$ , one may integrate the above equation by  $g \int \frac{d^3p}{(2\pi)^3} \frac{p^2 \hat{p}^i \hat{p}^j}{E} f_0$  and then apply the traceless projection operator. The collision terms are dependent on the polarization. Therefore, additional equations are needed to track the photon polarization. We do not present them here. The massless neutrinos follow the same radiation hierarchy with no scattering. The equations for massive neutrinos are discussed in section 4.1.

At this point, one may wonder why we bother with the radiation hierarchy. After all, equation 2.17 can be integrated like any other differential equation. Instead of an infinite set of equations, we would have one equation per species and the two equations for the metric perturbations. However,

there are a few reasons one needs the radiation hierarchy. First, the radiation hierarchy bypasses the need to know  $\mu$ . This dependence is removed once we take the Legendre moment. Second, the collision terms, which we have not expressed in full detail, depend on specific Legendre moments, namely the zero moment of the photon and the first moment of the baryons. Lastly, we need to integrate  $\Psi$  to obtain  $\delta$ ,  $\theta$ ,  $\sigma$ , and  $\delta P$  which are required for metric perturbations. These integrals are not analytic in general and maybe cumbersome since they are over the entire phase-space. The relationships in the radiation hierarchy are much simpler. The price of simplicity is the radiation hierarchy.

Full, detailed treatments are shown in various works. Lifshitz[25] was the first to develop the perturbations around the FRW metric. Peebles and Yu[26] applied these perturbations to photons and baryons. Bond and Efstathiou[27] included the effects of dark matter. A useful reference is Ma and Bertschinger[20] which develops the perturbation in 2 common gauges. Seljak and Zaldarriaga [28] develop the line of sight integration. For textbook treatments see Weinberg[22], Dodelson[24], Padmanabhan[29], Kolb and Turner [30], and Mukhanov[31].

# Chapter 3

## Observables

This section summarizes how we compare theory to observations. For this work, we discuss the measurements of the Hubble constant, CMB angular spectrum, and LSS power spectrum. The Hubble constant,  $H_0$ , is a geometric probe and intrinsically tied to the concept of a standard candle. Therefore, we begin by discussing the use of supernovae type Ia as standard candles as an example before we discuss the measurements of  $H_0$ . Geometric probes are insensitive to the perturbations. The CMB and LSS, on the other hand, give us observables of the perturbations today. The power of the CMB is its linearity and its ability to be directly measured. Precision cosmology is dependent on the precision of CMB measurements. The tightest bounds on model parameters stem from CMB measurements such as the WMAP satellite. For massive neutrinos, the largest effects occur in the matter power spectrum. However, except for weak lensing, we can not observe this directly since we can not measure the distribution of CDM directly. We do not utilize current weak lensing observations because of their current precision level and error analysis. Galaxy power spectrum are currently used to estimate the matter spectrum. However, it appears that normal matter only accounts for a fifth of the total matter in the universe. Also, the distribution of Baryons is not expected

to coincide with CDM. The final subsection discusses these difficulties and how cosmologists connect galaxy surveys with the theoretical matter power spectrum.

### 3.1 Geometric Probes

Geometric probes measure the expansion rate of the universe, namely the behavior of  $a(t)$ . They have the advantage of being independent of the perturbations, and therefore, they can also probe cosmology outside of any particular gravitational theory. Measuring the expansion of the universe over time is akin to basic kinematics. We want to know the velocity of the expansion over time. From this, we can easily tell if something is speeding up or slowing down. Alternatively, we could look at velocity versus distance. In cosmology, velocity is measured by the Doppler shift or red shift. The term red shift is used because our observations show that the vast majority of galaxies are moving away from us, starting with Hubble's observations[32]. This is the qualitative evidence that the universe is expanding. What we want is quantitative information. However, both time and distance are extremely difficult to measure directly at large scales.

For objects near the Earth, direct distance measurements are made through radar. For measurements within our galaxy, parallax measurements are done. Parallax measurements are based on geometry: when an observer shifts in position, the apparent velocity of an object decreases with distance. The *Hipparcos* satellite[33] measured the distances of 100,000 stars up to 1

kpc using parallax. A successor mission, the *Gaia* satellite, is scheduled to launch in 2012. However, the apparent motion for objects beyond these scales is too small to observe. Standard candles are used to determine distances at most astrophysical and cosmological scales. Standard candle measurements are based on the difference between the observed and intrinsic brightness of an object. However, these are not direct measurements and require calibrations with direct measurements. The next subsections discuss standard candles and a few important examples. Afterward, we discuss how measurements on the Hubble constant are made with the concept of cosmic distance ladders. An alternative measurement is based on standard rulers. Rather than measuring the brightness of an object, we could compare the observed size of an object versus its intrinsic size. The crucial point is to find an object with a uniform or known intrinsic size. Recently, the Baryon acoustic oscillation(BAO) has come up as a potential standard ruler. However, the BAO measurements are a subset of the galaxy power spectrum so we do not delve into the details. See [34] for a review.

### 3.1.1 Standard Candles

Standard candles work in the following way. We have some object with some intrinsic luminosity  $L$ .  $L$  is the amount of radiative energy the object emits per time; it is a measurement of power. We call it intrinsic to distinguish from our observed luminosity  $\ell$  which is the power per square area.  $\ell$  is the flux of radiative energy. To begin with, we assume a Euclidean space. If some

object has luminosity  $L$  and we observed it from distance  $d$  away, we should measure  $\ell$  to be

$$\ell = \frac{L}{4\pi d^2}. \quad (3.1)$$

This assumes that energy is emitted isotropically. From equation 3.1, we see that a measurement of flux is a measurement of  $d$  if  $L$  is known. As long as we can find some object where  $L$  is constant regardless of its environment, then this kind of inference works.

However, we do not live in a Euclidean space; in cosmology we need to work with the FRW metric. The  $\ell$  we observe should be

$$\ell = \frac{L|_{t_o}}{4\pi d^2(t_o)} = \frac{\frac{dE}{dt}|_{t_o}}{4\pi d^2(t_o)}$$

where  $t_0$  denotes the time today. For this section on geometric probes, it is useful to write out  $a(t_0)$  even though we have set it equal to 1 in prior sections. There are 3 main modifications we need to make to equation 3.1.

1. The surface area of the sphere surrounding the emitting object is larger when we observed it than when the light was emitted. We have moved further away from it. Therefore the surface area should be

$$4\pi d^2 \rightarrow 4\pi a(t_o)^2 x^2$$

where  $x$  is the co-moving coordinate distance between us and the object. Note that we assume that  $x$  is roughly constant. In other words we assume that we are not moving with respect to the object except for the universal expansion (no peculiar velocity).

2. The energy of the radiation decreases as the universe expands. A simplistic way to view this is that the wavelength stretches as the space stretches. Therefore frequency decreases. This energy decrease is proportional to  $a^{-1}$ . The same result can be more formally shown by examining the geodesic equation 2.8. The energy of the radiation we observe is less than its original energy at the time of emission. Since energy is inversely proportional to  $a$ , we have

$$E(t_o) = \frac{a(t_e)}{a(t_o)} E(t_e)$$

where  $t_e$  refers to the time of emission. The luminosity term in equation 3.1 is modified by

$$L \rightarrow L \frac{a(t_e)}{a(t_o)}$$

where  $L$  is now the original luminosity at time of emission.

3. The last modification deals with time. This is why the luminosity in equation 3.1 was written as a derivative of energy over time. The power or rate at which we observe the energy differs than when it was originally emitted. One way to see this is to look at the metric. For light  $ds = 0$ . Therefore,  $dt = adx$ . We can see that  $dt$  is smaller in the past. Therefore, the observed luminosity or rate is slower than when it was originally emitted.

$$\frac{dt|_{t_o}}{dt|_{t_e}} = \frac{a(t_o)}{a(t_e)}$$

The luminosity needs to be shifted lower.

$$L \rightarrow L \frac{a(t_e)}{a(t_o)}$$

Before we add all these modifications, note that

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

where  $z$  is the redshift. Finally, our modified flux or observed luminosity is

$$\ell = \frac{L \left( \frac{a(t_e)}{a(t_o)} \right)^2}{4\pi a(t_o)^2 x^2} = \frac{L}{4\pi a(t_o)^2 (1+z)^2 x^2} = \frac{L}{4\pi d_L^2} \quad (3.2)$$

where we have defined the luminosity distance:

$$d_L^2 = a^2(t_o)(1+z)^2 x^2. \quad (3.3)$$

The luminosity distance is something we can observe by measuring the observed luminosity or flux from an object with known intrinsic luminosity. Again, the important thing is to find something close to a standard candle such that  $L$  is consistent. Then, observations of  $\ell$  are identical to observations of  $d_L$ .  $d_L$  is also a value that can be calculated from theory easily. It is just a matter of calculating the co-moving distance  $x$

$$x = \int dx = \int \frac{dt}{a} = \int \frac{da}{a \dot{a}} = \int \frac{da}{a^2 H(a)}$$

for a given redshift.  $H(a)$  is just the Hubble parameter that can be calculated by knowing the energy density from equation 2.1. In summary, equation 3.2 defines the observation, and equation 3.3 defines the theory. We can plot the observed  $d_L$  against  $z$  and fit the theory.

One systematic effect that all standard candles must deal with is dust reddening or extinction. As light propagates through the universe from the

source to us, it can travel through non-vacuum mediums. This medium is usually dust of the interstellar medium which causes a reduction in brightness. Corrections to this extinction involve either the use of multi-wavelength observations or the use of small frequency observations. Reddening is empirically known to affect the longest wavelengths the least[35][36].

### 3.1.2 Supernovae Type Ia

The use of supernovae as standard candles is an important cosmological observable. Riess et. al.[37] and Perlmutter et. al.[38] used supernovae data to show that the universe was accelerating. Supernovae allow us to measure the expansion of the universe to high redshifts. In particular, supernovae type Ia have roughly constant luminosity. Although Sne Ia are not as common as other standard candles, their brightness makes it possible to observe them even to redshifts beyond  $z = 1$ . For this reason, Sne Ia have been a pivotal piece of information.

In general, supernovae are categorized as type I or II based on their spectroscopy. Type I have no hydrogen lines in their optical spectra whereas Sne II do. Sne I are further categorized into subtypes a, b, and c. Sne Ia have a strong absorption line near 6150 Angstroms. This line is believed to come from Si II. Sne Ib have no Si II line but a strong He I line. Sne Ic have neither absorption lines[39]. Sne Ia are believed to come from C-O white dwarf stars accruing mass from a companion star. This model is opposed to the model for type II which are core collapses of massive supergiant stars. In

the Sne Ia model, the white dwarf is in a binary system with a companion star. White dwarfs are the stable final state of main sequence stars. However, as the white dwarf collects mass, it ultimately reaches a critical stage in which the electron degeneracy pressure can not maintain the additional mass[40][41]. This limit is the Chandrasekhar limit[42] and is about 1.4 solar masses. Sne Ia are powered by the decay of  $\text{Ni}^{56}$  for the early peak brightness and  $\text{Co}^{56}$  after a few weeks[43]. Uncertainties of this model still exist, although there is a consensus on the general framework. One example of an observed Sne Ia which can not be explained is the so called “champagne” supernova (SNLS-03D3bb)[44]. Investigations into the ejecta spectrum such as in [45] show that we do not fully understand the exact ignition mechanisms.

It turns out that Sne Ia are close but not exactly standard candles. However, Phillips[46] found an empirical correlation between the peak brightness and the rate of luminosity decline. With this correction, Sne Ia are almost perfect standard candles. The methods to make Sne Ia standard candles are called light curve fitting. Two leading methods are SALT[47] and MLCS2k2[48]. In addition, the Carnegie Supernova Project has introduced a new method for their purposes[49]. These methods are based on empirically fitting the brightness profiles with multiple wavelength observations. It has been observed that for higher frequencies, larger corrections are needed. In fact, [50] show that in the near infrared, the absolute magnitude of luminosity has the same spread as the corrected optical light curves. Enough Sne type Ia have been observed such that statistical errors are smaller than the systematic errors. These sys-

tematic errors stem from the differences in the light curve fitting and handling of reddening effects[51]. A review of Sne Ia and its general application to cosmology can be found by Filippenko[52]. The strength of Sne Ia as standard candles is not only the large magnitude of brightness but also the small spread in the apparent brightness. In other words measurements of  $\ell$  are done with great precision. However, the intrinsic brightness is still a source of systematic error. In fact, the intrinsic brightness is usually calibrated to another standard candle: cepheids.

### 3.1.3 Cepheids

Cepheids are a class of pulsating variable star. Within this class are population I and II stars. Cepheids are distinguished from other pulsating variable star by their short periods, usually around days to months. A similar type of star, RR Lyrae variables, are sometimes referred to as cluster cepheids. However, the current consensus is that RR Lyrae are a distinct class from cepheids due to a few factors: spectral class, population II stars, shorter period ( $< 1$  day), and smaller magnitude[53]. Population I Cepheids are younger stars and are more massive than Population II. Both are brighter than RR Lyrae. The mechanism behind the pulsating stems from dynamics of helium ionization. As helium is heated, singly ionized atoms can be doubly ionized. These double ionized helium allow more radiation to escape and cooling resumes. This mechanism causes the periodic expansion and contraction of the star. The model is sometimes referred to as the Eddington valve[54].

For a reference, see Percy[55].

Cepheids are standard candles based on the pulsating period-luminosity relation (P-L). The longer the period, the brighter the star. This relation was discovered by Leavitt[56][57] and is sometimes referred to as Leavitt's law. In Sne Ia, the physics of the supernova explosion was the source the standard brightness. In the case of Cepheids, the absolute luminosity is not constant but is known empirically.

Just as Sne Ia are not perfect candles, Cepheids require corrections beyond reddening effects. The main systematic issue with Cepheids is their metallicity. Improper or lack of classification of Cepheids within their subclasses caused large discrepancies in reported  $H_0$ .

### 3.1.4 Determining the Hubble Constant

The measurement of  $H_0$  is intrinsically tied to the concept of cosmic distance ladders. Standard candles are necessary to measure large distances. However, they are not direct measurements and there will always be some uncertainty to their intrinsic brightness. Therefore, the intrinsic brightness is calibrated with other distance measures. This is the concept of a cosmic distance ladder. The first rung of the ladder is the calibration of standard candles of a few kpc with some direct measurements such as parallax. Therefore, the first standard candles usually need to be part of our local group. Standard candles themselves usually have a range mostly dependent on how bright they are. Standard candles based on stars like cepheids or RR Lyrae

are not bright enough to be seen beyond the 5 – 10 Mpc scale i.e. the closest galaxies. This rung calibrates the standard candles for beyond the 10 Mpc scale such as Sne Ia and velocity-size relationships in elliptical galaxies. Each new standard candle in the next rung is calibrated with the previous rung of the ladder. Historically, there have been large discrepancies of the measured  $H_0$  based on systematic errors stemming from absolute distance calibration. Naturally, a systematic error on the first rung of the ladder propagates to the higher distance measures.

The dynamics of the recent universe expansion simplifies greatly. In fact it is model independent. The proper distance is the comoving distance multiplied by the scale factor.

$$s = a(t)x$$

The derivative of  $s$  has two components: one due to the expansion of the universe and one due to the motion in comoving coordinates. These two terms are sometimes referred to as the Hubble flow and the peculiar velocity, respectively. If we assume the peculiar velocity is small compared to the Hubble flow, then the velocity  $v = ds/dt$  is the Hubble law[32].

$$v = H(t)s, \quad \frac{\dot{x}}{x} \ll H(t) \tag{3.4}$$

We have left the Hubble rate to be a function of  $t$  as it generally is. For the Hubble constant, we want to observe  $v$  and  $s$  closer to our present time i.e. at small redshifts. Taylor expansion of the scale factor results in

$$a(t) \approx a(t_0)(1 + H_0(t - t_0)).$$

Since  $1/(1+z) = a(t)/a(t_0)$ , the above equation becomes

$$z \approx H_0(t_0 - t) = H_0 d \quad (3.5)$$

with  $c = 1$ . Note that  $H_0$  is a constant number which can be measured independent of any particular cosmological theory. Therefore, these measurements of  $H_0$  are direct. Equation 3.5 did not use equation 3.4 explicitly. However, equation 3.5 does assume that the redshift is completely due to the Hubble flow. We presented equation 3.4 to illustrate that this is not always true. In fact for objects very close to us, the local expansion rate is small compared to the peculiar velocities. The trick to measuring the Hubble constant is to find a large amount of standard candles within a perfect range. If they are too far away, then equation 3.5 is no longer valid. If they are too close, then  $v$  and  $z$  are dominated by the peculiar velocities. It turns out that enough objects have been observed such that the statistical errors associated with this limited range are growing smaller with respect to the systematic errors of distance calibration.

For our purposes, we discuss one particular ladder using cepheids and Sne Ia. This ladder was used by the Hubble space telescope key project. In this ladder, cepheids are calibrated using the distance to the Large Magellanic cloud (LMC). The cepheids were then used to calibrate Sne Ia and other distance measures such as the Tully-Fisher relation[58] and surface-brightness fluctuations[59] to push the observed distances. At the time, the largest uncertainties stemmed from the absolute distance to LMC and the effect of metallicity on

the cepheid P-L relation. The resultant measure of the Hubble constant was  $72 \pm 3(\text{stat}) \pm 7(\text{sys})$ [60]. This result is obtained by weighting the several distance ladders used; all of which were calibrated by cepheids. Since the HST key project measurement, other groups have attempted to reduce the errors with the same type of distance ladder. [61] calibrate the cepheids using NGC 4258 rather than the LMC. While NGC 4258 is further away, its distance measure seems more stable and precise. These measures come from the motion of its water masers[62]. The new calibration leads to  $H_0 = 74.2 \pm 3.6$ . [63] used the same distance ladder but with a new calibration of the cepheid absolute magnitude with galactic parallax[64] and additional supernovae data[65]. They found an  $H_0 = 73 \pm 2(\text{stat}) \pm 4(\text{sys})$ . In the future, errors on the  $H_0$  could be as low as 2% with additional data and reduction of uncertainties[63].

Other types of cosmic distance ladders use RR Lyrae stars which exist in older population II containing galaxies. These stars can be used instead of cepheids for the first rung of the ladder. They are more abundant but less luminous than cepheids[66]. RR Lyrae can then calibrate the next rung in the ladder such as the use of elliptical galaxy diameter-velocity dispersion relations[67]. There are also indirect measurements of  $H_0$ . These include the use of x-ray data[68] and gravitational lensing data[69].

### 3.2 Cosmic Microwave Background

The CMB is a result of the acoustic oscillations in the primordial plasma. In the early universe, baryons and photons are tightly coupled due

to Compton scattering. The CDM drag the baryons into their gravitational potential. Thus, the photons also feel this gravitational effect due to the tight coupling. In turn, the photon's radiative pressure pulls the baryons-photons out of the potential. These opposing forces cause the oscillation. This behavior continues until the Hubble rate is large enough such that Compton scattering is not fast enough to keep photons and baryons together. This event is determined by recombination physics. Since the early universe is very dense, neutral atoms can not form; they are continuously ionized by the background photons. However, as the universe expands, the density diminishes as seen in section 2.1. At some point, the electrons can combine to form neutral atoms. The photons that were once coupled now freely propagate through the universe. The small perturbations in the uniformity of photon energy are indicative of the gravitation perturbations at this time of *last scattering*.

Our observations of the CMB are a snap shot of the gravitational perturbations at this early stage of the universe, roughly 400,000 years after the big bang. The measured CMB is dominated by these early universe physics; these anisotropies are called primary. Secondary anisotropies exist due to more recent events. The reionization of the recent universe starts with the formation of the first stars and galaxies. The radiation from these objects ionizes the neutral medium which in turn scatters with the CMB. Reionization dampens the level of anisotropies at smaller scales and polarizes the CMB. Another effect is the Sunyaev-Zeldovich effect due to inverse Compton scattering. The CMB photons scatter off the hot electrons surrounding galaxy clusters.

The Sunyaev-Zeldovich effect occurs at even smaller scales than reionization since it is tied to galaxy clusters. This effect can be further divided into two categories: kinetic and thermal. Lastly, an effect known as the integrated Sachs-Wolfe (ISW) effect alters the anisotropies at large scales. ISW is due to a non-constant gravitational potential. It is seen in equation 2.14 and is due to photons climbing out of potential well. This causes a red or blue shift. However, if the gravitational potential is constant, then this effect is non-existent. This is the case where there is no Dark Energy.

The CMB was first predicted as a consequence of the Big Bang by Gamow[70][71] and Alpher & Herman[72]. The first detection was made by Penzias and Wilson[73]. Their work resulted in half of the Nobel prize in 1978. The NASA probe COBE verified the black-body nature of the CMB and made the first detection of anisotropies[23][74]. This work resulted in the Nobel Prize for Smoot and Mather in 2006. The successor of COBE, WMAP, observed the anisotropies up to the first 2 acoustic peaks[75]. It helped usher in the era of precision cosmology. An important measurement is the level of the perturbations. COBE found temperature differences only at a level of 0.001%[74]. Therefore, we can assume that the gravitational perturbations were small, verifying the validity of linear perturbation theory. Because of the linearity of the theory, high precision inference can be obtained by analyzing the CMB. In this section, we utilize several statistical properties without proof. Appendix A lists these theorems.

For this section, we use the following notation.

- $\Delta(x^\mu, \hat{n})$  is the theoretical temperature anisotropy observed at  $x^\mu$  in direction  $\hat{n}$  where  $T(x^\mu, \hat{n}) = T(1 + \Delta(x^\mu, \hat{n}))$ .
- $a_\ell^m(x^\mu)$  are the coefficients of  $\Delta$  in the spherical harmonics expansion. They depend on  $x^\mu$ .
- $A_\ell^m$  are the coefficients of the observed anisotropies in the spherical harmonics expansion. They do NOT depend on  $x^\mu$ .
- $C_\ell$  are the theoretical variances for the distribution of  $a_\ell^m(x^\mu)$ .
- $C_\ell^{\text{sky}}$  is the unbiased estimator of  $C_\ell$  based on our observations.
- $S_i(x^\mu)$  is the theoretical temperature anisotropy at pixel  $i$ . It is dependent on the theoretical  $\Delta$  and the instrument beam design.
- $\Delta_i$  is the observed temperature anisotropy at pixel  $i$ .
- $C_{S,ij}$  is the covariance of  $S$  between pixel  $i$  and pixel  $j$ .

### 3.2.1 Theoretical Angular Power Spectrum

First, what does theory predict? Theory does not predict an exact  $\Delta(x^\mu, \hat{n})$ . That would be quite bold. What theory does predict is a probability distribution for  $\Delta$ . First, we expand  $\Delta$  into the spherical harmonics.

$$\begin{aligned}\Delta(x^\mu, \hat{n}) &= \sum_{\ell, m} a_\ell^m(x^\mu) Y_\ell^m(\hat{n}) \\ a_\ell^m(x^\mu) &= \int d\Omega \Delta(x^\mu) Y_\ell^m(\hat{n})\end{aligned}\tag{3.6}$$

The simplest inflation theories predict that the coefficients will follow a normal distribution

$$a_\ell^m \sim \text{norm}(0, C_\ell)$$

or, in other words

$$\begin{aligned} \langle a_\ell^m \rangle &= E[a_\ell^m] = 0 \\ \langle a_\ell^{m'} * a_\ell^m \rangle &= \text{Var}[a_\ell^m] = C_\ell \delta_{\ell'\ell} \delta_{mm'}. \end{aligned} \quad (3.7)$$

The averaging is done over the entire support/domain of  $a_\ell^m$  which is the entire space of the universe. Note that since the  $a_\ell^m$ 's are normal distributed then  $\Delta$  is also normally distributed with mean 0 and variance =  $\sum |Y_\ell^m|^2 C_\ell$ .

The  $C_\ell$ 's are found by theory. We have a set of differential equations modeling the physics. The physics here being the evolution of perturbations to the FRW metric, perturbations to the homogeneous phase space distribution of the matter/energy components, and the interactions. We integrate these differential equations until today. What we have is an initial value problem. Chapter 2 discussed only the differential equations, not the initial conditions. The initial conditions are stochastic in nature and come from models of inflation. Inflation is a theoretical period of accelerated expansion in the very beginning of the universe. Inflation is a desired addition to current cosmological models for several reasons: it explains the homogeneity of the observed CMB, it explains the observed flatness of our universe without fine-tuning, and it provides an explanation to the origin of the perturbations. For an introductory treatment of inflation, see [76]. For this work, we assume adiabatic initial

conditions. This condition is based on the simplest of single-field inflation models and seems the best fit to current observations.

We now begin the mathematical formalism for calculating  $C_\ell$ . From equations 3.6 and 3.7,

$$C_\ell = \int d\Omega d\Omega' Y_\ell^m(\hat{n}') Y_\ell^m(\hat{n}) \langle \Delta(x^\mu, \hat{n}') \Delta(x^\mu, \hat{n}) \rangle \quad (3.8)$$

In section 2.3.1, the dynamical equation for  $\Delta$  was in the Fourier space. Therefore, we transform the above equation. The dynamical equations in section 2.3.1 were completely deterministic. To take into account the stochastic initial conditions, we let the Fourier transformed  $\Delta$  to be

$$\Delta(t, \vec{q}, \hat{p}) \rightarrow \mathcal{R}(\hat{q}) \Delta(t, q, \mu). \quad (3.9)$$

The initial conditions are encoded in  $\mathcal{R}$  which is a random variable and only depends on the direction of  $\vec{q}$ . The  $\Delta$  on the right hand side of equation 3.9 is still a completely deterministic variable. It depends only on the magnitude of  $\vec{q}$  and the inner product  $\mu \equiv \hat{q} \cdot \hat{p}$ . The initial conditions state that  $\mathcal{R}$  comes from a standard Gaussian distribution with variance

$$\langle \mathcal{R}^*(\vec{q}') \mathcal{R}(\vec{q}) \rangle \propto \delta^3(\vec{q}' - \vec{q}) P_p(q)$$

where  $P_p(q)$  is the primordial power spectrum. In the simplest model, the primordial power spectrum is parameterized as

$$P_p(q) = A_s \left( \frac{q}{q_*} \right)^{1-n_s}.$$

There are two parameters:  $A_s$  is the amplitude and  $n_s$  is the spectral index or tilt.  $q_*$  is an arbitrary value. These two parameters encode the initial conditions or the physics of inflation.

Returning to equation 3.8, we now work in the Fourier space and with the parameterization of the initial conditions in equation 3.9. The equation for the  $C_\ell$ 's becomes

$$C_\ell \propto \int d^3q d\Omega d\Omega' Y_\ell^m(\hat{n}') Y_\ell^m(\hat{n}) P_p(q) |\Delta(t, q, \mu)|^2$$

To solve for  $C_\ell$ , we need to integrate the differential equation of  $\Delta$  to the present. We saw in section 2.3.3 that solving for  $\Delta$  would require the integration of the radiation hierarchy. Therefore, we take the inverse Legendre moment of  $\Delta$ . The inverse moment is

$$\Delta(t, q, \mu) = \sum_\ell (2\ell + 1) P_\ell(\mu) \Delta_\ell(t, q).$$

Further reduction can be made by taking the properties of Legendre polynomials and spherical harmonics. The final result is

$$C_\ell \propto \int_0^\infty dq q^2 P_p(q) |\Delta_\ell(t, q)|^2 \quad (3.10)$$

where  $t$  should be today if we want to compare with observation. The steps to calculate the theoretical  $C_\ell$  are now clear. We integrate  $\Delta_\ell$  based on the radiation hierarchy equation in section 2.3.3 along with the other differential equations. These need to be done for an appropriate number of  $q$ 's such that the integral in equation 3.10 can be done. The actual number of  $q$ 's and for that matter  $\ell$  depends on the desired accuracy as will be discussed in chapter 5.

### 3.2.2 Estimation of the Angular Power Spectrum

From theory, the angular power spectrum is based on the variance of some distribution. First, we look at a simplistic statistical example which outlines the basic concepts. Say we have made  $n$  measurements:  $X_1, X_2, \dots, X_n$ . Now assume that these measurements were made from a normal/Gaussian distribution denoted  $X_i \sim \text{norm}(\mu, \sigma^2)$  where  $\mu$  and  $\sigma^2$  are the population mean and variance. The unbiased estimators for  $\mu$  and  $\sigma^2$  are the sample mean and the sample standard deviation.

$$\hat{\mu} = \bar{x}$$

$$\hat{\sigma}^2 = s^2 = \frac{\sum (x_i - \bar{x})^2}{n - 1}$$

We divide by  $n - 1$  because we have to estimate  $\mu$ . In a case closer to the CMB, we let  $\mu$  be known and estimate  $\sigma^2$ . Without loss of generality, we set  $\mu = 0$ . In this case, the estimator of  $\sigma^2$  loses the  $n - 1$  term and becomes

$$\hat{\sigma}^2 = \frac{\sum x_i^2}{n}.$$

We now apply this example to the CMB. From theory and observation, we know that  $\Delta$  is very close to normal. Therefore, the coefficients,  $a_{\ell m}$ , also follow a normal distribution (see appendix).  $C_\ell$  is defined as the variance of the distribution of the  $a_\ell^m$ 's.

$$a_{\ell m} \sim \text{norm}(0, C_\ell)$$

Theory gives us  $C_\ell$ . Our estimate of this is then

$$\hat{C}_\ell = C_\ell^{\text{sky}} = \frac{\sum_m A_{\ell m}^2}{n}$$

In this case,  $n$  is known. For each  $\ell$ , there are  $2\ell + 1$   $m$ 's. This statement is just the property of the spherical harmonics. Therefore we have

$$C_\ell^{\text{sky}} = \frac{1}{2\ell + 1} \sum_m A_{\ell m}^2.$$

The only differences between the CMB and our simple example are the number of measurements and the variance. The  $N$  and variance in the CMB case is dependent on  $\ell$ . We can think of our measurements as coming from a mixture of normal distributions. One caveat is that the  $a_{\ell m}$ 's are complex numbers but this is a trivial change to the analysis.

We mentioned the cosmic variance before. The cosmic variance is just the variance of our estimator,  $C_\ell^{\text{sky}}$ . For the normal distribution, if both the mean and variance are unknown, we would use  $s^2$  to estimate  $\sigma^2$ . The variance of this estimator is known since

$$\frac{(n-1)s^2}{\sigma^2} \sim \chi_{n-1}^2.$$

The variance of a  $\chi^2$  distribution is two times the degrees of freedom; in this case, it is equal to  $n - 1$ . Therefore the variance of  $s^2$  is

$$\text{Var}[s^2] = \frac{2\sigma^4}{n-1}.$$

In our case, the mean is known so  $n - 1$  becomes  $n$ . Therefore the variance of our estimator is

$$\text{Var}[C_\ell^{\text{sky}}] = \frac{2C_\ell^2}{2\ell + 1}$$

This expression is the cosmic variance. Note that the higher  $\ell$  is, the smaller the variance is. We have a fundamental limit on the certainty of our estimate  $C_\ell^{\text{sky}}$ . At low  $\ell$  this cosmic variance becomes an issue.

### 3.2.3 'Theoretical' Pixelated Sky

The next step to realism is to know that we can not observe  $\Delta$  at an exact direction. There is some thickness to our instrumentation. What we want to know is the “theoretical” anisotropy,  $S_i(x^\mu)$ , at pixel  $i$ . We do not go into detail into how we determine pixel shape or distribution. It is sufficient to know that we must pixelize the sky and that the size is dependent on how good our instrumentation is.  $S_i$  depends both on the theoretical  $\Delta$ 's and on the instrument design or beam design. For neatness, we envelope the stochastic initial conditions into  $\Delta$ . Note that  $S_i(x^\mu)$  does depend on the position in space-time and is probabilistic.  $S_i$  is

$$S_i(x^\mu) = \int d\hat{n} \Delta(x^\mu) B_i(\hat{n})$$

where  $B_i$  is the beam profile for pixel  $i$  in direction  $\hat{n}$ . Since  $E[\Delta] = 0$ ,  $E[S_i] = 0$ , too. We define

$$C_{S,ij} \equiv \langle s_i s_j \rangle$$

as the covariance between pixel  $i$  and pixel  $j$ .

What we want is to relate the theoretical  $C_\ell$  to  $C_{S,ij}$ .

$$C_{S,ij} = \int d\Omega d\Omega' B_i(\hat{n}) B_j(\hat{n}') \langle \Delta(x^\mu, \hat{n}) \Delta(x^\mu, \hat{n}') \rangle$$

Note that the variance term is equal to  $\sum_{\ell,m} C_\ell Y_\ell^m(\hat{n}) Y_\ell^{m*}(\hat{n}')$  based on spherical harmonic decomposition. Also, a property of the spherical harmonics is that  $\sum_m Y_\ell^m(\hat{n}) Y_\ell^{m*}(\hat{n}') = \frac{(2\ell+1)}{4\pi} P_\ell(\hat{n} \cdot \hat{n}')$  where  $P_\ell$  are the Legendre polyno-

mials. Then

$$C_{S,ij} = \sum_{\ell} \frac{2\ell + 1}{4\pi} C_{\ell} \int d\hat{n}d\hat{n}' B_i(\hat{n}) B_j(\hat{n}') P_{\ell}(\hat{n} \cdot \hat{n}').$$

Note that the integral term has nothing to do with the theory. It depends only on the instrument's beam profile. This integral is termed the window function,  $W_{\ell,ij}$ . Then the relationships between the theoretical variance,  $C_{\ell}$ , and the pixel variance,  $C_{S,ij}$  is

$$C_{S,ij} = \sum_{\ell} \frac{2\ell + 1}{4\pi} C_{\ell} W_{\ell,ij}$$

From this section, we see that

$$\vec{S}(x^{\mu}) \sim \text{norm}(0, C_S)$$

where  $\vec{S}$  is the set of all the pixel anisotropies and  $C_S$  is the entire covariance matrix with elements defined by  $C_{S,ij}$ . The probability density function of  $\vec{S}$  is

$$f(\vec{S}|B, C_{\ell}) = \left( \frac{1}{2\pi \det C_S} \right)^{M/2} e^{-\frac{1}{2} \vec{S}^T C_S^{-1} \vec{S}} \quad (3.11)$$

where M is the number of pixels.  $B$  and  $C_{\ell}$  are not explicitly in the PDF but  $C_S$  depends on these two.

### 3.2.4 Observed Anisotropy and Noise

In the previous section, we related  $C_{\ell}$ 's to  $C_{S,ij}$  by taken into account the finite shape of our instrumental beam. However, this is still theoretical in the sense that  $S_i$  is some random variable whose distribution is related to the

true  $\Delta$  distribution. In this section, we want to relate our observed anisotropy,  $\Delta_i$ , to  $S_i(x^\mu)$ . In other words, we want to estimate  $C_{S,ij}$  using our observed anisotropy,  $\Delta_i$ . In the language of Bayesian statistics, we want to know the posterior distribution of  $C_{S,ij}$  conditioned on the data  $\Delta_i$

$$f(C_{S,ij}|\Delta_i) \propto f(\Delta_i|C_{S,ij})$$

where we have assumed no prior information on  $C_S$ . First we need to work out the above equation by including noise and equation 3.11. Then, we can find the form of the posterior distribution.

The noise is modeled as a Gaussian distribution. In other words

$$\Delta_i \sim \text{norm}(S_i, C_N)$$

where  $C_N$  contains all the information about the noise. For our purposes, we assume that  $C_N$  is known. Equation 3.11 shows that  $S_i$  comes from a Gaussian distribution as well.

$$S_i \sim \text{norm}(0, C_{S,ij})$$

In statistics, this is a specific case of a hierarchy model or mixture model. We refer the reader to chapter 4.4 in [77] for a statistical treatment of hierarchy models. A statistical hierarchy model is when the data is taken from a distribution whose parameters also come from a distribution. The hierarchy model in the CMB case is as so.

- Data:  $\Delta_i$  from distribution  $\text{norm}(S_i, C_N)$

- Parameter:  $S_i$  from distribution  $\text{norm}(0, C_{S,ij})$
- Hyper-parameter:  $C_{S,ij}$  assumed to be constant

We assume a non-informative hyper-prior on  $C_{S,ij}$ . In general, Bayesian inference looks at the joint posterior.

$$f(S_i, C_{S,ij} | \Delta_i, C_N)$$

However we are only concerned with the marginal posterior distribution of  $C_{S,ij}$ .

$$f(C_{S,ij} | \Delta_i, C_N)$$

As shown in equation 2.9 and 2.10 in [78], the marginalization of two normals also results in a normal distribution. Based on equation 5.18 in [78], the resulting distribution is

$$f(C_{S,ij} | \Delta_i, C_N) \propto \frac{1}{\sqrt{\text{Det}(C_S + C_N)}} e^{-\frac{1}{2} \Delta_i (C_S + C_N)^{-1} \Delta_i} \quad (3.12)$$

In other words, the distribution of  $C_S$  has the same functional form as  $\Delta_i \sim \text{norm}(0, C)$  where  $C = C_S + C_N$ . In this discussion, we have assumed a known  $C_N$ .  $C_N$  comes from the mapmaking process which we do not discuss. This process takes the raw or time-ordered-data and assembles the sky map. We refer readers to [79] for details.

### 3.2.5 Error Analysis and Forecasting

Equation 3.12 shows the posterior distribution of  $C_S$  which is used to estimate  $C_\ell$ . With the inclusion of noise in the analysis, we can investigate

the error in the estimated  $C_\ell$  for current and future probes. To this end, we build on the simple example in section 3.2.2. The variance of  $C_\ell^{\text{sky}}$  is known since it comes from a Gaussian distribution. Based on equation 3.12, we infer that  $\Delta_i|C \sim \text{norm}(0, C)$ . Therefore, the variance of our estimation of  $C_S$  is

$$\frac{2}{N} (C_S + C_N)^2 \quad (3.13)$$

To relate this to the  $C_\ell$ , we note that the window function in section 3.2.3 is orthogonal. Specifically,

$$\int_{-1}^1 d\mu W_{\ell'} W_\ell = \delta_{\ell'\ell} \frac{2}{2\ell + 1}.$$

Integrating equation 3.13 by  $4\pi \int_{-1}^1 (d\mu/2) W_\ell$  and taking the square root results in the error of our estimate of  $C_\ell$ .

$$\sqrt{\frac{2}{N}} \left( C_\ell + 4\pi \int \frac{d\mu}{2} W_\ell C_N \right)$$

For simplicity, we assume a Gaussian beam profile and constant white noise.

This results in

$$\sqrt{\frac{2}{(2\ell + 1)f_{\text{sky}}}} (C_\ell + w e^{\ell(\ell+1)\sigma}). \quad (3.14)$$

$N$  becomes  $(2\ell + 1)f_{\text{sky}}$  since there are  $2\ell + 1$   $m$ 's for each  $\ell$ . The  $f_{\text{sky}}$  part comes from the reality that some of the sky is too contaminated to be analyzed. This contamination comes from foreground signals dominated by our galaxy.  $\sigma$  defines the beam size, and  $w$  defines the noise per pixel. Our treatment of window functions and beam profiles is simplistic. For an example of the calculation of realistic window functions, see [80].

Note that to find specific error values demands a specific value of  $C_\ell$ . We must know  $C_\ell$  to know the error in our estimate of it. In other words, these forecasted error bars not only depend on the chosen model but on the chosen parameter values. Therefore, one must take the specific values from forecasting analysis with caution. The analysis presented here is somewhat different in literature. Here, we have based most of our analysis on Bayesian analysis and general statistical analysis in hierarchy models. In literature, forecasting analysis of the CMB has been done by Fisher matrix analysis[81][82] and Monte Carlo simulation[83][1]. The Fisher matrix analysis takes the curvature of the likelihood function which in this case would be equivalent to equation 3.12. The inverse of the Fisher matrix is related to the error. See page 337-338 in [77] for more details. Alternatively, Monte Carlo simulations of the intrinsic variance and noise can be performed. These forecasting methods can be expanded to place bounds on the model parameters themselves.

### 3.3 Large Scale Structure

While the CMB measures the perturbations in the photon distribution, LSS measures the perturbations in the matter distribution. The CMB photons after recombination freely propagate towards us today. In the case of matter, perturbations to their distribution grow over time. As the matter gets clumpier and clumpier, stars and galaxies begin to form. This non-linear growth, among other factors, means that the linear perturbation model is not accurate at small scales. Where the precision of CMB inferencing is limited by the cosmic

variance, the precision of LSS inferencing is limited by our calculations of theory. The linear matter power spectrum is limited not only to the scales we can compare with observation but also on the exact observation we use. In the future, it might be possible to obtain a direct measure of the matter power spectrum through weak lensing surveys. However, at this time, these surveys do not have the precision necessary for neutrino inferencing. For a review of weak lensing, see [84]. Another probe is the Lyman alpha forest. This is the sum effect of light coming from distant objects such as Quasars and being absorbed by the intervening intergalactic medium. These absorption lines can be used to help constrain the amplitude and shape of the matter power spectrum and the Hubble rate at higher redshifts than galaxy surveys[85]. The most precise measurements of the matter power spectrum come from galaxy surveys. However, these galaxy surveys do not measure the matter power spectrum directly.

The first large survey was the Cfa survey done between 1977-1982[86]. The 2dFGRS in 1997 - 2002 had enough precision to detect the Baryon Acoustic Oscillation: the expected wiggles in the matter power spectrum due to the same physics as the CMB angular power spectrum[87]. The Sloan Digital Sky Survey began observing in 2000 and is currently in its third phase[88]. The SDSS is currently the largest survey. Future galaxy survey projects include the Large Synoptic Survey Telescope[89], the Dark Energy Survey[90], and HETDEX[91]. The BAO has garnered much attention for its ability to be a standard ruler. This standard ruler has been a target to gather more data

about dark energy. The information from the BAO is a subset of the total matter power spectrum. We choose to examine the entire power spectrum data and do not focus on the BAO.

### 3.3.1 Theoretical Matter Power Spectrum

The theoretical matter power spectrum is in some ways a simpler calculation than the CMB angular power spectrum. First, we do not work on a spherical domain and do not have to work in the spherical harmonic basis. Second, the evolution of matter perturbations is much simpler as seen in section 2.3.2. One simply needs to integrate the ODE equations up to some point (the effective redshift of a galaxy survey) and use the value of the matter perturbations to solve for the power spectrum. Unlike the CMB, we can observe structure at various positions so we are no longer constrained by the cosmic variance.

We can follow similar steps to find the matter power spectrum as we did for the CMB. The matter power spectrum is

$$P(q) \propto \delta^3(\vec{q}' - \vec{q}) \langle \delta(\vec{q}') \delta(\vec{q}) \rangle .$$

Note that  $\delta$  is already written in the Fourier space. In real space, the power spectrum becomes the correlation function,  $\xi(\vec{x}' - \vec{x})$ . If we rewrite  $\delta$  into the deterministic part and the initial conditions, the matter power spectrum can be written as

$$P(q) \propto P_p(q) |\delta(t, q)|^2.$$

Unlike the CMB, galaxy surveys have an effective redshift so  $\delta$  may not necessarily be integrated to today. The measure of the matter power spectrum is a comparison of observation with the first moment of the matter perturbations. It also does not require integration over the Fourier space unlike the CMB angular power spectrum. In this respect, the calculation of the matter power spectrum is simpler. However, this is where the simplicity ends as the next couple of sections discuss.

Before moving to the next section, we now discuss the window function for galaxy surveys. We limit ourselves to volume galaxy surveys rather than deep, narrow surveys. The covariance of the signal is

$$C_{S,ij} \propto \int d^3q P(q) w_i(\vec{q}) w_j^*(\vec{q})$$

where  $w_i$  is a weighting function based on the pixel scheme. A Fourier pixel scheme results in

$$w_i = \int \frac{d^3x}{V} e^{-i\vec{q}\cdot\vec{x}} e^{i\vec{q}_i\cdot\vec{x}}.$$

The window function is defined as the integral over the angular parts in the signal covariance definition. If we integrate the weighting functions over the entire solid angle and divide by  $4\pi$ , the signal covariance matrix becomes

$$C_s = \int_0^\infty \frac{dq}{q} W(q) \frac{q^3 P(q)}{2\pi^2}$$

where  $W(q)$  is the window function. Unlike the CMB analysis, the signal covariance and window function integrals do not come out nicely. In general, the power spectrum from a survey is not measured well on scales near the size

of the survey. Therefore, we are only concerned with  $q \gg 1/R$  where  $R$  is the radius of the survey. In this case, the window function behaves close to a delta function picking out only the very nearby modes. The signal covariance simplifies to

$$C_{S,ii} \approx \frac{P(q)}{V}.$$

See [24] for more discussion.

### 3.3.2 Galaxy Surveys

The power spectrum from galaxy surveys differ from the theoretical matter power spectrum in three ways:

1. Nonlinear Gravitational Evolution[92]
2. Redshift space distortion[93]
3. Bias[94]

The theory behind calculations of the CMB and LSS power spectrum is a linear theory. It assumes that the perturbations are small. This assumption is valid for the CMB. However, for matter at small scales, this assumption no longer holds true. The small gravitational perturbations at the time of last scattering grow as time progresses. At some point, these perturbations grow large enough such that gravitationally bound structure form. In a universe in which matter is dominated by CDM, structure begins at small scales first. In fact

$$q > q_{\text{NL}} \approx 0.2h \text{ Mpc}^{-1}$$

are non-linear today.  $q_{\text{NL}}$  is a time dependent term. The earlier the time, the smaller the scales that have become non-linear. Calculation of the non-linear power spectrum is difficult. N-body simulations have been carried out by groups[95][96] but these can only be done for a small set of parameter values. At this point it is unrealistic for statistical inferencing. Analytic fitting schemes to calculate the non-linear effect have been studied such as HALOFIT[97].

Galaxy surveys do not measure position of objects but rather their redshifts. This fact causes a distortion from the real space distribution due to gravitational effects. Galaxies form cluster structures among themselves. If a galaxy is closer to us than the rest of its cluster, it is gravitationally pulled towards the cluster center. This causes a redshift since the galaxy has a peculiar velocity away from us. The opposite effect occurs when the galaxy is behind the cluster from our point of view. This scenario causes a blueshift since the galaxy is attracted toward us. The total effect is to stretch the apparent distribution in redshift space. This effect is sometimes referred to as the Fingers-of-God effect.

Finally, the bias is the relationship between the distribution of CDM and normal matter. CDM is modeled as non-interacting whereas normal matter interacts with photons through Compton scattering. This interaction results in a different power spectrum for normal matter. Therefore, the power spectrum of galaxies does not reflect the total matter spectrum exactly. This bias must be accounted for when comparing with theory. However, this bias is also a non-linear effect. [98] proposed an empirical equation to relate the

galaxy power spectrum to the linear matter perturbation power spectrum and account for non-linearity.

### 3.3.3 Halo Power Spectrum

The analysis by Reid et. al.[99] reduced the effect of bias in LRG by comparing the Halo power spectrums. As stated before, previous analysis measured the galaxy power spectrum and related this to the matter power spectrum with an empirical equation for the bias. Reid et. al. take the measured galaxy survey and infer a halo power spectrum. This measured halo power spectrum is then compared to a theoretical halo power spectrum derived from the matter power spectrum. The advantage of this method is that the bias is greatly reduced. The disadvantage is the need of N-body simulations in order to calculate the theoretical halo power spectrum. In [100], the preliminary work on the N-body simulations was done.

In this approach, the theory and the measurement meet halfway. The halo power spectrum must be calculated from the linear power spectrum. The measurement of the galaxy survey must be turned into an estimate of the halo power spectrum. To calculate the theoretical halo power spectrum, three corrections to the linear power spectrum were performed. First a suppression of the BAO signal was done analytically using the empirical formula in [101]. The second correction uses Halofit to model the nonlinear corrections. Even at  $q < q_{nl}$ , there are small corrections that need to be taken into account. The third is the bias. While the bias is smaller than the galaxy power spectrum, it

still exists. As in [98], an empirical bias equation is formed and marginalized over during statistical analysis.

On the observational side, the calculation of the halo spectrum involves understanding the errors. Understanding the errors means one must understand  $C_S$ . Unlike the CMB, the  $C_S$  in the matter power spectrum does not follow a simple solution. Only in the extreme simple case do we get an analytic expression. The  $C_S$  is estimated using mock catalogues of the Halo power spectrum, using the experimental window function. Instead of N-body simulations, log-normal distributions are assumed as the distribution of the underlying density[98].

### 3.3.4 Forecasting

Similar to the CMB, we present the forecasting methods for the matter power spectrum. Section 3.3.1 discussed the signal covariance. We assume the simplification of  $C_S \approx P(k)/V$  for this subsection. The noise for a galaxy survey can be summed up as

$$\hat{\delta}_i = \int d^3x w_i(\vec{x}) \frac{n(\vec{x}) - \bar{n}(\vec{x})}{\bar{n}(\vec{x})}.$$

This shows the difference between the observed average density and the density at a certain pixel.  $w_i$  here is the Fourier pair that was seen in section 3.3.1. We assume Poisson noise; this is the appropriate model when we have discrete tracers of an underlying continuous field. The Poisson distribution has its expectation value equal to its variance. We set this value equal to  $\lambda$ .

Based on the above equation and the properties of the Poisson distribution, the noise covariance matrix is

$$C_N = \int d^3x \frac{w_i(\vec{x})w_j^*(\vec{x})}{\bar{n}(\vec{x})}.$$

This assumes that  $n$  comes from a Poisson distribution with  $\lambda = \bar{n}$ . We can follow the exact same steps of the CMB to find the error in the estimate of the power spectrum.

$$\sqrt{\frac{2}{NV}} \left( P_i + \frac{1}{\bar{n}} \right) \quad (3.15)$$

Here  $P_i$  is the power spectrum for bin  $i$ .  $N = 4\pi q_i^2 \delta q$  which takes into account the bin process. This case is somewhat simpler than the CMB only because we have made very rough approximation on  $C_S$ . In [99], the  $q$ 's studied were restricted to roughly  $0.02h \text{ Mpc}^{-1}$  to  $0.2h \text{ Mpc}^{-1}$ . The latter number is due to the nonlinear growth of structure as discussed before. To push this value higher would require more precise knowledge of the nonlinear power spectrum and bias. The former number is mostly due to the geometry of the survey. Since only a finite size is observed, not all  $q$  modes can be observed with precision. This number can be pushed lower with greater size. For our purposes, we restrict our analysis to this range.

## Chapter 4

### Massive Neutrinos in Cosmology

In chapter 2, we outlined the theoretical equations governing expansion and the perturbations. We implicitly restricted the discussion to massless neutrinos for simplicity. In this section, we discuss neutrinos with small masses. This small mass creates more complexities than seen in chapter 2. Massive neutrinos act as hot dark matter (HDM). Unlike CDM, HDM decouples from the primordial plasma while it is relativistic. In section 2.1, we saw that the energy density of non-relativistic matter and massless particles had simple behaviors with respect to the scale factor. HDM does not and generally requires numerical calculations for  $\rho$ . Since HDM does not in general have  $p \ll E$ , we can expect the need for the radiation hierarchy to describe the perturbations as seen in section 2.3.3. Connecting these perturbations to the stress energy tensor also requires numerical integration.

The study of massive neutrinos in the cosmological framework has been done for a number of years. In 1966, Gershtein and Zeldovich placed an upper bound of 15 eV based solely on the idea that neutrinos should not be so heavy as to overclose the universe[102]. In 1977, Sato and Kobayashi studied the mass bounds of neutrinos from decays of heavy leptons[103]. [104] studied the

effect of massive neutrino on large scale structure in 1980. Other investigations on the effect on cosmic observables by neutrinos include [105][106]. In more recent studies, [107] examined the accuracy of the fluid approximation against the full radiation hierarchy equations. [108] investigated the massive neutrino anisotropy background. In a work similar to our aims, [109] found that the matter power spectrum needed a precision less than 0.5% to differentiate between hierarchy using a 2 mass eigenstates model. Bounds on the neutrino mass have been studied by a variety of researchers. [110] found an upper bound of 1.5 eV by combining WMAP 5 and SDSS LRG DR7 data. [111] found an upper bound of 0.44 eV without axions and 0.41 eV with axions when combining WMAP 7, LRG 7, and measurement of  $H_0$ . Recent studies have investigated the effects on the nonlinear power spectrum and observations from weak lensing[112], the growth function and the observed galaxy luminosity function[113], and the nonlinear halo based on hydrodynamical simulations[114]. For a review of the effects of massive neutrinos in cosmology, see [115].

#### 4.1 Effect on the FRW universe and perturbations

As seen in section 2.1, the behavior of the energy density for photons and CDM is simplistic. The radiation and non-relativistic matter energy density scale as  $a^{-4}$  and  $a^{-3}$ , respectively. HDM does not have an analytic relation. In the early universe,  $\rho_{\text{HDM}}$  behaves like radiation. At the present time, HDM behaves like CDM. In general, the energy density for HDM must

be numerically calculated.

$$\rho = g \int \frac{d^3p}{(2\pi)^3} \sqrt{p^2 + m^2} f(x, p, t)$$

The aspects of the numerical integration is discussed later. For this discussion, it is sufficient to recognize the complexity of  $\rho_{\text{HDM}}$ . Because massive neutrinos behave like non-relativistic matter today, we can relate their relative fraction to their absolute mass. First, the relative fraction is

$$\Omega_i = \frac{\rho_i}{\rho_{cr}}$$

where

$$3H_0^2 = 8\pi G\rho_{cr}.$$

Also, it is common to break up the Hubble constant into a unitless parameter:  $H_0 = \left(100 \frac{\text{km/s}}{\text{Mpc}}\right) h$  where  $h$  is unitless. Based on equation 2.6, for a non-relativistic species,  $\rho = nm$ . Also for HDM,  $n \propto a^{-3}$ . Therefore, if we know the current number density of neutrinos today, we have a simple identity between the absolute mass and the energy density today.

$$\Omega_\nu = \frac{\sum m_\nu}{94.1 \text{eV}}$$

For the perturbations, we saw that in the massless case, the radiation hierarchy followed equation 2.22. For massive neutrinos, we have a similar equation only with factors of  $p/E$ . Taking the Legendre moments of equation 2.19 yields

$$\dot{\Psi}_\ell = -\frac{1}{a} \frac{qp}{E} \frac{1}{2\ell + 1} [(\ell + 1)\Psi_{\ell+1} - \ell\Psi_{\ell-1}] \quad (4.1)$$

for  $\ell \geq 2$ . For  $\ell$  equal to 0 or 1, extra terms from the gravitational perturbations emerge as seen in section 2.3.3. In the strictly non-relativistic case,  $p \ll E$  and this equation truncates after  $\ell = 1$ . This is not necessarily so for the neutrino case. In the next section, we discuss the qualitative effects of this equation on the observables. Section 4.3 discusses how these equations are numerically solved. Because of the nature of HDM, the radiation hierarchy calculation is somewhat different than massless particles.

## 4.2 Effect on Observables

Geometric probes are restricted to the near present universe. Therefore, the effects of neutrino mass are difficult to observe. Even with its seemingly small absolute mass, massive neutrinos act like non-relativistic matter. For measurements of  $H_0$ , HDM and CDM are indistinguishable. Also, measurements of Sne Ia will mostly likely not exceed a redshift of 10. Therefore, massive neutrinos are no different than non-relativistic matter for these probes. However, while the expansion history of the universe can only be directly observed in recent times, its effects can be seen in the perturbative observations of the CMB and LSS.

The unique modifications of the expansion rate from HDM can be seen in perturbations. One such effect occurs to the time of matter-radiation equality. As discussed in section 3.2, the CMB is a result of the acoustic oscillation in the primordial plasma. These oscillations occur at scales that have entered the horizon during the radiation dominated period. Perturbations at scales

which enter the horizon during the matter dominated period do not oscillate. Therefore, the time of matter-radiation equality is an important marker. If neutrinos are massless, this time can be calculated quite simply. However, the inclusion of HDM makes this calculation non-analytical. This effect can be easily seen in the matter power spectrum. Figure 4.1 is an example of an estimate matter power spectrum from various probes[2]. Note that in both figure 4.1 and figure 4.2, the x-axis label  $k$  is equivalent to  $q$  in our work. The solid curve is the theoretical power spectrum for the best fit values of the cosmological parameters. The turnover point indicates the scale that entered the horizon at equality. For scales smaller than this (larger  $q$ ), the power spectrum decreases. These perturbations are erased by the radiation dominated universe. The earlier the scale enters the horizon, the more time there is to interact and to dampen these perturbations. Larger scales have smaller amplitude since they have had less time inside the horizon. Less time inside the horizon means less time for gravity to clump up matter.

We have seen one of the effects of HDM on perturbations. This effect was actually due to modifications in the background evolution, not the actual perturbations themselves. Another effect due to HDM is the suppression of perturbations at small scales. The behavior of HDM in this case is scale dependent. At present times, HDM at large scales behave the same as CDM. In fact they are indistinguishable. However, at small scales, HDM's relativistic behavior emerges. The large kinetic energy counteracts the gravitational pull of metric perturbations. In effect, part of the matter can not clump at small

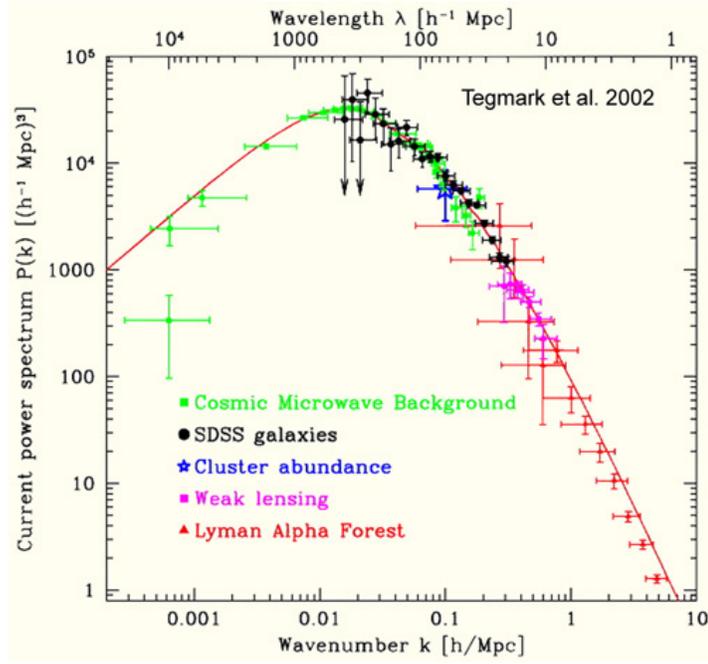


Figure 4.1: Estimated Power Spectrum from [2]

scales. Therefore, the matter power spectrum is damped at small scales with respect to a non-HDM universe. We can examine equation 4.1 to understand this behavior. The right hand side is sensitive to  $\frac{q}{a} \frac{p}{E}$ . Massive neutrinos will behave like radiation if this prefactor is significant. Even if  $\frac{p}{E}$  is small,  $q$  might be large enough such that massive neutrinos behave like radiation. Therefore, we expect at large  $q$ , massive neutrinos still have radiative behavior and do not clump. This prefactor also depends on the scale factor and therefore the Hubble rate. The larger the universe becomes, the smaller this prefactor is. The scale at which suppression occurs is sometimes called the freescale. The dependence of the freescale through the pre-factor times time to take make it

unitless. The scale would be the inverse of  $q$  with  $p$  scaling as inverse of  $a$  and  $E$  equal to the mass. Figure 4.2 demonstrates this effect[3]. The dotted and dashed curves are for massive neutrino models. The solid curve is the power spectrum for the massless case.

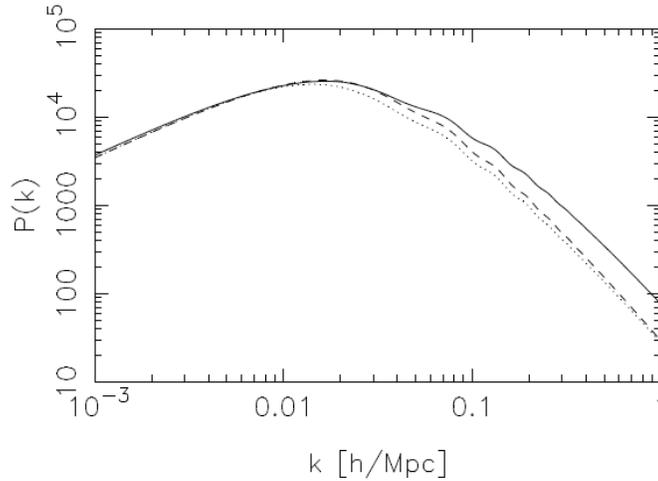


Figure 4.2: Effect of Massive Neutrinos on Power Spectrum[3]

The CMB and LSS are more sensitive to the effects of HDM than nearby observations of the expansion rate. As discussed in the last paragraph of section 4.1, we expect a suppression of the matter power spectrum at large  $q$ . This effect is perhaps our best chance of observing the effects of massive neutrinos. Unlike the CMB, the LSS is a measurement of the matter perturbations. Therefore, LSS observations are probably the closest thing to a direct observation of cosmic neutrinos as we can practically expect. Also, the matter power spectrum does not have the fundamental uncertainty of the

cosmic variance as the CMB does. However, several factors exist which hamper the precision of LSS as discussed in section 3.3. Some of the suppression effects of massive neutrinos fall in the non-linear regime where perturbation calculations are not accurate. The bias problem also needs to be addressed precisely in order for more precise inferencing from galaxy surveys. The CMB, on the other hand, provide more precise measurements of the perturbations. However, these measurements are perturbations of the photons. These perturbations themselves are indicative of the gravitational perturbations of the early universe. Therefore, the effects of the modified expansion history and dampening at small scales is not as strong in the CMB. The dampening effect is an integrative effect over time. The CMB is more analogous to a snapshot of the early universe.

### 4.3 Numerical Calculations

The calculation of massive neutrinos effects both the FRW energy density & pressure and the perturbations. For the energy density,

$$\rho = g \int \frac{d^3p}{(2\pi)^3} \frac{\sqrt{p^2 + m^2}}{e^{E/T} + 1} \quad (4.2)$$

where we now assume a Fermi-Dirac phase space distribution. By taking this form, we assume that the neutrinos were in thermal contact with the primordial plasma in the early universe. After decoupling, the phase-space distribution keeps its thermal shape but with the temperature falling as  $a^{-1}$ . Similarly the

pressure is

$$P = g \int \frac{d^3p}{(2\pi)^3} \frac{p^2}{\sqrt{p^2 + m^2}} \frac{1}{e^{E/T} + 1}. \quad (4.3)$$

We see that the calculations of  $\rho_\nu$  and  $P_\nu$  require numerical integration for every  $T$  needed. This scenario can be quite cumbersome. Rather than calculate this for every instance of  $a$ , interpolation tables are calculated first. Then a spline method is used for interpolation. The current CAMB code's default setting calculates  $\rho_\nu$  and  $P_\nu$  at 2000 points to build the interpolation tables. At these points, the integrals in equations 4.2 and 4.3 are solved with an extended trapezoidal rule plus the next order term in the Euler-Maclaurin summation formula (see Appendix). We find that the current settings provide very good accuracy for the calculation of  $\rho_\nu$ . An eight-fold increase of the number of interpolation points has almost no effect. Figure 4.3 illustrates this error. The x-axis is just an index value. The largest difference is  $1.4 \times 10^{-10}$ . Therefore, we are not concerned with the error from these numerical integrations.

As stated before, HDM can not be approximated by fluid equations. Just like the photons, the neutrinos follow a hierarchy equations stemming from the anisotropic pressure. Equation 4.1 defines the radiation hierarchy for massive neutrinos. Accuracy of the code can be improved by increasing the number of equations actually calculated. This statement is also true for photons and is discussed in section 5.1.2. However, as discussed in chapter 2, we have to connect  $\Psi$  with the metric perturbations through the stress energy

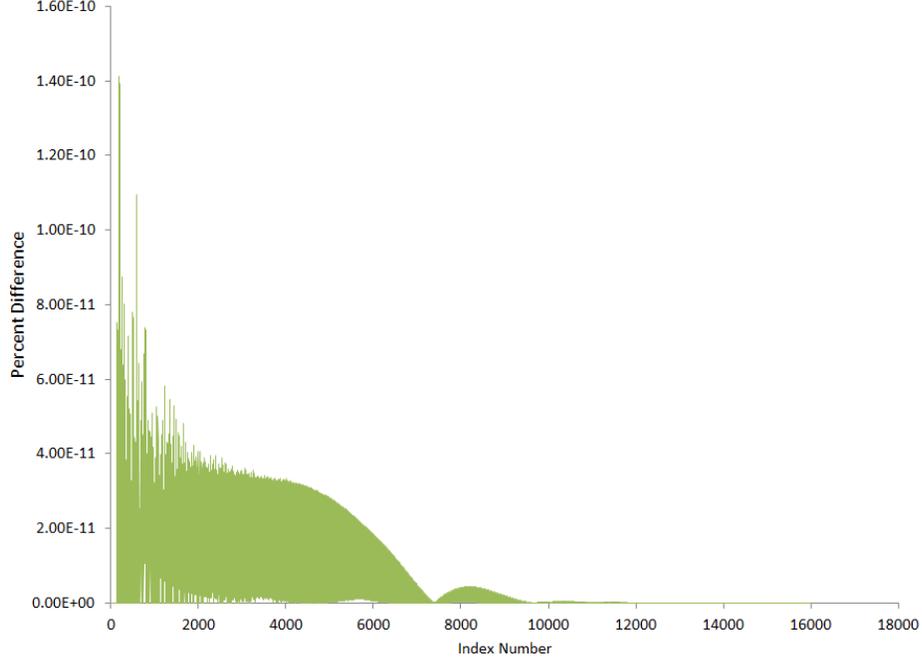


Figure 4.3: Interpolation error from 8-fold increase in calculation of  $\rho_\nu$ . Integral calculated at points from 0.01 eV to 600 eV. Numbers on horizontal axis are the array indexes for points between this range.

tensor. Specifically, we need to calculate

$$\delta = \frac{1}{\rho} g \int \frac{d^3 p}{(2\pi)^3} \sqrt{p^2 + m^2} f_0 \Psi$$

$$\delta T_j^i = g \int \frac{d^3 p}{(2\pi)^3} \frac{p^2 \hat{p}^i \hat{p}^j}{E} f_0 \Psi$$

$$v^i = \frac{1}{\rho_0 + P_0} g \int \frac{d^3 p}{(2\pi)^3} p \hat{p}^i f_0 \Psi$$

Just as in the zero order  $\rho$  and  $P$  case, these integrals are numerically calculated using the extended trapezoidal rule. Since the methodology and parameters of these methods are the same as the calculation of  $\rho$ , we expect the error to be of the same order of magnitude.

## 4.4 Effects from Hierarchy

The previous parts of this chapter focused on the difference between massive and massless neutrinos. Here, we present numerical calculations to demonstrate effects from neutrino hierarchy. In particular, we examine the possible effects of neutrino mass hierarchy on the matter power spectrum and CMB angular spectrum. Based on our earlier discussion, we do not present any comparisons of geometric probes. We assume a flat universe with adiabatic initial conditions and a cosmological constant to explain the present acceleration. Tables 4.1 and 4.2 show the chosen parameter values. The  $\tau$  parameter determines the amount of reionization. The figures below show

Table 4.1: Testing Values for Non-Neutrino Parameters.  $H_0$  in units of (km/s)/Mpc

$\Omega_b h^2$	0.02219
$\Omega_c h^2$	0.1122
$H_0$	65.0
$\tau$	0.086
$n_s$	0.953
$A$	$2.5 \times 10^{-9}$

Table 4.2: Testing Values for Neutrino Masses

	Normal	Inverted	Degenerate
m1	0.015 eV	0.001 eV	0.038 eV
m2	0.024 eV	0.051 eV	0.038 eV
m3	0.074 eV	0.061 eV	0.038 eV
$\sum m_\nu$	0.113 eV	0.113 eV	0.114 eV
$\Omega_\nu h^2$	0.0012	0.0012	0.0012

the difference between the hierarchy schemes and the degenerate case. We

present the differences in terms of percentages. Figure 4.4 shows the difference between the two neutrino mass hierarchies and the degenerate case for the matter power spectrum. The percent differences are on the order of a tenth of a percent. Note that we can not accurately model the matter power spectrum much beyond the  $q \approx 0.2$  mark. Nonlinear effect begin to take effect here. Figure 4.5 shows the difference between the two hierarchies for the

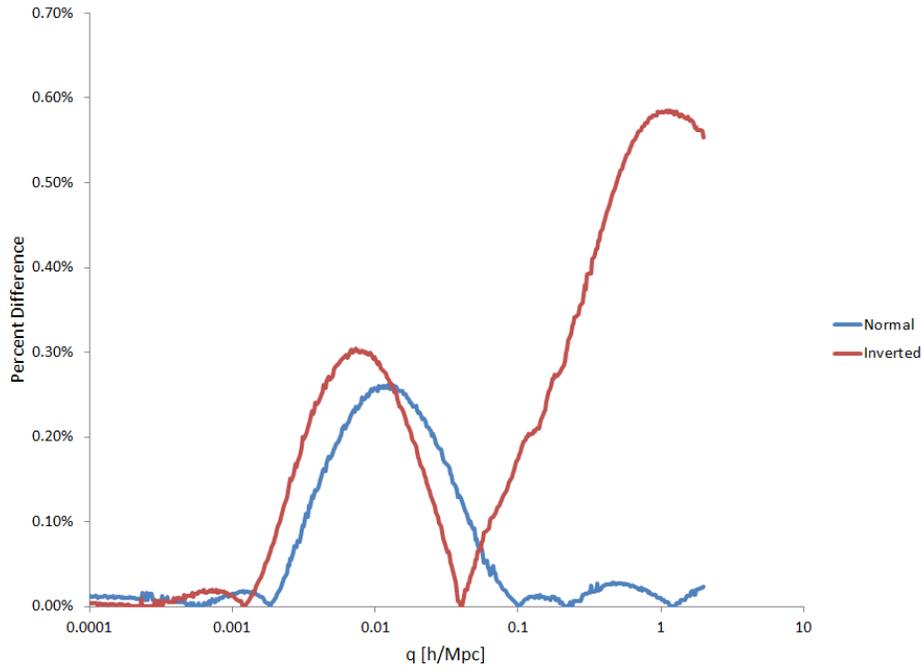


Figure 4.4: Percent difference between Mass hierarchies and Degeneracy

matter power spectrum. The percent difference between the two is similar to the difference with the degenerate case. Therefore, we can expect that at the same level where the degenerate case is no longer accurate, we would be able to differentiate between the two hierarchies. In other words, based on figures

4.4 and 4.5, when the degenerate case is no longer accurate, our data should be able to reveal which hierarchy fits the observations best.

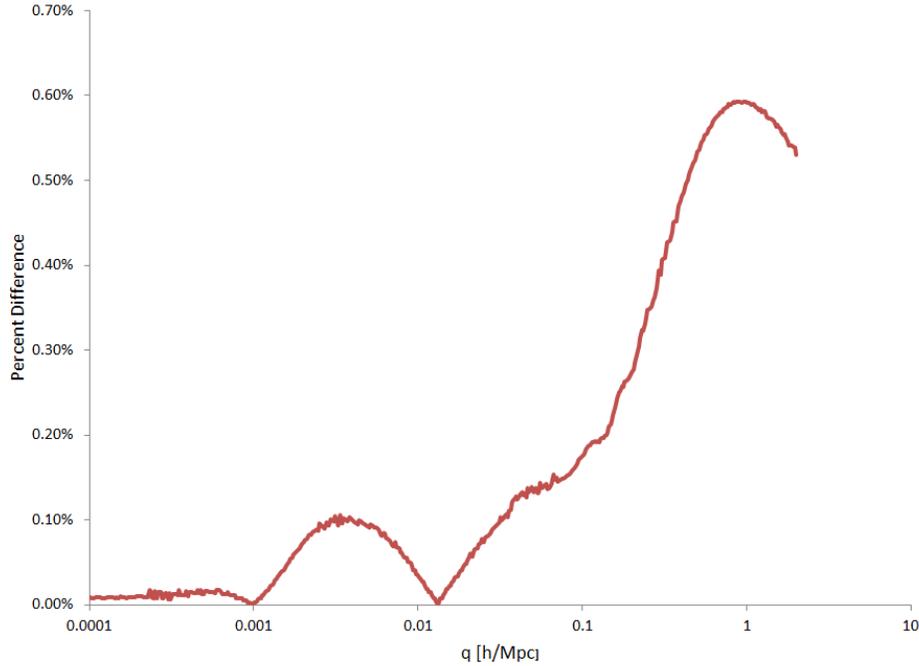


Figure 4.5: Percent difference between Normal and Inverted Hierarchy

Figure 4.6 shows the percent difference between the mass hierarchies and the degenerate case for the CMB angular spectrum. The percent level difference is on order of hundredth percents. The difference in the CMB is an order of magnitude lower than the matter power spectrum. Note that the cosmic variance is roughly 2% at the lowest. Even with the precision of the CMB probes, the difference between hierarchy and non-hierarchy can not be inferred from the CMB alone. The cosmic variance prevents this.

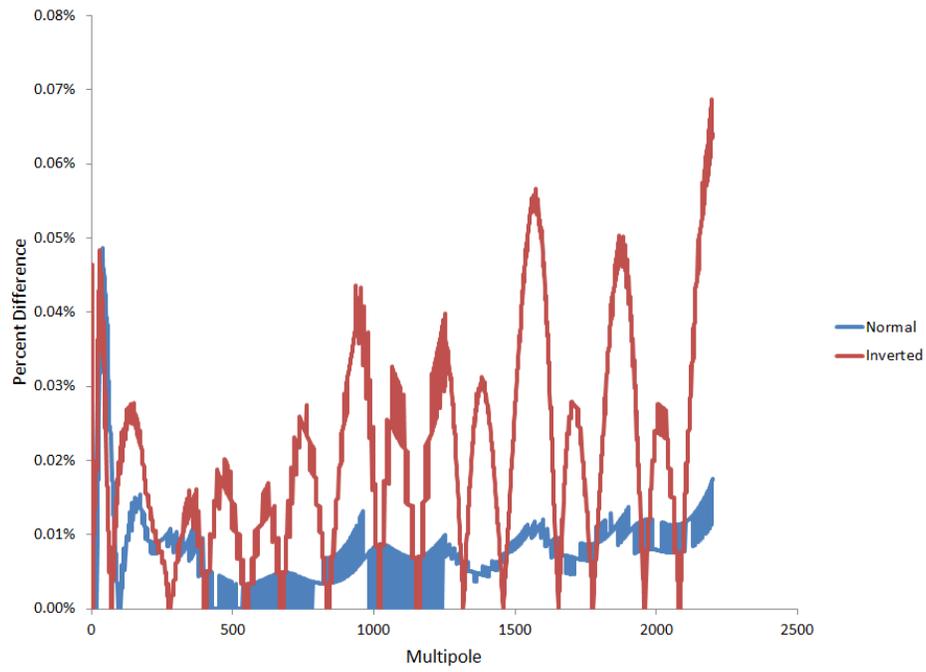


Figure 4.6: Percent Difference in the CMB between Neutrino Hierarchy and Degeneracy

## Chapter 5

### Numerical and Computational Aspects

The physics of cosmology is encoded in the set of coupled ODEs from chapter 2. Here, these perturbation equations are examined from a numerical point of view. These issues are important to understand the accuracy of CAMB. From a numerical point of view, the core of the calculation is the integration of the cosmology ODEs in the Fourier space. The errors associated with numerical integration methods are akin to Taylor series approximation. In addition, the Boltzmann equation for relativistic species involve a hierarchy of ODEs as seen in section 2.3.3. This infinite hierarchy must be truncated which introduces another source of error. Finally, calculation of observables requires interpolation. Solving for ever  $C_\ell$  is not practical. We test the accuracy of the CAMB code for models with neutrino mass hierarchies.

#### 5.1 General Accuracy

In this section, we discuss factors that effect the calculations of CMB and LSS for any models. The three main factors are: numerical integration, radiation hierarchy truncation, and calculation of the  $C_\ell$ 's. The first factor stems from any numerical integration of ODEs. The second factor was dis-

cussed in section 2.3.3. We discuss the use of the line-of-sight algorithm which helps alleviate this issue. Finally, for the CMB angular spectrum, the  $C'_\ell$ s are integrated over the entire Fourier space. This numerical integration and interpolation leads to errors.

### 5.1.1 Integration Scheme

For cosmology, we are interested in the initial value problem. We have a set of ODEs with some defined initial conditions. Our discussion of errors is general. More details about the exact schemes can be found in the appendix. We focus only on explicit (time-forward) schemes. The errors introduced in numerical integration are independent of any physics or cosmology models. In fact, the stochastic initial conditions are removed from the actual ODEs as seen in equation 3.9. Most of the computation can be done without regard to the primordial power spectrum. Once the integration has completed, then the initial conditions can be applied; this statement is true as long as we work in the linear regime. For a reference to numerical integration of ODEs, see [116].

Explicit schemes for ODE integration stem from the finite difference approximation to derivatives. For example, if we have

$$\frac{\partial y}{\partial t} = f(y, t),$$

where  $f(y, t)$  is some function, the explicit finite difference approximation would be

$$\frac{y(t+h) - y(t)}{h} \approx f(y, t)$$

where  $h$  is some finite number. The simplest numerical scheme for integration would be

$$y_{i+1} = y_i + hf_i$$

where  $y_i$  and  $f_i$  are  $y$  and  $f(y, t)$  evaluated at  $t_i = ih + t_0$ , and  $t_0$  is the initial time. This is why these methods are also called “time-forward” methods.  $y$  at the next time step is calculated based only on the current time step. The accuracy of the code depends on the choice of  $h$ . To estimate the error, we look at the Taylor expansion.

$$y(t + h) = y(t) + \frac{\partial y}{\partial t} h + \frac{\partial^2 y}{\partial t^2} \frac{h^2}{2} + \dots$$

Each step introduces an error of  $O(h^2)$ . The total error then is

$$\left( N = \frac{t_f - t_0}{h} \right) (h^2) = O(h)$$

where  $N$  is the number of steps. This type of error is termed “truncation” error. Truncation errors are based on finite difference approximation of the derivative. Even if a computer could hold infinite number of bytes per number, this error still occurs. Ideally we could set  $h$  as low as possible. However, there is a minimum value which  $h$  can be due to the finite bytes of memory. This type of error is termed “round-off” error. If  $h$  is very low, then the difference between  $y(t+h)$  and  $y(t)$  becomes extremely small. If this difference is smaller than the smallest number the memory can hold, then the output is erroneous. This round-off error is dependent on the precision or the number of bits used per value.

The simple example above was the Euler method. It is considered a first order integration method. In general, extensions of this method, called Runge-Kutta (RK) methods, are used in scientific computing. Trial steps are taking between the time interval to obtain a more accurate estimate of the derivative. For example, a second order RK method uses a middle point to increase the accuracy.

$$\begin{aligned}
 k_1 &= hf(y, t) \\
 k_2 &= hf\left(y + \frac{1}{2}k_1, t + \frac{1}{2}h\right) \\
 y_{i+1} &= y_i + k_2
 \end{aligned}$$

This RK2 scheme produces an error of order  $h^3$  per step and  $h^2$  overall. RK2 is also referred to as the midpoint method. A fourth order RK method calculates

$$\begin{aligned}
 k_1 &= hf(y, t) \\
 k_2 &= hf\left(y + \frac{1}{2}k_1, t + \frac{1}{2}h\right) \\
 k_3 &= hf\left(y + \frac{1}{2}k_2, t + \frac{1}{2}h\right) \\
 k_4 &= hf(y + k_3, t + h) \\
 y_{i+1} &= y_i + \frac{k_1}{6} + \frac{k_2}{3} + \frac{k_3}{3} + \frac{k_4}{6}
 \end{aligned}$$

This method results in an order  $h^5$  error per step and  $h^4$  overall. RK4 is a popular method for ODE integration. It is a rather robust integrator which is also rather easy to implement. It is not always the most accurate but its balance of simplicity and accuracy makes it a common choice. The downside

is the possibility of instabilities. Time-steps must be made small enough such that the RK algorithm is stable. The actual time-steps also control the accuracy of the integrals. The CAMB code uses a RK45 code. This method is a slight variation on the Runge-Kutta scheme.

The cosmology ODEs are Fourier transformed. This removes the spatial derivatives and also connects better with our initial conditions which are probabilistic statements in Fourier space. In first order linear perturbation theory, each  $q$ -mode evolves independent from one another. This simplification allows calculations to be performed on high-performance parallel computers. One may simply distribute the various  $q$ 's to multiple processors to calculate. This feature is important because RK methods and MCMC statistical methods do not parallelize well if at all. In fact, the higher the accuracy of the code, the more time steps are required. These steps themselves cannot be parallelized as they are explicit.

### 5.1.2 Numerical Radiation Hierarchy

As seen in section 2.3.3, relativistic species have an infinite hierarchy of ODEs to describe the evolution of their phase space distribution. Initial calculations were limited by the size of hierarchy they could calculate. If one wished to calculate  $C_\ell$  to some  $\ell_{\max}$  then one must include all the ODES in the hierarchy beyond this point to ensure accuracy for  $C_{\ell_{\max}}$ . In addition, one must integrate over sufficient number of Fourier modes to obtain accurate values. The line-of-sight integration approach by [28] changed this scenario.

This algorithm works as follows. First, one integrates equation 2.15 along the photon past light cone. This results in an integration expression for  $\Delta$ . The Legendre moments of  $\Delta$  then follow as

$$\Delta_\ell(t_0, q) = \int_0^{t_0} dt S(t, q) j_\ell(q(t_0 - t)).$$

Here  $j_\ell$  is the Bessel function,  $t_0$  is the time today, and  $S(t, q)$  is called the source function. The source function contains all the physics and is the result of the integration of the ODEs. Note that  $S(t, q)$  does not depend on  $\ell$ . For all  $\ell$ 's, we only need one  $S(t, q)$ . This method effectively eliminates the need for a large radiation hierarchy.

The hierarchy is not completely eliminated. The source function depends on specific moments stemming from the collision terms. For photons, the monopole and quadrupole moments of the photons intensity and polarization is needed. Therefore, the radiation hierarchy needs to go far enough such that these moments are integrated accurately. However, this hierarchy is much smaller than the work prior to [28] and is no longer dependent on the desired  $\ell_{\max}$ . Current practice is to keep the hierarchy up to 7 or 8 moments. For neutrinos, this method simplifies the neutrino hierarchy into another source function. The source function for neutrinos only depends on the metric perturbations.

### 5.1.3 Calculation of Observables

The two main observables that are calculated based on perturbations are the CMB angular spectrum and LSS power spectrum. The LSS matter

power spectrum is calculated based on the transfer functions:  $\delta_i$  where  $i$  indexes the different types of matter. Solving for the matter power is just a matter of integrating  $\delta_i$  to the desired redshift, find the total matter  $\delta$ , and then multiply by the primordial power spectrum. The accuracy of this calculation depends on the previous two factors.

For practical reasons, the  $C_\ell$ 's of the CMB angular spectrum are not calculated for every  $\ell$ . The smoothness of the CMB at large  $\ell$  allows one to interpolate with confidence. The CAMB code calculates  $C_\ell$  at intervals of 50 for  $\ell \geq 10$ . For low  $\ell$ , the discrete nature of the angular power spectrum requires narrow sampling. The  $C_\ell$  also require integration over the Fourier modes. The number of  $q$ 's for which the differential equations are solved can lead to more accurate  $C_\ell$ 's. For our purposes, this factor is not relevant since we focus on the LSS. We include this discussion for completeness.

## 5.2 Testing Results

We test the CAMB accuracy for the three factors described in section 5.1. The accuracy was checked for normal hierarchy, inverted hierarchy, and the degenerate case. The CAMB code has established parameters for these three factors: `boost`, `lboost`, and `lsampleboost`. The `lsampleboost` increases the number of points the spline routine uses for interpolation. `lboost` increase the number of terms in the radiation hierarchy for both photons and neutrinos. `boost` decreases the time steps and other factors related to numerical integration of the ODEs. It also controls the number of Fourier modes

used. Our goal is to find the right values for `boost` and `lboost` such that the precision of the calculated matter power spectrum is sufficient for neutrino mass hierarchy. In other words, we want the error in  $P(q)$  to be less than the differences seen in section 4.4. Since we are only interested in the matter power spectrum we do not present the differences from increasing `lsampleboost` since that only affects the CMB calculations.

Figures 5.1 and 5.2 show the estimated accuracy from decreasing the time integration steps and the increase of terms in the radiation hierarchy. As expected, the precision appears independent of the actual mass hierarchy model. In both cases, the percent differences are of the same magnitude as those seen in the previous chapter. In fact, the maximum error due to the radiation hierarchy is barely under 1%. Therefore, the default accuracy settings are not sufficient when examining differences in the neutrino mass models.

Figures 5.3 and 5.4 show the estimated accuracy when `boost` and `lboost` equal 2, respectively. Note that these parameters are not restricted to integer values. An increase in the number of time steps seems to have very little effect on the accuracy of the code. However, increasing the number of terms in the radiation hierarchy does affect accuracy significantly. The maximum error is decreased by a factor of almost 10. In retrospect, this result is not unexpected. After all, at larger  $q$  values, the radiative terms in the neutrino perturbation become more relevant. Therefore, it seems intuitive that the radiation hierarchy becomes more important to the accuracy in massive neutrino models. Since the neutrino hierarchy causes effects at the 0.1% level,

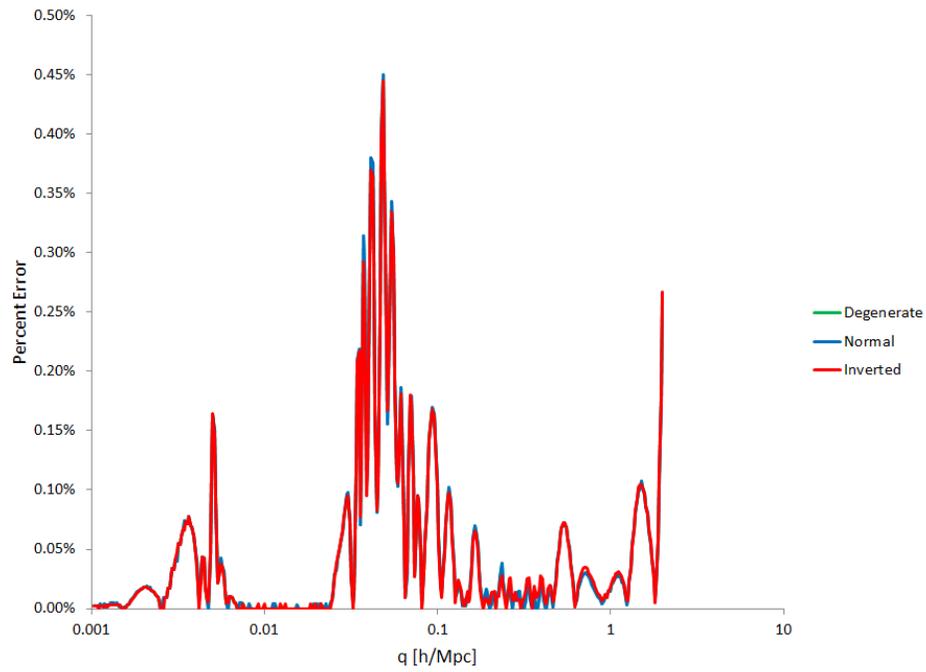


Figure 5.1: Accuracy of Power Spectrum for `boost = 1`

the error in the calculation of the power spectrum should be slightly more accurate than `lboost = 2`. A setting of `boost = 1.5` and `lboost = 2.5` should be sufficient.

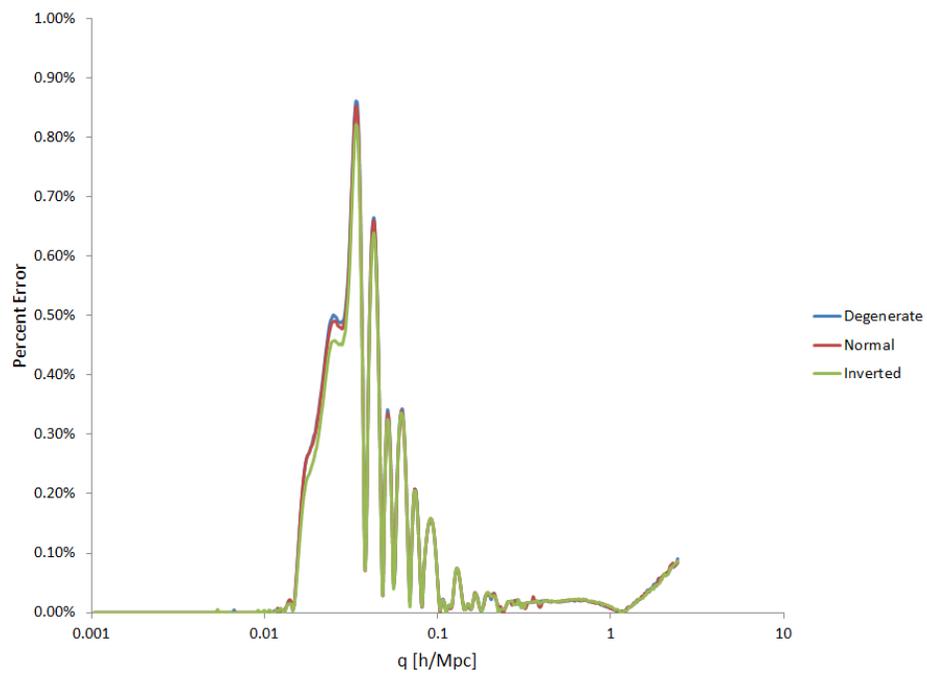


Figure 5.2: Accuracy of Power Spectrum for  $1\text{boost} = 1$

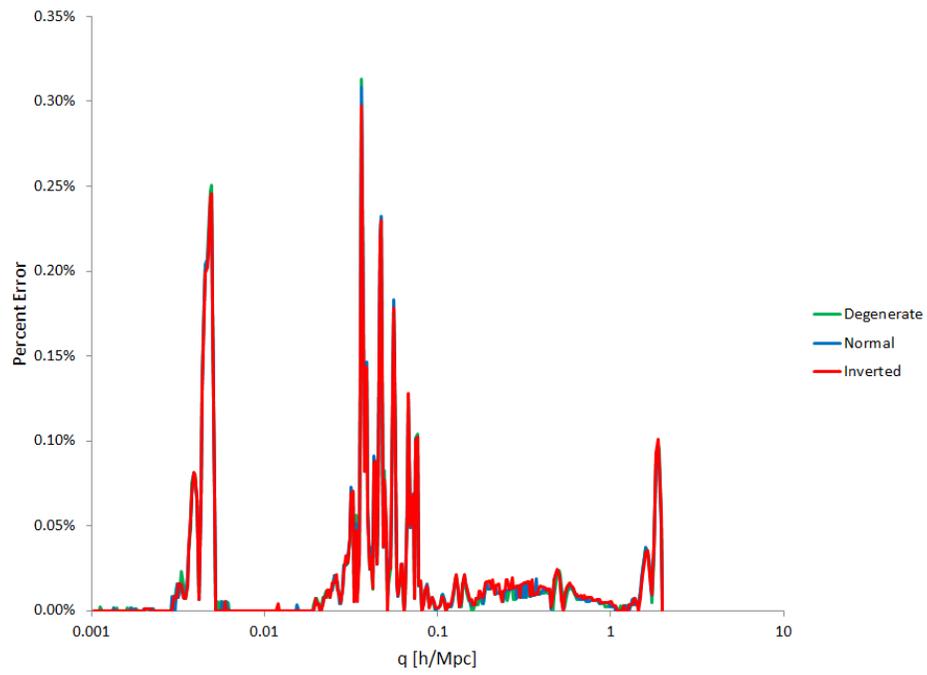


Figure 5.3: Accuracy of Power Spectrum for `boost = 2`

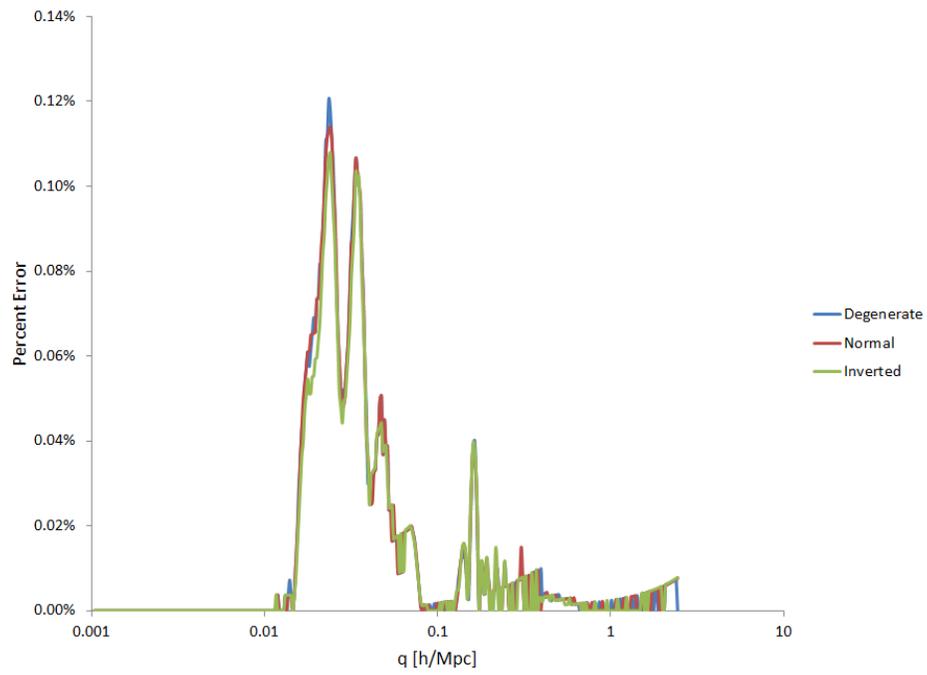


Figure 5.4: Accuracy of Power Spectrum for  $1\text{boost} = 2$

## Chapter 6

### Inference with Importance Sampling

In chapter 4, we examined the impact of neutrino hierarchy on the CMB and LSS. The differences in the CMB were well below the cosmic variance. Therefore, we do not expect the CMB to have the statistical power to differentiate between hierarchies. Rather than use the standard MCMC sampler, we propose the use of importance sampling. MCMC methods explore the likelihood using a Markov algorithm. In the limit of infinite sample size, the Markov chain converges to a true sample of the likelihood. However, this convergence may be slow depending on the exact type of MCMC. The increasing amount of data in the future will increase the amount of computational time to calculate the likelihood. In addition, since we are dealing with neutrino hierarchies, the theoretical calculations of power spectrum require more time.

An alternative route can be done with importance sampling. Importance sampling is based on weighting. Say we have a distribution  $g(x)$  which we have sampled from. Now we want to make statistical inferences on distribution  $f(x)$ . We can use those samples from  $g(x)$  and weight them by  $w = f/g$ . Then inferences can be made based on those weights and the original sample. For instance, if we want to know the mean of  $f(x)$ , denoted  $E_f[x]$ , we can

approximate it as

$$E_f[x] \approx \frac{1}{N} \sum x_i w(x_i)$$

where  $x_i \sim g(x)$ . If we already have samples from  $g(x)$ , we can forgo sampling from  $f(x)$  and weight the samples from  $g(x)$ . This technique can also be helpful if  $f(x)$  is hard to sample from but  $g(x)$  is not.

We implement importance sampling to take advantage of the CMB's lack of sensitivity to neutrino hierarchy. We propose to sample from the CMB likelihood based *only* on the degenerate model. The lack of sensitivity of the CMB means that these samples would represent not only the degenerate model but also the hierarchy models. Then these samples are weighted to the LSS based on importance sampling. This technique reduces the computational time in many ways. First, the MCMC sampler only needs to be done on the CMB. This sampling can be done with only the degenerate model. Once the MCMC sampling of CMB is done, the importance sampling weights can be calculated. At this step, we do not have to implement a random search as in MCMC. Therefore, importance sampling is a much more efficient algorithm.

## 6.1 General Framework of Importance Sampling

In this section, we explain importance sampling more formally. Let  $h(x)$  be a function where  $x$  is a random variable from distribution  $g(x)$ . For example, we might want to know the expected value/mean of  $g(x)$ . Then  $h(x)$

would be  $h(x) = x$ . The expected value is

$$E_g[h(x)] = \int_{\mathcal{X}} h(x)g(x)dx$$

where  $\mathcal{X}$  is the support or domain of the distribution  $g(x)$  and  $E_g[h(x)]$  denotes the expected value of the function  $h(x)$  when  $x \sim g(x)$ . This integral may be difficult to solve analytically. What if we could sample from  $g(x)$ ? Then the expected value can be approximated as

$$E_g[h(x)] = \int_{\mathcal{X}} h(x)g(x)dx \approx \frac{1}{N} \sum_i^N h(x_i)$$

where  $x_i$  is a set of  $N$  samples from  $g(x)$ . This approximation is akin to Monte Carlo integration.

Suppose we want to make inferences on  $h(x)$  but from a different distribution,  $f(x)$ . If the sampling from  $f(x)$  is possible and not too difficult, then we could do the same thing as before. But what if  $f(x)$  was too difficult to sample from or we could not sample from it at all? In some situations, it is easier to calculate the value of the distribution than to sample from it. Importance sampling addresses this issue. The expected value of  $h(x)$  can be

written as

$$\begin{aligned} E_f[h(x)] &= \int_x h(x)f(x)dx \\ &= \int_x h(x)\frac{f(x)}{g(x)}g(x)dx \\ &= \int_x h(x)w(x)g(x)dx \\ &= E_g[h(x)w(x)] \\ &\approx \frac{1}{N} \sum_i^N h(x_i)w(x_i). \end{aligned}$$

In the last line,  $x_i \sim g(x)$  not  $f(x)$ . The weights,  $w(x_i)$ , are equal to the ratio of the distributions:  $\frac{f(x)}{g(x)}$ . As a result, we can use our samples from  $g(x)$  to make inferences on  $f(x)$  if we weight them by the ratio of the distributions at each sample point. We have bypassed the need to sample from  $f(x)$ . All we need to be able to do is calculate the value of the distributions at different points.

What happens if we only know the distribution functions  $f(x)$  and  $g(x)$  up to proportionality? In other words,

$$\int_x f(x)dx = c \neq 1.$$

Technically,  $f(x)$  would not be a probability density function(PDF);  $\frac{1}{c}f(x)$  would be the PDF. We now explain how this affects our procedure.

Let

$$\int_x f(x)dx = c$$

and

$$\int_x g(x)dx = d.$$

Then the weights should be

$$w(x) = \frac{\frac{1}{c}f(x)}{\frac{1}{d}g(x)} = \frac{d f(x)}{c g(x)}.$$

Now inferences on the function  $h(x)$  can be approximated by

$$\begin{aligned} E_f[h(x)] &\approx \frac{1}{N} \sum h(x_i)w(x_i) \\ &= \frac{1}{N} \sum \frac{d f(x)}{c g(x)} h(x_i) \\ &= \left(\frac{1}{N}\right) \left(\frac{d}{c}\right) \sum \frac{f(x)}{g(x)} h(x_i) \\ &= \left(\frac{1}{N}\right) \left(\frac{d}{c}\right) \sum w^*(x_i)h(x_i) \end{aligned}$$

where  $w^*(x_i) = \frac{f(x_i)}{g(x_i)}$  and  $x_i \sim \frac{1}{d}g(x)$ . We do not know what  $c$  and  $d$  are so we need to estimate them.

Remember that

$$c = \int_x f(x)dx.$$

We can modify this similarly as we did before. We change the integral such

that we can sample from  $g(x)$ .

$$\begin{aligned}
c &= \int_x f(x) dx \\
&= \int_x \frac{f(x)}{g(x)} g(x) dx \\
\frac{c}{d} &= \int_x \frac{f(x)}{g(x)} \frac{1}{d} g(x) dx \\
&= \int_x w^*(x) \frac{1}{d} g(x) dx \\
&\approx \frac{1}{N} \sum w^*(x_i)
\end{aligned}$$

where  $x_i \sim \frac{1}{d}g(x)$ . Now we have a numerical estimate of  $\frac{c}{d}$ . The estimation of  $E_f[h(x)]$  is now

$$\begin{aligned}
E_f[h(x)] &\approx \left(\frac{1}{N}\right) \left(\frac{d}{c}\right) \sum w^*(x_i) h(x_i) \\
&= \frac{\left(\frac{1}{N}\right) \sum w^*(x_i) h(x_i)}{\frac{1}{N} \sum w^*(x_i)} \\
&= \frac{\sum w^*(x_i) h(x_i)}{\sum w^*(x_i)}
\end{aligned}$$

where  $x_i \sim \frac{1}{d}g(x) \propto g(x)$  and  $w^*(x) = \frac{f(x)}{g(x)}$ .

So far in this discussion, we have assumed that  $g(x)$  and  $f(x)$  have the same support. What if they have different support? First, in order for importance sampling to be a valid method, the support of  $g(x)$  must contain the support of  $f(x)$ . In other words,  $\mathcal{Y} \supset \mathcal{X}$  where  $\mathcal{Y}$  is the support of  $g$  and  $\mathcal{X}$  is the support of  $f$ . For those samples of  $g$  which do not fall in  $\mathcal{X}$ , these samples are discarded. Since the expectation value of  $f$  is an integration over  $\mathcal{X}$ , then these outliers from the sample of  $g$  do not contribute.

## 6.2 Application to Cosmological Inference

We now apply this technique to cosmology.  $x$  becomes  $\theta$ : the set of parameters for the desired cosmological model. In the case of a flat  $\Lambda$ CDM model,  $\theta = \{\Omega_b h^2, \Omega_c h^2, \Omega_\Lambda, n_s, A, \tau\}$ . The distributions that we want to infer from are

$$\theta \sim f(\theta|\text{data})$$

where data is some desired combination of cosmological observables. In the language of Bayesian statistics,  $f(\theta|\text{data})$  is called the posterior distribution.

For example, we want to make inference on the combination of WMAP 7[6] and SDSS LRG DR7[99]. The posterior distribution is

$$f(\theta|\text{LRG} + \text{WMAP}) = f(\theta|\text{LRG})f(\theta|\text{WMAP})$$

where we have assumed that the two data sets are independent. Samples from  $f(\theta|\text{WMAP})$  can be acquired from [117]. In this case, the weights become

$$w(\theta) = \frac{f(\theta|\text{LRG})f(\theta|\text{WMAP})}{f(\theta|\text{WMAP})} = f(\theta|\text{LRG}).$$

If we want to make inferences on  $\theta \sim f(\theta|\text{LRG} + \text{WMAP})$ , we can use the samples from  $\theta \sim f(\theta|\text{WMAP})$  and weight them by  $w(\theta) = f(\theta|\text{LRG})$ .

The posterior distribution is often hard to calculate. By definition,

$$f(\theta|\text{LRG}) = \frac{f(\theta, \text{LRG})}{f(\text{LRG})} = \frac{f(\theta, \text{LRG})}{\int f(\theta, \text{LRG})d\theta}.$$

Using Bayes Theorem,

$$\begin{aligned} f(\theta|\text{LRG}) &= \frac{f(\text{LRG}|\theta)f(\theta)}{f(\text{LRG})} \\ &\propto f(\text{LRG}|\theta)f(\theta) \end{aligned}$$

where  $f(\text{LRG}|\theta)$  is called the likelihood, which we denote as  $\mathcal{L}_{\text{LRG}}(\theta)$ , and  $f(\theta)$  is called the prior distribution, which we denote as  $\pi(\theta)$ . The prior distribution can represent prior data or prior beliefs on the parameters. For now we let the priors be uniform/flat ( $\pi(\theta) \propto \text{constant}$ ). We can now calculate the posterior distribution and therefore the weights up to proportionality:  $w(\theta) \propto \mathcal{L}_{\text{LRG}}(\theta)$ .

In summary, if we want to know the expectation value of  $h(\theta)$  where  $\theta \sim f(\theta|\text{LRG} + \text{WMAP})$ , we can use the samples from  $f(\theta|\text{WMAP})$ . The weights are  $w(\theta) = \mathcal{L}_{\text{LRG}}(\theta)$ . Inference of  $h(\theta)$  becomes

$$E_{f(\theta|\text{LRG}+\text{WMAP})}[h(\theta)] \approx \frac{\sum h(\theta)\mathcal{L}_{\text{LRG}}(\theta)}{\sum \mathcal{L}_{\text{LRG}}(\theta)}, \theta \sim f(\theta|\text{WMAP})$$

We could generalize this procedure for many different data sets. For example, we may want to make inferences based on WMAP, LRG, current measurements of the Hubble rate, and expansion from Supernova. The posterior would be

$$f(\theta|\text{WMAP} + \text{LRG} + \text{Ho} + \text{Sne}) = f(\theta|\text{WMAP})f(\theta|\text{LRG})f(\theta|\text{Ho})f(\theta|\text{Sne}).$$

Importance sampling could use the samples from WMAP

$$\theta \sim f(\theta|\text{WMAP})$$

and use the weights

$$w(\theta) = f(\theta|\text{LRG})f(\theta|\text{Ho})f(\theta|\text{Sne}) \propto \mathcal{L}_{\text{LRG}}\mathcal{L}_{\text{Ho}}\mathcal{L}_{\text{Sne}}$$

to make inference on this posterior.

### 6.3 Obtaining Upper Bounds on Mass

If we apply importance sampling to most parameters, finding point or interval estimates is straight-forward.  $g(\theta) = \theta$  gives us the means. Intervals can be found by estimating the variance,  $g(\theta) = (\theta - E[\theta])^2$ . Then,  $\hat{\theta} \pm \hat{s}_\theta$  where  $\hat{\theta}$  is the estimate of the mean and  $\hat{s}_\theta$  is the estimate of the standard deviation. This interval estimate is a  $1 - \sigma$  confidence interval based on the assumption that the underlying distribution is normal.

The problem with this is that a normal distribution seems a poor model for the distribution of  $\sum m_\nu$ . For example, figure 6.1[4] shows the estimated distribution from MCMC sampling. While the distributions do appear bell-

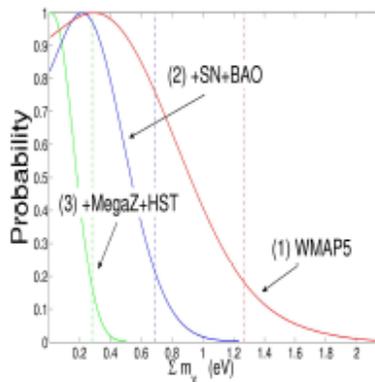


Figure 6.1: Estimated distribution of Neutrino mass for Various data sets[4]

shaped, they are cut-off at 0. A normal distribution would not be appropriate. Instead we assume a truncated normal: a normal distribution cut-off at the lower and upper points  $a$  and  $b$ . We allow  $b$  to be  $\infty$ . For the degenerate, normal hierarchy, and inverted hierarchy cases, we set  $a = \{0, 0.059, 0.099\}$  eV

as the lower bounds, respectively (see chapter 7).

How can we use this to get upper bounds on the neutrino hierarchy? This question is a little bit harder than it first seems. If we have a sample from MCMC, we would simply order the samples and find where 95% of them fall under. For the neutrino mass PDF, we propose a fit to the truncated normal distribution. This fit would be based on the sample mean and standard deviation. The upper bound would be found by solving for  $x$  when  $\text{Prob}(\sum m_\nu < x) = 0.95$ . Alternatively, one may use a non-parametric estimate based on the weighted histogram.

### 6.3.1 Summary of the Truncated Normal

Truncating any distribution causes some modifications to the probability distribution. For our case, the PDF becomes

$$\phi(x|\mu, \sigma) = \frac{f(x|\mu, \sigma)}{1 - F(a|\mu, \sigma)}$$

where  $f(x|\mu, \sigma)$  is the PDF for the untruncated normal distribution with population mean  $\mu$  and standard deviation  $\sigma$ ,  $F(a|\mu, \sigma)$  is the cumulative distribution function (CDF) which is  $F(x|\mu, \sigma) \propto \int_{-\infty}^x e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ , and  $a$  is the appropriate lower cut off. The CDF of a normal distribution is sometimes called the error function,  $\text{Erf}(x)$ , and can not be determined analytically. Therefore, the PDF for a truncated normal can only be found numerically.

Also,  $\mu$  and  $\sigma$  are no longer the mean and standard deviation of the distribution. The truncation modifies this. For our case, with a lower bound

of  $a$ , the mean and variance are

$$\mathbb{E}[x] = \mu + \sigma^2 \frac{f(a|\mu, \sigma)}{1 - F(a|\mu, \sigma)} = \mu + \sigma^2 \phi(a|\mu, \sigma)$$

and

$$\text{Var}[x] = \sigma^2 \left[ 1 - \frac{f(a|\mu, \sigma)}{1 - F(a|\mu, \sigma)} \left( \mu - a + \sigma^2 \frac{f(a|\mu, \sigma)}{1 - F(a|\mu, \sigma)} \right) \right],$$

respectively. See Appendix B for derivations. Solving for  $\mu$  and  $\sigma$  in terms of the mean and variance can not be done in closed-form and requires numerical techniques.

We also need the CDF of the truncated normal when we want to get upper bounds on the neutrino mass. The CDF for a truncated normal with lower bound at  $a$  is

$$\begin{aligned} \Phi(x|\mu, \sigma) &= \int_a^x \phi(x'|\mu, \sigma) dx' \\ &= \int_a^x \frac{f(x'|\mu, \sigma)}{1 - F(a|\mu, \sigma)} dx' \\ &= \frac{1}{1 - F(a|\mu, \sigma)} \int_a^x f(x'|\mu, \sigma) dx' \\ &= \frac{1}{1 - F(a|\mu, \sigma)} \left( \int_{-\infty}^x f(x'|\mu, \sigma) dx' - \int_{-\infty}^a f(x'|\mu, \sigma) dx' \right) \\ &= \frac{F(x|\mu, \sigma) - F(a|\mu, \sigma)}{1 - F(a|\mu, \sigma)} \end{aligned}$$

For a 95% upper bound, what we want to find is  $U$  such that

$$\begin{aligned} \Phi(U|\mu, \sigma) &= 0.95 \\ \frac{F(U|\mu, \sigma) - F(a|\mu, \sigma)}{1 - F(a|\mu, \sigma)} &= \\ F(U|\mu, \sigma) &= F(a|\mu, \sigma) + 0.95(1 - F(a|\mu, \sigma)) \end{aligned}$$

which most statistical software packages can solve.

### 6.3.2 Estimation of the Truncated Normal

Our current situation is this. We can use importance sampling to estimate the sample mean and variance, denoted  $\bar{x}$  and  $s^2$ . However, we do not have estimates on  $\mu$  and  $\sigma$  and therefore can not get an estimate on the 95% upper bound. What we want are ways to estimate  $\mu$  and  $\sigma$  based on  $\bar{x}$  and  $s^2$ . The two most straight-forward estimators are method of moments (MOM) and maximum likelihood estimators (MLE). MOMs are relatively straight forward; set the sample mean,  $\bar{x}$ , equal to the population mean,  $E[x]$ , and similarly for the standard deviation.

$$\bar{x} = \hat{\mu} + \hat{\sigma}^2 \frac{f(a|\hat{\mu}, \hat{\sigma})}{1 - F(a|\hat{\mu}, \hat{\sigma})}$$

and

$$s^2 = \hat{\sigma}^2 \left[ 1 - \frac{f(a|\hat{\mu}, \hat{\sigma})}{1 - F(a|\hat{\mu}, \hat{\sigma})} \left( \hat{\mu} - a + \hat{\sigma}^2 \frac{f(a|\hat{\mu}, \hat{\sigma})}{1 - F(a|\hat{\mu}, \hat{\sigma})} \right) \right]$$

with the hats meaning the estimators for the parameters. Just like there was no closed form solution for  $\mu$  and  $\sigma$ , there is no closed form for  $\hat{\mu}$  and  $\hat{\sigma}$ . This problem is a set of 2 nonlinear equations that must be solved numerically.

The MLEs are defined as the values of  $\mu$  and  $\sigma$  that maximizes the

likelihood. The log-likelihood for our truncated normal distribution is

$$\begin{aligned}
\ell \equiv \log \mathcal{L} &= \log \left( \prod_i^N \phi(x_i | \mu, \sigma) \right) \\
&= \log \left[ \left( \frac{1}{1 - F(a | \mu, \sigma)} \right)^N \prod_i^N f(x_i | \mu, \sigma) \right] \\
&= \log \left[ \left( \frac{1}{1 - F(a | \mu, \sigma)} \right)^N \left( \frac{1}{2\pi\sigma^2} \right)^{N/2} \exp \left( \frac{-1}{2\sigma^2} [(N-1)s^2 + N(\bar{x} - \mu)^2] \right) \right] \\
&= -N \log(1 - F(a | \mu, \sigma)) - \frac{N}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} [(N-1)s^2 + N(\bar{x} - \mu)^2]
\end{aligned}$$

We are left to numerically optimize  $\ell$  for parameters  $\mu$  and  $\sigma$  with  $s$  and  $\bar{x}$  known from importance sampling.

### 6.3.3 Making Bounds on the Neutrino Mass

We now have all the statistical tools needed. The random variable  $x$  in the previous subsections is now  $\sum m_\nu$ . We can use importance sampling to obtain the mean and variance of  $\sum m_\nu$  from a set of cosmological data.  $\mu$  and  $\sigma$  are estimated numerically using MOM or MLE. Then one solves for  $\sum m_\nu$  in

$$F(\sum m_\nu | \mu, \sigma) = F(a | \mu, \sigma) + 0.95(1 - F(a | \mu, \sigma))$$

to find the upper bound at the 95% level.

## 6.4 Testing Importance Sampling

In this section, we verify the use of importance sampling. First we test whether the truncated normal distribution is a good fit to the data. In other

words, we examine whether the bounds from a fitted truncated normal match the bounds from a full MCMC analysis. Next we test if importance sampling gives us the same bounds as a full MCMC analysis on the joint posterior. All numerical routines were done on Matlab®. Evaluations of the PDF and CDF for the untruncated normal distribution are also done in Matlab®.

#### 6.4.1 Testing the Truncated Normal Fitting

We compare the upper bounds from WMAP 7[6] with the upper bounds assuming a truncated normal distribution. Since we are checking only the validity of the truncated normal, we do no importance sampling. We simply use the mean and variance of the WMAP 7 chains from the  $\Lambda$ CDM +  $\sum m_\nu$  model. We then calculate the 95% upper limit using both MOMs and MLEs. If we obtain the same upper bounds and if the estimates on  $\mu$  and  $\sigma$  seem reasonable, then we will conclude that the truncated normal is a good description of the distribution for  $\sum m_\nu$ . As discussed in section 6.3.2, the MOMs require us to solve a set of non-linear equations. The MLEs require a numerical optimization routine. The non-linear equations are solved with a Levenberg-Marquardt algorithm[118][119]. The optimization routine uses a Nelder-Mead simplex search algorithm[120].

The results are in table 6.1. All parameters are in units of eV. WMAP places an upper bound of  $\sum m_\nu < 1.3\text{eV}$  at the 95% confidence level. The MOM estimation has a slightly lower bound and very different estimate of  $\mu$ . This estimate seems poor when we consider the actual shape of the distribu-

tion. Fig. 6.1 does not seem to indicate a negative  $\mu$ . MLEs have effectively the same bound as the WMAP case. The estimates of  $\mu$  and  $\sigma$  also seems reasonable when compared to the estimate of the distribution(see the red curve in Fig. 6.1).

Table 6.1: Estimation of  $\mu$ ,  $\sigma$  and  $\sum m_\nu$  based on MCMC chain

	MOM	MLE
$\hat{\mu}$	-1.272	0.395
$\hat{\sigma}$	0.901	0.513
$\sum m_\nu$	1.213	1.299

Note that this only tests the validity of the truncate normal. In other words, we are checking if fitting to a truncated normal distribution is reasonable. The use of MLEs do seem to give us the correct results while MOMs do not. In the rest of this paper, we use the MLE estimation of  $\mu$  and  $\sigma$ . The next subsection checks the importance sampling technique.

#### 6.4.2 Testing Importance Sampling

We test importance sampling for three cases.

- $\Lambda$ CDM weighting with  $H_0$ .
- $\Lambda$ CDM weighting with LRG.
- $\Lambda$ CDM +  $\sum m_\nu$  weighting with LRG and  $H_0$ .

We chose these 3 sets based on a few aspects. All three cases are done by Komatsu et. al.[6] using MCMC so we can check our results. The first case is

the simplest. We do not have to do any theoretical calculation of the CMB or matter power spectrum. It is a simple weighting. The second case tests the weighting of LRG without adding massive neutrinos. Finally, the third set tests importance sampling for models with massive neutrinos. To get intervals on all the cosmological parameters except  $\sum m_\nu$ , we assume a Gaussian distribution. Therefore, we can not obtain different upper and lower bounds on interval estimates. MCMC estimates are done non-parametrically and therefore can produce asymmetric intervals. However, this asymmetry appears small in current cosmological inferences. In tables 6.2, 6.3, and 6.4, the units of  $H_0$  are (km/s)/Mpc, the units of  $\sum m_\nu$  are eV, and  $A$  is written in factors of  $10^{-9}$ .

For the first case, the simple weighting is done by using the measured Hubble rate of  $74.2 \pm 3.6$  from Riess et. al.[61] as discussed in section 3.1.4. We check the means and intervals we obtain from weighting and compare them with the results of the MCMC done by WMAP[6]. Table 6.2 shows our results. This is the  $\Lambda$ CDM model; neutrinos are treated massless. Table 6.2 shows almost identical inferences. For this simple case, importance sampling and MCMC yield equivalent results.

Table 6.2: Importance Sampling of WMAP7 by  $H_0$  for  $\Lambda$ CDM Model

	Imp. Samp.	MCMC
$A$	$2.39 \pm 0.10$	$2.39 \pm 0.10$
$n_s$	$0.967 \pm 0.013$	$0.967 \pm 0.013$
$\Omega_c h^2$	$0.1088 \pm 0.0048$	$0.1087 \pm 0.0047$
$\Omega_b h^2$	$0.02270 \pm 0.00057$	$0.02270^{+0.00055}_{-0.00056}$
$H_0$	$72.1 \pm 2.1$	$72.1^{+2.1}_{-2.0}$

Table 6.3 shows the results of the second case. The weighting is done by calculating the likelihood of LRG as explained in section 3.3.3. An algorithm to calculate this likelihood is publicly available at [117]. In this case we use CAMB along with the modifications by Reid et. al.[100] to compute the theoretical halo power spectrum. Likelihood calculations are done as discussed in previous sections. As in the first case, we see almost identical results. The differences between importance sampling and MCMC do appear larger than in table 6.2. However, these estimates are statistically the same. The differences can probably be attributed to sampling variation and differences in parametric/non-parametric estimations.

Table 6.3: Importance Sampling of WMAP7 by LRG for  $\Lambda$ CDM Model

	Imp. Samp.	MCMC
$A$	$2.489 \pm 0.100$	$2.491 \pm 0.099$
$n_s$	$0.960 \pm 0.013$	$0.959 \pm 0.013$
$\Omega_c h^2$	$0.1161 \pm 0.0037$	$0.1161 \pm 0.0036$
$\Omega_b h^2$	$0.02244 \pm 0.00055$	$0.02247^{+0.00053}_{-0.00054}$
$H_0$	$68.8 \pm 1.7$	$68.7 \pm 1.6$

Table 6.4 shows the results of the third case. This model does include degenerate massive neutrinos. Table 6.3 and table 6.4 show similar results. While there are slight differences, the estimates are statistically equivalent. This last case was important since it was the only case with massive neutrinos. Based on these 3 test cases, we are confident that importance sampling does obtain similar bounds as those obtained from MCMC. The truncated normal distribution does appear to be a good fit to the  $\sum m_\nu$  distribution.

Table 6.4: Importance Sampling of WMAP7 by  $H_0$  & LRG for  $\Lambda$ CDM +  $\sum m_\nu$  Model

	Imp. Samp.	MCMC
$A$	$2.437 \pm 0.096$	$2.442 \pm 0.096$
$n_s$	$0.966 \pm 0.013$	$0.965 \pm 0.013$
$\Omega_c h^2$	$0.1133 \pm 0.0034$	$0.1136 \pm 0.0035$
$\Omega_b h^2$	$0.02266 \pm 0.00053$	$0.02265 \pm 0.00054$
$H_0$	$68.5 \pm 1.8$	$68.5 \pm 1.8$
$\sum m_\nu <$	0.45	0.44

## Chapter 7

# Current and Forecasted Bounds on Neutrino Mass

As was discussed in section 4.4, the effect of hierarchy is roughly an order of magnitude smaller on the CMB than on the matter power spectrum. This motivates our importance sampling. Rather than do a full MCMC analysis with neutrino hierarchy, we importance sample based on the CMB samples for the degenerate neutrino case. For instance, the MCMC results for WMAP are publicly available[117]. We can then weight them according to either the normal or inverted neutrino hierarchy case with data from galaxy surveys or other cosmological data. For the current galaxy survey, we use the SDSS LRG as in section 6.4.2.

Neutrino oscillation experiments are sensitive to mass differences squared. The latest results are

$$\Delta m_{\odot}^2 \equiv |m_b^2 - m_a^2| = 7.65_{-0.60}^{+0.69} \times 10^{-5} \text{eV}^2$$

and

$$\Delta m_A^2 \equiv |m_c^2 - m_a^2| = 2.40_{-0.33}^{+0.35} \times 10^{-3} \text{eV}^2$$

at the 99.73% confidence level. The former comes from solar oscillations and the later from atmospheric oscillations [121]. Here  $m_{a,b,c}$  are set mass eigen-

states. We denote  $m_{1,2,3}$  as the mass eigenstates from lightest to heaviest, respectively. This notation differs from the convention of the Particle Data Group. For the hierarchies, we need to relate the  $\sum m_\nu$  to the individual mass eigenstates. In general

$$\sum m_\nu = m_1 + \sqrt{m_1^2 + \Delta m^2} + \sqrt{m_1^2 + a^2} \quad (7.1)$$

where  $a^2 = \Delta m_\odot^2 + \Delta m_A^2$ .  $\Delta m^2$  depends on which hierarchy we desire. For a normal hierarchy,  $\Delta m^2 = \Delta m_\odot^2$  while for the inverted hierarchy,  $\Delta m^2 = \Delta m_A^2$ . Equation 7.1 is solvable; we have one unknown and one equation. However,  $m_1$  can not be solved analytically. For a given  $\sum m_\nu$ , we employ a root-finding algorithm; specifically, we use a Newton-Raphson/Bisection hybrid algorithm from Numerical Recipes[116]. With regard to importance sampling, we assume a lower truncated normal distribution for the  $\sum m_\nu$ . Based on equation 7.1, we set the truncations points for the normal and inverted hierarchy at 0.058511 and 0.098754, respectively.

For the lightest eigenstate, one may wish to know if the mass is consistent with zero. In other words, one may wish to know if the lightest mass eigenstate is actually massless. As seen in equation 7.1, equating  $m_1$  from  $\sum m_\nu$  can not be done analytically. Therefore, it would be mathematically difficult to derive the PDF of  $m_1$  based on the known PDF of  $\sum m_\nu$ . For this work, we assume a truncated normal distribution for  $m_1$  with truncation at 0. Note that this is not based on any mathematical reasoning. Formally, one must transform the PDF of  $\sum m_\nu$  based on equation 7.1. However, this seems

difficult to accomplish. Therefore, we assume a truncated normal distribution for  $m_1$  for the same reasons we assumed a truncated normal for  $\sum m_\nu$ . Parameters in all tables in this chapter follow the convention in chapter 6.  $A$  is written in factors of  $10^{-9}$ .  $H_0$  is in units of (km/s)/(Mpc). Both  $\sum m_\nu$  and  $m_1$  are in units of eV.

## 7.1 Bounds from Current Data

For current data, we use the same combination of data as table 6.4. Table 7.1 shows our results from importance sampling. Figures 7.1 and 7.2

Table 7.1: Importance Sampling of WMAP7 by  $H_0$  & LRG for  $\Lambda$ CDM +  $\sum m_\nu$  Model

	Deg.	Normal	Invert
$A$	$2.437 \pm 0.096$	$2.439 \pm 0.083$	$2.451 \pm 0.090$
$n_s$	$0.966 \pm 0.013$	$0.967 \pm 0.013$	$0.966 \pm 0.013$
$\Omega_c h^2$	$0.1133 \pm 0.0034$	$0.1140 \pm 0.0034$	$0.1139 \pm 0.0033$
$\Omega_b h^2$	$0.02266 \pm 0.00053$	$0.02264 \pm 0.00044$	$0.02263 \pm 0.00044$
$H_0$	$68.5 \pm 1.8$	$68.4 \pm 1.7$	$68.3 \pm 1.7$
$\sum m_\nu <$	0.45	0.43	0.45
$m_1 <$	0.15	0.14	0.14

are the estimates of the probability densities for  $\sum m_\nu$  and  $m_1$ . These figures are just the truncated normal PDF's for the estimated  $\mu$  and  $\sigma$ . The estimates of  $\mu$  and  $\sigma$  for the truncated normal was discussed in section 6.4.1. The peaks of the PDF's in figure 7.1 are quite different. The degenerate case has a  $\mu$  well above 0. The variance is also quite broad relative to the hierarchy case. Both hierarchy cases have truncations to the right of their untruncated peaks. The variance of the normal case appears slightly broader than the inverted case.

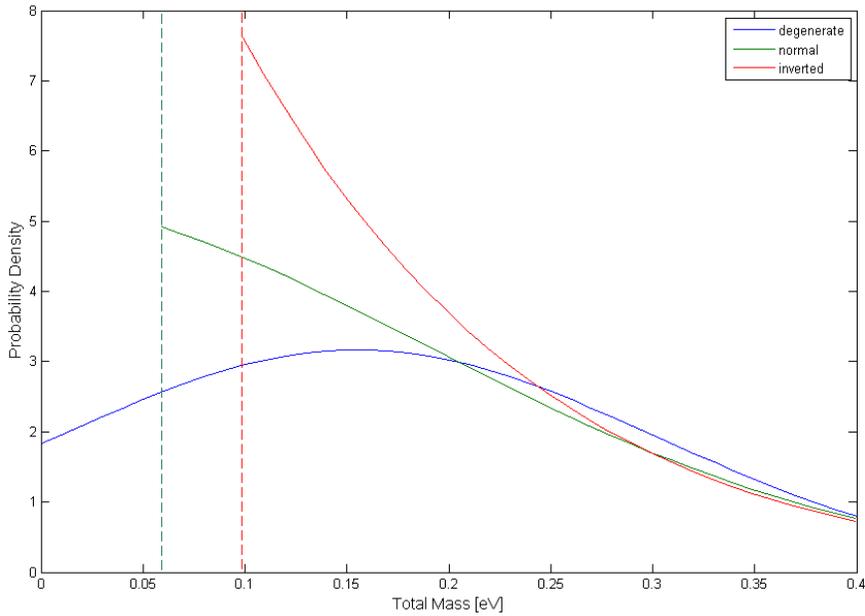


Figure 7.1: Probability Density of  $\sum m_\nu$  for WMAP + SDSS LRG +  $H_0$

The PDFs of  $m_1$  are quite similar. The truncation is not as large an effect here. Interestingly enough, the PDF of the inverted case is quite similar to the degenerate case. Even with the distinct PDFs in figure 7.1, the upper bounds on the total mass are roughly the same.

## 7.2 Effect of Future CMB Measurements

The most precise measurement of the CMB was done by the WMAP probe. Its data helped usher in the era of precision cosmology. The next generation space probe is Planck[122]. Planck was launched in 2009. At the time of this paper, the first year analysis of the primary signal had not been

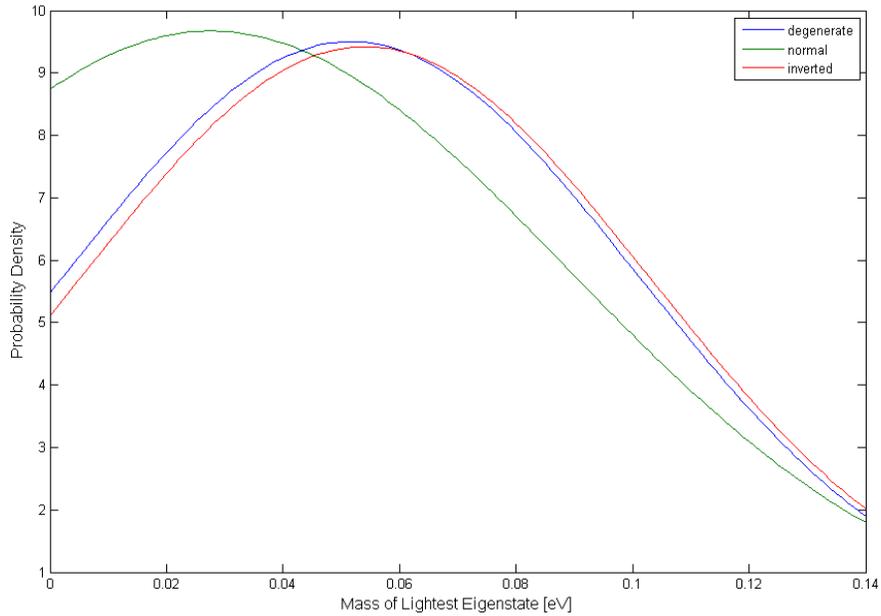


Figure 7.2: Probability Density of  $m_1$  for WMAP + SDSS LRG +  $H_0$

completed. The Planck data are expected to measure up to the third acoustic peak with high precision. The statistical analysis should be dominated by cosmic variance up to this scale. Another probe, CMBPol[123], is planned for after Planck. Its primary goal is the measurement of the polarization with a focus on extracting information about inflation. Forecasting has been done on both Planck and CMBPol by [1]. Table 7.3 shows the experimental set up of both these probes.

Earlier in section 4.4, we argued that the neutrino mass hierarchy has little to no effect on the CMB power spectrum. Therefore, we can forecast the parameter estimates from the degenerate massive neutrino models. For

Table 7.2: Experimental Set-ups[1]

Experiment	Channels	FWHM	Noise/pixel $10^{-6}$	$f_{\text{sky}}$
Planck	70	14'	4.7	0.85
	100	10'	2.5	
	143	7.1'	2.2	
CMBPol	150	5.6'	0.037	0.72

our purposes, we set the following means and uncertainties based on [1]. The mean values were chosen to roughly fit the current parameter estimates to the degenerate model. The mean value of  $\sum m_\nu$  was chosen to be lower than both minimums of normal and inverted hierarchy. This choice is rather arbitrary and would be an interesting case if it were true.

Table 7.3: Forecasted Errors on Parameters from Planck and CMBPol[1]

Experiment	Mean	Planck Error	CMBPol Error
$\Omega_b h^2$	0.0227	0.00014	0.000033
$\Omega_{CDM} h^2$	0.110	0.0017	0.00071
$n_s$	0.963	0.0034	0.0016
$A$	2.44	0.032	0.016
$H_0$	68.5	0.53	0.12
$\sum m_\nu$	0.03	0.13	0.07

To illustrate the impact of better CMB measurements, we importance sample on the current LRG data. We can then compare the estimates on the neutrino mass and see how the future CMB probes improve the current estimates. Table 7.4 shows the estimates of the cosmological parameters. The values statistically agree with the forecasted values. This result is expected since the precision of the parameters is dependent on CMB. Also, we would not expect a current probe of LSS to have more effect than a future CMB probe.

Even with the current SDSS LRG data, the upper bounds on the absolute mass have shrunk. These upper bounds are now at the level of [4]. Also, the differences in the upper bounds for the total mass between different hierarchy schemes is approaching 10%. These values are still above the minimum  $\sum m_\nu$  values for both normal and inverted hierarchies. Figures 7.3 and 7.4 show the

Table 7.4: Importance Sampling of Planck forecasts by LRG

	Deg.	Normal	Invert
$A$	$2.435 \pm 0.034$	$2.438 \pm 0.032$	$2.439 \pm 0.032$
$n_s$	$0.9637 \pm 0.0033$	$0.9635 \pm 0.0035$	$0.9634 \pm 0.0034$
$\Omega_c h^2$	$0.1103 \pm 0.0016$	$0.1103 \pm 0.0017$	$0.1105 \pm 0.0016$
$\Omega_b h^2$	$0.02270 \pm 0.00012$	$0.02269 \pm 0.00014$	$0.02271 \pm 0.00013$
$H_0$	$68.43 \pm 0.60$	$68.33 \pm 0.54$	$68.41 \pm 0.53$
$\sum m_\nu <$	0.230	0.258	0.278
$m_1 <$	0.077	0.080	0.084

estimated PDFs of the total mass and lightest mass eigenstate. From figure 7.3, the shapes of the PDFs here are similar to the WMAP case. One difference is the general reduction in the variance of the PDFs. Another difference is that the peaks of the  $m_1$  PDFs are closer to 0.

Table 7.5 and figures 7.5 & 7.6 show the same results as above but for the CMBPole forecasts. Again, the precision of the CMBPol experiment shrinks the upper-bound. The differences between the degenerate case and hierarchy cases are approaching 20%. However, these bounds are nevertheless impressive when we consider we are using current galaxy surveys.

The PDFs for  $\sum m_\nu$  and  $m_1$  are very similar to the Planck case. Note that these estimates assume that the chosen hierarchy is the correct model.

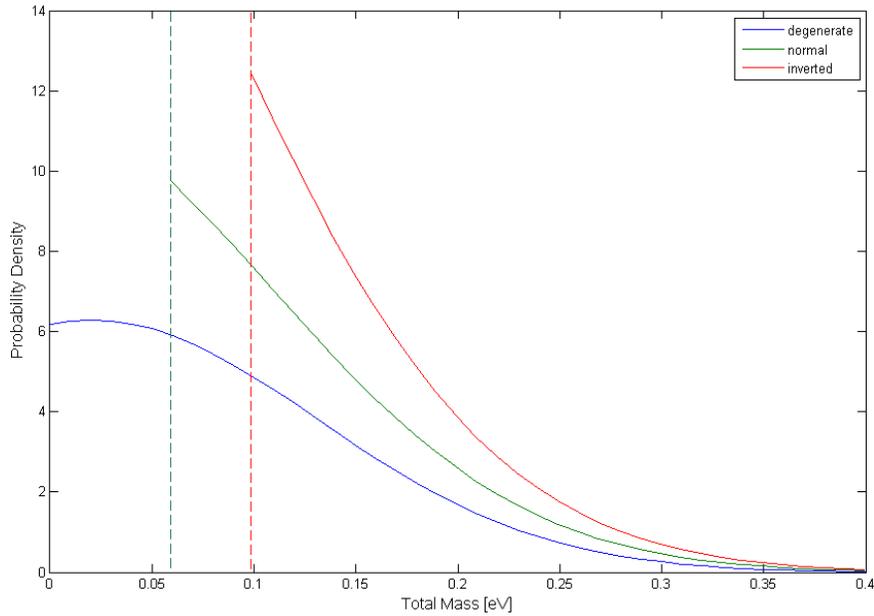


Figure 7.3: Probability Density of  $\sum m_\nu$  for Planck + LRG

We have not checked how well they fit the actual data.

### 7.3 Effect of Hubble Constant

When inferencing from the CMB alone, the Hubble constant is a degenerate parameter. Tighter bounds from direct measurements of  $H_0$  should result in tighter bounds on other parameters. As mentioned in section 3.1.4, we can expect the errors to reach around 2% in future measurements. At this level, one must be careful on the selection of the forecasted value. Based on the CMB, the median values for  $H_0$  are consistently under 70 for massive neutrino models. Recent direct measures of the  $H_0$  seem to prefer median values

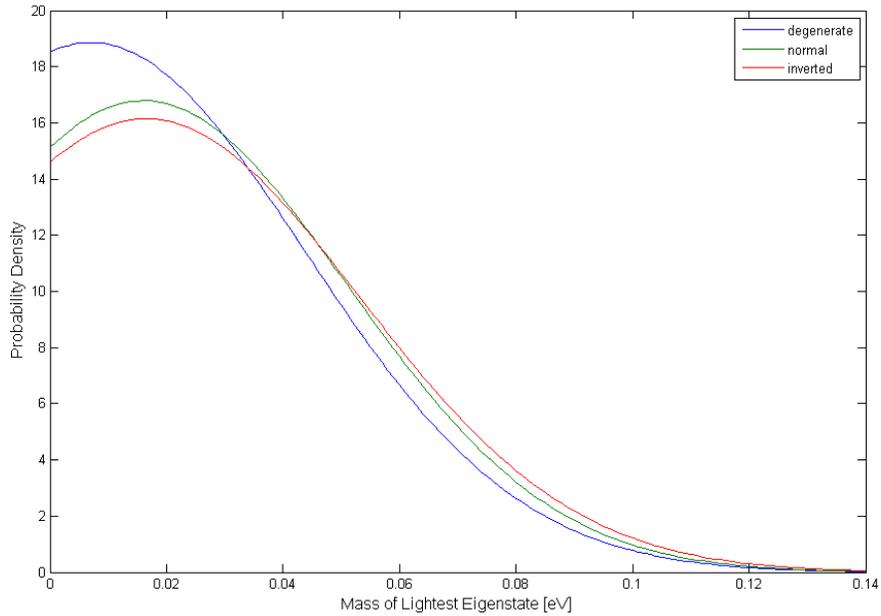


Figure 7.4: Probability Density of  $m_1$  for Planck + LRG

around 72. If the errors from both the CMB and  $H_0$  decrease to forecast values and their median values stay the same as today, then these two probes could be in disagreement. If this were to occur, then we would rule out the massive neutrino model with a cosmological constant. If one were very secure in the fact that neutrinos are massive and treated properly, then one could use this as evidence against a cosmological constant.

For our purposes, we set a forecasted measurement of  $H_0 = 71.5 \pm 1.5$ . At this value, the Planck forecasted values should still be within uncertainty. Figures 7.7 and 7.8 show the estimated PDFs of the total mass and lightest eigenstate. These estimates of the PDF are very similar to the Planck only

Table 7.5: Importance Sampling of CMBPol forecasts by LRG

	Deg.	Normal	Invert
$A$	$2.438 \pm 0.017$	$2.439 \pm 0.016$	$2.442 \pm 0.016$
$n_s$	$0.9633 \pm 0.0016$	$0.9631 \pm 0.0016$	$0.9631 \pm 0.0016$
$\Omega_c h^2$	$0.11006 \pm 0.00067$	$0.11004 \pm 0.00065$	$0.11010 \pm 0.00073$
$\Omega_b h^2$	$0.022702 \pm 0.000029$	$0.022698 \pm 0.000033$	$0.022700 \pm 0.000034$
$H_0$	$68.51 \pm 0.14$	$68.49 \pm 0.12$	$68.49 \pm 0.13$
$\sum m_\nu <$	0.147	0.173	0.184
$m_1 <$	0.049	0.050	0.047

case. Table 7.6 shows the parameter estimates. Estimates of the upper bound of the total neutrino mass differ a little less than 10%. These estimates are almost the same as those down without  $H_0$ . The small error already on  $H_0$  from Planck relative to direct measurements weights the sample closer to the Planck level. There is a slight rise in the median value but well within statistical variance. The differences in the neutrino bound are only seen in the degenerate case.

Table 7.6: Importance Sampling of Planck forecasts by  $H_0$  Forecast & LRG

	Deg.	Normal	Invert
$A$	$2.433 \pm 0.033$	$2.436 \pm 0.032$	$2.440 \pm 0.031$
$n_s$	$0.9636 \pm 0.0034$	$0.9632 \pm 0.0034$	$0.9633 \pm 0.0034$
$\Omega_c h^2$	$0.1106 \pm 0.0016$	$0.1102 \pm 0.0016$	$0.1105 \pm 0.0015$
$\Omega_b h^2$	$0.02272 \pm 0.00012$	$0.02269 \pm 0.00014$	$0.02272 \pm 0.00014$
$H_0$	$68.87 \pm 0.58$	$68.68 \pm 0.49$	$68.76 \pm 0.50$
$\sum m_\nu <$	0.219	0.258	0.278
$m_1 <$	0.073	0.080	0.083

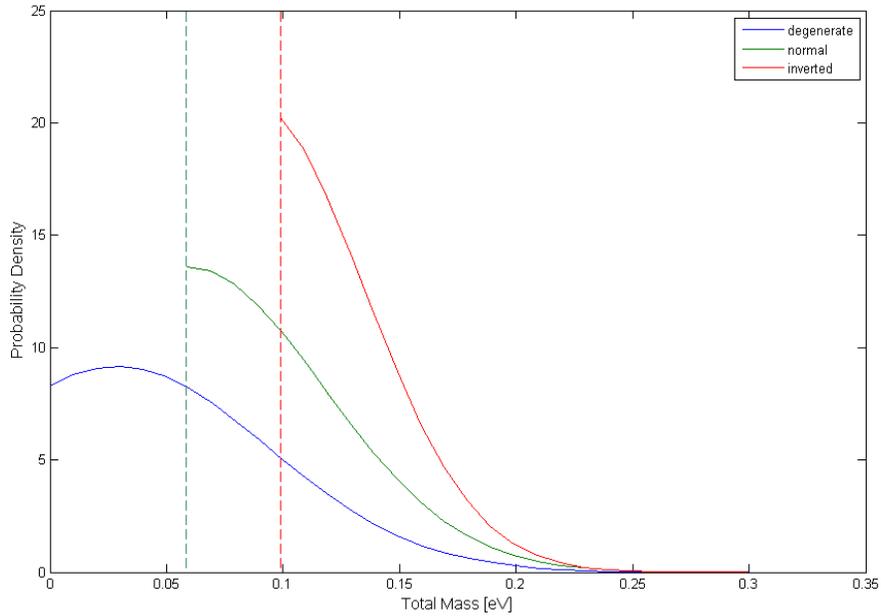


Figure 7.5: Probability Density of  $\sum m_\nu$  for CMBPol + LRG

## 7.4 Future Galaxy Surveys

In the previous section, we examined the effects of future CMB probes on the current SDSS LRG survey. Rather than follow a particular survey, we forecast the galaxy survey power spectrum to find what value is suitable to see neutrino mass hierarchy. To this end, we assume a simplistic likelihood with a multivariate normal distribution with diagonal covariance matrix. This analysis is equivalent to a  $\chi^2$  goodness-of-fit. To estimate the error, we refer to equation 3.15, the forecasted errors in galaxy surveys. First, we need to determine the parameters of this future galaxy survey so that the error bars are lower than the differences caused by neutrino mass hierarchy. An ambitious

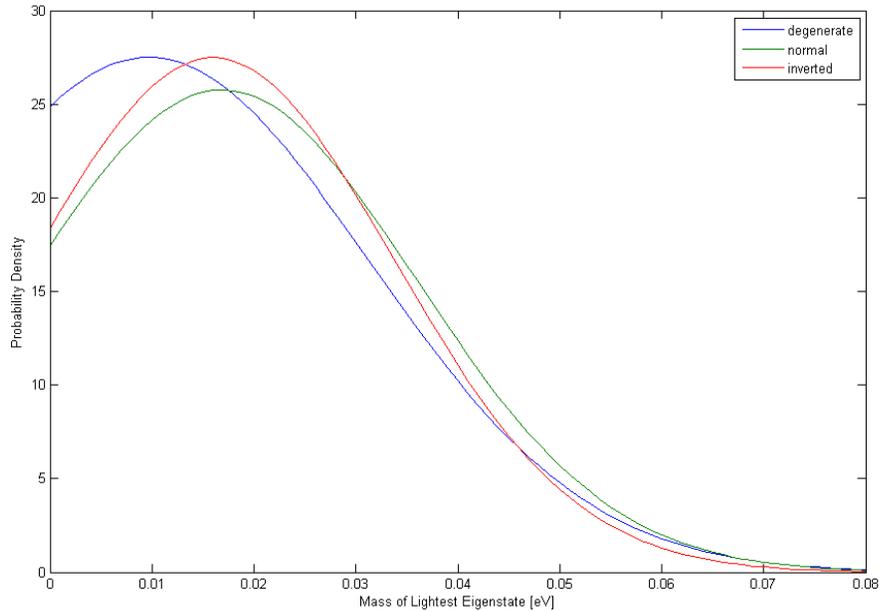


Figure 7.6: Probability Density of  $m_1$  for CMBPol + LRG

goal would be errors below 0.1%. We assume a galaxy survey of roughly 15,000 square degrees based on proposed observations such as BigBOSS project on SDSS stage IV[124] and JDEM project using WFIRST[125]. This area would roughly be twice that of SDSS LRG DR7. We also assume a center red shift of 1 with width around 0.4. Based on these choices, we estimate a volume survey of  $20h^{-3}\text{Gpc}^3$ . We assume  $\bar{n} \approx 1$ . We restrict the likelihood calculation to scale lengths 0.02 to 0.2  $h/\text{Mpc}$ . Figure 7.9 shows the estimate percent error for a survey of this kind.

We simulate the matter power data by assuming true parameter values and then calculating the power spectrum. For all the non-neutrino parameters,

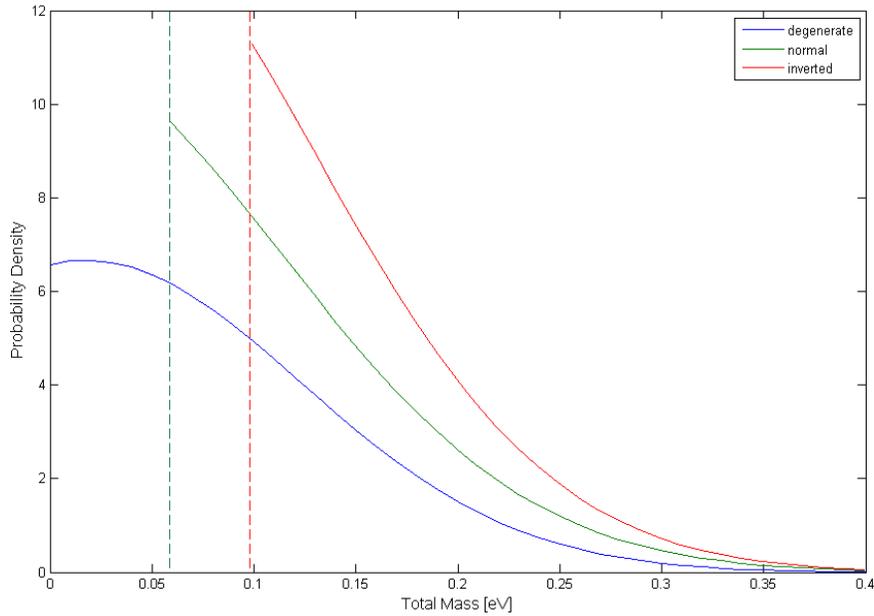


Figure 7.7: Probability Density of  $\sum m_\nu$  for Planck +  $H_0$  + LRG

we assume the central values in table 7.3 We find that the high precision of this kind of survey would make the importance sampling technique a poor tool. The motivation for importance sampling was the precision data of the CMB. Currently the CMB is a more precise observation that can place bounds on all the parameters. However, based on our work, the precision of this fictitious survey would make this tool ineffective. At this level, even a CMBPol experiment would not provide a necessary precision. Even with forecasted error values half of CMBPol led to only a few points in the importance sampling space. In other words, the samples from future CMB experiments would have too much variation with respect to a matter power spectrum of this precision.

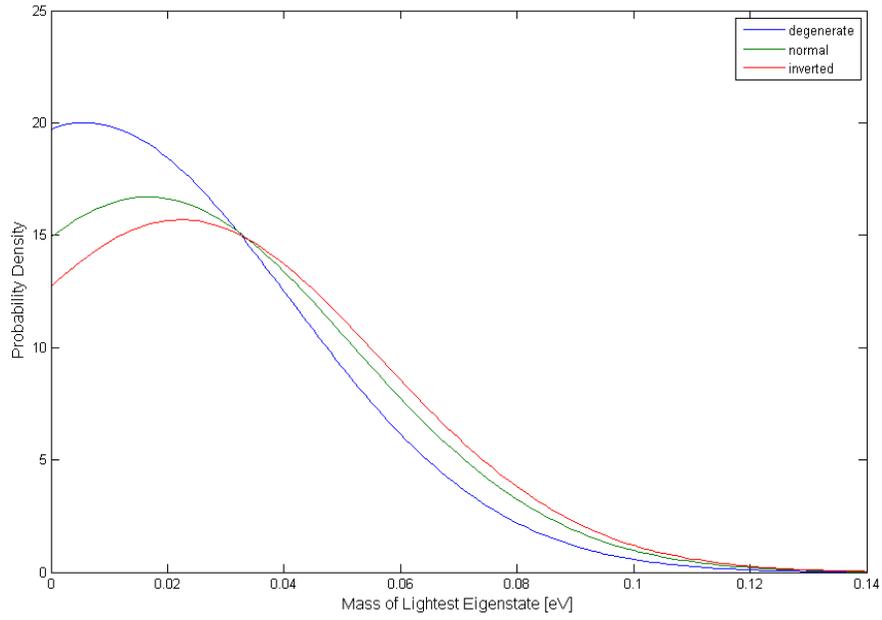


Figure 7.8: Probability Density of  $m_1$  for Planck +  $H_0$  + LRG

Most of the sample values would yield a very small likelihood beyond the precision of computation. In this case, one would choose an MCMC algorithm to sample the matter power spectrum.

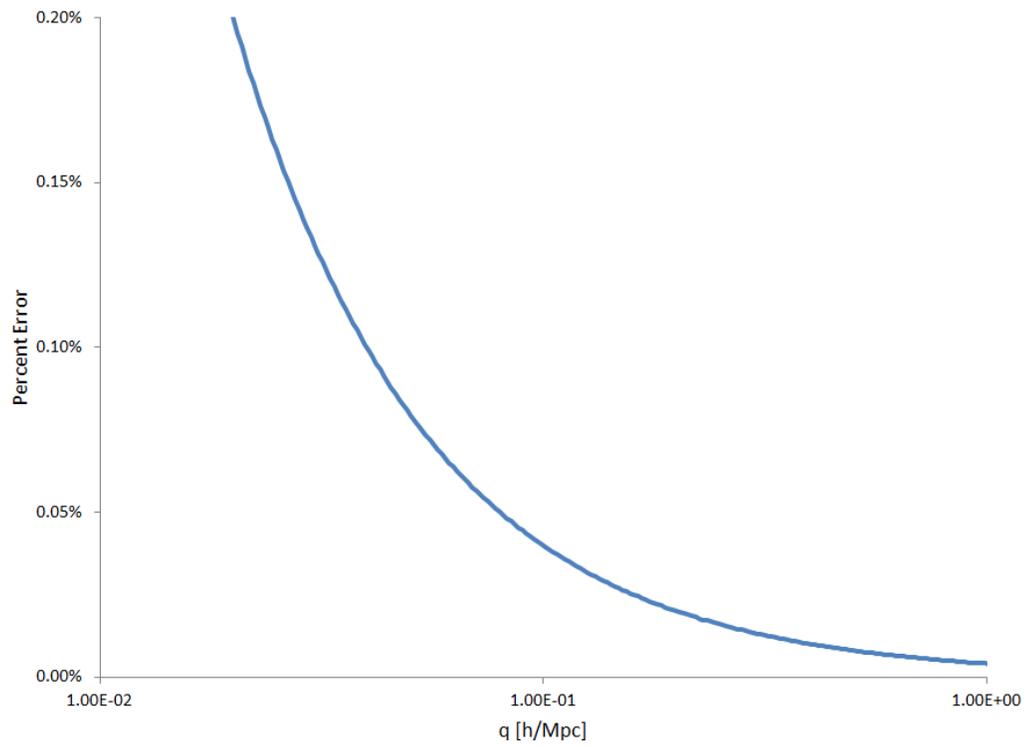


Figure 7.9: Forecasted Errors of Matter Power Spectrum

# Chapter 8

## Summary

In chapters 2 and 3, we reviewed the theoretical cosmological equations and how they are connected to our observations. We emphasized our review on the source of the radiation hierarchy. The three observables we examined were the Hubble constant, CMB, and LSS. In chapter 4, we examined the effects of massive neutrinos on these same observables. A qualitative explanation of these effects was given that focused on the matter power spectrum. We examined quantitatively the effect of the neutrino hierarchy on both the LSS and CMB. We found that the effect on the CMB was an order of magnitude lower than the matter power spectrum. These effects were also much smaller than the cosmic variance. Therefore, we concluded that neutrino hierarchy effects must be observed through LSS observations. The effects on the matter power spectrum were found to be roughly 0.1 – 0.2%. Since we are dealing with sub-percent level effects, we investigated the accuracy of the Boltzmann code in chapter 5. For the CAMB code, we found that the number of radiation hierarchy terms kept during integration was the key factor in terms of accuracy. At default settings, the error peaks at 0.9%. We found that settings of `boost = 1.5` and `lboost = 2.5` would be sufficient for these levels. In chapter 6, we discussed the use of importance sampling in inferencing the neutrino masses.

Since the CMB is insensitive to neutrino hierarchy, we can sample from the CMB likelihood based only on the degenerate neutrino case. Then importance sampling can weight these points in the parameter space based on the LSS likelihood. We verify that this technique acquires the same bounds on the neutrino mass as the conventional MCMC. In chapter 7, we implemented this technique to current and future probes. We used forecasted samples of the parameters based on the Planck and CMBPol experimental setups. While increased precision of CMB probes reduced the upper bound on neutrino mass, it could not determine the hierarchy alone. We determine the properties of a future galaxy survey necessary to observe the effects of neutrino hierarchy.

As stated before, the effect of the neutrino hierarchy is smaller on the CMB than the LSS. Our study indicated effects on the order of 0.10% on the LSS and effects on the order of 0.01% on the CMB. In addition, the effects on the CMB are well below the cosmic variance. This result motivated our importance sampling technique. We found that this technique to be much more efficient than a brute-force MCMC sampler. However, the importance sampling technique required a statistical model for the  $\sum m_\nu$ . We found that a truncated normal distribution fitted well. Specifically, we found the bounds from an assumed truncated normal to be in good agreement with non-parametric bounds. Our implementation of importance sampling took advantage of the CMB's insensitivity to neutrino hierarchy and reduced the amount of computation time. Once a sample of the CMB posterior is done, then the weight calculation reduced to calculations of the LSS likelihood. Our technique would

reduce the burden of the original sampling of the CMB posterior by allowing one to work with only degenerate neutrinos. Unlike MCMC techniques, once the CMB sample is performed, we no longer have to worry about convergence. The resultant inferences from our technique were equivalent to the common MCMC methods.

Based on our work, we concluded that the improved precision of CMB probes will most likely decrease the upper bounds on the neutrino mass. In addition, the differences in these bounds differ around 10% based on the neutrino hierarchy. In the case of Planck, the difference between the degenerate bound and the inverted bound reached 20%. We conclude that future bounds on the neutrino mass must take into account the hierarchy if one wishes for precision better than 10%. Upper bounds that fall below the minimum mass in the inverted hierarchy may be reached with future CMB and LSS probes. While it might be tempting to rule out the inverted hierarchy in this case, the effect of the neutrino hierarchy on these bounds needs to be taken into account.

We also investigated the desired precision of a galaxy survey to detect the neutrino mass. A survey of 15,000 square degrees and mean redshift 1 should come close to the type of precision needed to detect neutrino hierarchy differences. However, at these precision levels, the importance sampling technique is not useful. A measure of the matter power spectrum at this precision level outweighs the CMB precision level. Even a CMBpol level probe will not reach the precision of this galaxy survey. However, we note that this error

analysis was done simplistically and realistic errors could be higher.

In this work, we have focused on inferring the mass bounds of neutrinos. We do not discuss which neutrino mass model fits the data best. What we have done is examine the effect of the models on these upper bounds. One may wish to statistically find which model has a larger probability with respect to the data. These statistical techniques fall under the category of model selection. Since model selection techniques are more precarious than inferencing techniques, we do not implement them here. Bayes factors is one method that could be used.

## Appendices

# Appendix A

## Statistical Properties

Here we list any statistical theorems we used without discussion. We refer readers to the texts Casella & Berger[77] and Gelman et. al.[78] for more details. The page numbers to the relevant text are noted below.

### **The sum of normal random variables yields a normal random variable**

This statement can easily be verified by moment generating functions.

The moment generating function (MGF) is defined as

$$M(t) = E[e^{xt}]$$

where  $x$  is the random variable. The MGF of a normal distribution with mean and variance  $\mu$  and  $\sigma^2$  is  $e^{\mu t + \sigma^2 t^2 / 2}$ . The MGF for a sum of  $x$ 's is then

$$M(t) = \prod_i E[e^{x_i t}] = e^{t \sum \mu + t^2 / 2 \sum \sigma^2}.$$

Therefore the sum of normal random variables follows a normal distribution with mean equal to the individual  $\mu$ 's and  $\sigma^2$ , respectively. See page 214 in Casella & Berger.

### Distribution of sample variance is $\chi^2$

Theorem 5.3.1 on page 218 in Casella and Berger states that

$$\frac{(n-1)s^2}{\sigma^2} \sim \chi_{n-1}^2.$$

Note that we have assumed that  $\sigma^2$  is a known parameter. A  $\chi^2$  distribution has expected value equal to the number of degrees of freedom. Its variance is also the number of degrees of freedom. The support of a  $\chi^2$  distribution is positive real values with the degrees of freedom restricted to positive integers.

### The marginalization of normal hierarchy model

We define the hierarchy model as follows. We have data  $x$  from a normal distribution with mean  $\theta$  and variance  $\sigma^2$ . We then let  $\theta$  come from a normal distribution with mean  $\mu$  and variance  $\tau^2$ . In our example,  $\mu = 0$ . We often want the marginal distributions over this joint posterior distribution. Equation 5.1 in Gelman et. al. shows the form of  $f(\tau|x)$ . In this case

$$f(\tau|x) \propto \left(\frac{1}{\sigma^2 + \tau^2}\right)^{-N/2} \prod \exp\left(-\frac{x}{2(\sigma^2 + \tau^2)}\right)$$

In terms of  $\tau$ , this is a complicated probability but it has the same functional form as a normal distribution for  $x$  with mean 0 and variance  $\sigma^2 + \tau^2$ . See pages 134 - 136 for details.

## Appendix B

### Derivation of Mean and Variance of Truncated Normal

The mean or expected value is defined as

$$E[x] = \int_{\mathcal{X}} x \phi(x|\theta) dx$$

where  $\phi(x|\theta)$  is the probability density function (PDF),  $\theta$  are the parameters of the PDF, and  $\mathcal{X}$  is the support or domain of the PDF. The PDF of the truncated normal with lower cut-off  $a$  is

$$\phi(x|\mu, \sigma) = \frac{f(x|\mu, \sigma)}{1 - F(a|\mu, \sigma)}$$

where  $f(x|\mu, \sigma)$  and  $F(a|\mu, \sigma)$  are the PDF and cumulative distribution function (CDF) of the normal distribution, respectively.

## B.1 Mean

The expected value can be derived as so.

$$\begin{aligned}
 \mathbb{E}[x] &= \int_a^\infty x \phi(x|\mu, \sigma) dx \\
 &= \int_a^\infty x \frac{f(x|\mu, \sigma)}{1 - F(a|\mu, \sigma)} dx \\
 &= \frac{1}{1 - F(a|\mu, \sigma)} \int_a^\infty x f(x|\mu, \sigma) dx \\
 &= \frac{1}{1 - F(a|\mu, \sigma)} \int_a^\infty \frac{x}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx
 \end{aligned}$$

Now standardize the normal distribution. In other words, let  $z = \frac{x-\mu}{\sigma}$  thus  $dz = \frac{1}{\sigma} dx$ .

$$\begin{aligned}
 &\frac{1}{1 - F(a|\mu, \sigma)} \int_a^\infty \frac{x}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx \\
 &= \frac{1}{1 - F(a|\mu, \sigma)} \int_{\frac{a-\mu}{\sigma}}^\infty \frac{z\sigma + \mu}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz \\
 &= \frac{1}{1 - F(a|\mu, \sigma)} \left( \int_{\frac{a-\mu}{\sigma}}^\infty \frac{z\sigma}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz + \frac{\mu}{\sqrt{2\pi}} \int_{\frac{a-\mu}{\sigma}}^\infty e^{-\frac{z^2}{2}} dz \right) \\
 &= \frac{1}{1 - F(a|\mu, \sigma)} \left( \int_{\frac{a-\mu}{\sigma}}^\infty \frac{z\sigma}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz + \frac{\mu}{\sqrt{2\pi}} \int_{\frac{a-\mu}{\sigma}}^\infty e^{-\frac{z^2}{2}} dz \right)
 \end{aligned}$$

Let's work with the first integration term and then come back to the second term. The first term can be done with a simple transformation.

$$\int_{\frac{a-\mu}{\sigma}}^\infty \frac{z\sigma}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz$$

Now let  $u = z^2/2$  thus  $du = z dz$ . Then

$$\begin{aligned}
 \int_{\frac{a-\mu}{\sigma}}^{\infty} \frac{z\sigma}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz &= \\
 &= \frac{\sigma}{\sqrt{2\pi}} \int_{\frac{1}{2}\left(\frac{a-\mu}{\sigma}\right)^2}^{\infty} e^{-u} du \\
 &= -\frac{\sigma}{\sqrt{2\pi}} e^{-u} \Big|_{\frac{1}{2}\left(\frac{a-\mu}{\sigma}\right)^2}^{\infty} \\
 &= \sigma^2 \left( \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{a-\mu}{\sigma}\right)^2} \right) \\
 &= \sigma^2 f(a|\mu, \sigma)
 \end{aligned}$$

where again  $f(a|\mu, \sigma)$  is the PDF of the normal distribution at the lower cut-off.

As for the second integral term

$$\frac{\mu}{\sqrt{2\pi}} \int_{\frac{a-\mu}{\sigma}}^{\infty} e^{-\frac{z^2}{2}} dz$$

we will be able to write this in terms of the CDF.

$$\begin{aligned}
 \frac{\mu}{\sqrt{2\pi}} \int_{\frac{a-\mu}{\sigma}}^{\infty} e^{-\frac{z^2}{2}} dz &= \\
 &= \frac{\mu}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{z^2}{2}} dz - \frac{\mu}{\sqrt{2\pi}} \int_{-\infty}^{\frac{a-\mu}{\sigma}} e^{-\frac{z^2}{2}} dz \\
 &= \mu - \mu F\left(\frac{a-\mu}{\sigma} | 0, 1\right) \\
 &= \mu(1 - F(a|\mu, \sigma))
 \end{aligned}$$

In the last line we have used the fact that  $F(a|\mu, \sigma) = F\left(\frac{a-\mu}{\sigma} | 0, 1\right)$  where again  $F(x|\mu, \sigma)$  is the CDF of the normal distribution. This property can be

easily seen by the definition and standardization.

$$\begin{aligned}
 F(x|\mu, \sigma) &= \\
 &= \int_{-\infty}^x \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx \\
 &= \int_{-\infty}^{\frac{x-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz \\
 &= F\left(\frac{x-\mu}{\sigma}, 0, 1\right)
 \end{aligned}$$

Back to the expectation value we have

$$\begin{aligned}
 E[x] &= \\
 &= \frac{1}{1 - F(a|\mu, \sigma)} \left( \int_{\frac{a-\mu}{\sigma}}^{\infty} \frac{z\sigma}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz + \frac{\mu}{\sqrt{2\pi}} \int_{\frac{a-\mu}{\sigma}}^{\infty} e^{-\frac{z^2}{2}} dz \right) \\
 &= \frac{1}{1 - F(a|\mu, \sigma)} (\sigma^2 f(a|\mu, \sigma) + \mu(1 - F(a|\mu, \sigma))) \\
 &= \sigma^2 \frac{f(a|\mu, \sigma)}{1 - F(a|\mu, \sigma)} + \mu
 \end{aligned}$$

## B.2 Variance

The Variance is defined as

$$\text{Var}[x] = E[(x - E[x])^2] = E[x^2] - (E[x])^2$$

We are left to evaluate  $E[x^2]$ .

$$\begin{aligned}
E[x^2] &= \\
&= \int_a^\infty x^2 \phi(x|\mu, \sigma) dx \\
&= \int_a^\infty x^2 \frac{f(x|\mu, \sigma)}{1 - F(a|\mu, \sigma)} dx \\
&= \frac{1}{1 - F(a|\mu, \sigma)} \int_{b=\frac{a-\mu}{\sigma}}^\infty \frac{(\sigma z + \mu)^2}{\sqrt{2\pi}} e^{-z^2/2} dz \\
&= \frac{1}{1 - F(a|\mu, \sigma)} \int_b^\infty \frac{1}{\sqrt{2\pi}} (\sigma^2 z^2 + 2\mu\sigma z + \mu^2) e^{-z^2/2} dz
\end{aligned}$$

We are now left with three integrals to perform, effectively. The first integral is in the form of the second moment. However, we are not integrating over the entire real line so this is of no use. This integral can be done through integration by parts (done two times). The result is

$$\begin{aligned}
\frac{\sigma^2}{\sqrt{2\pi}} \int_b^\infty z^2 e^{-z^2/2} dz &= \\
&= \frac{\sigma^2}{\sqrt{2\pi}} \left( -z e^{-z^2/2} \Big|_b^\infty + \int_b^\infty e^{-z^2/2} dz \right) \\
&= \frac{\sigma^2}{\sqrt{2\pi}} \left( b e^{-b^2/2} + \sqrt{2\pi} [1 - F(a|\mu, \sigma)] \right) \\
&= \sigma^2 (a - \mu) f(a|\mu, \sigma) + \sigma^2 (1 - F(a|\mu, \sigma))
\end{aligned}$$

The second integral can be done with a substitution  $u = z^2/2$ . This results in

$$\begin{aligned}
\frac{2\mu\sigma}{\sqrt{2\pi}} \int_b^\infty z e^{-z^2/2} dz &= \\
&= \frac{2\mu\sigma}{\sqrt{2\pi}} \left( e^{-b^2/2} \right) \\
&= 2\mu\sigma^2 f(a|\mu, \sigma)
\end{aligned}$$

The third integral is simply a function of the CDF

$$\frac{\mu^2}{\sqrt{2\pi}} \int_b^\infty e^{-z^2/2} dz = \mu^2(1 - F(a|\mu, \sigma))$$

The first part of this is canceled when we solve for the variance. After combining all these terms, we are left with

$$\begin{aligned} \text{Var}[x] &= \mu^2 + \frac{1}{1 - F(a|\mu, \sigma)} [\sigma^2(a - \mu) + 2\mu\sigma^2] + \sigma^2 \\ &\quad - \left[ 2\mu\sigma^2 \frac{f(a|\mu, \sigma)}{1 - F(a|\mu, \sigma)} + \sigma^4 \left( \frac{f(a|\mu, \sigma)}{1 - F(a|\mu, \sigma)} \right)^2 \right] \\ &= \sigma^2 \left[ 1 + \frac{f(a|\mu, \sigma)}{1 - F(a|\mu, \sigma)} \left( a - \mu - \sigma^2 \frac{f(a|\mu, \sigma)}{1 - F(a|\mu, \sigma)} \right) \right] \end{aligned}$$

# Appendix C

## Numerical Techniques

In this appendix we briefly review some of the numerical algorithms used in cosmology calculations. For full discussion, we refer readers to [116]. In the main body of this work, we mention some techniques without any discussion. In chapter 5, we discuss the numerical integration of ODEs in some detail so that the reader may understand the sources of error. This section of appendix begins with numerical integration of functions as these are needed in calculating the neutrino energy density. Next, we discuss further the numerical integration of ODES. Finally, the optimization routines used in chapter 6 are discussed.

### C.1 Numerical Integration of functions

We discuss the techniques used in CAMB to integrate the massive neutrino energy density, pressure and perturbations. These integrals are one-dimensional; therefore, no Monte Carlo techniques are used. These techniques are sometimes referred to as quadrature techniques. In short, integrals are approximated by summations

$$\int_a^b f(x)dx \approx \sum_i^N c_i f(x_i)$$

where  $c_i$  depends on the exact algorithm. The trapezoidal rule takes the average of the function at the end points.

$$\int_a^b f(x)dx \approx h \left( \frac{1}{2}f(a) + \frac{1}{2}f(b) \right)$$

where  $h = b - a$ . In effect, they are assuming a straight line between points  $a$  and  $b$  and calculating that area. The error of this algorithm is of order  $h^3 f''$ . The extended trapezoidal rule breaks the interval  $b - a$  into many intervals and then applies the same rule within these smaller intervals. This is called the extended trapezoidal rule or composite rule. The result is

$$\int_a^b f(x)dx \approx h \left( \frac{1}{2}f(a) + f(a+h) + f(a+2h) \cdots + f(b-h) + \frac{1}{2}f(b) \right)$$

where  $h = (b - a)/N$  and  $N$  is the number of intervals. The error is on order  $h^2(b-a)f''$ . This assumes equal intervals. This approximation can be furthered by adding the next terms in the Euler-Maclaurin summation formula.

The integrals over the phase-space results in improper integrals from 0 to  $\infty$ . Therefore, the above algorithm is not optimal. The improper integral can be broken into two pieces.

$$\int_a^\infty f(x)dx = \int_a^b f(x)dx + \int_b^\infty f(x)dx$$

For our purposes, we know  $f(x)$  is a function with a proper integral. All that needs to be done is to integrate from  $a$  to  $b$  using the extended trapezoidal rule and then calculate the last integral as

$$\int_b^\infty f(x)dx = \int_0^{1/b} \frac{1}{x^2} f(1/x)dx$$

In the case of the phase-space distribution,  $f(y) \rightarrow 0$  as  $y \rightarrow \infty, 0$ . Therefore, if one were to apply a trapezoidal rule to the above integral, one would only need to evaluate  $f(1/x)$  at  $x = 1/b$ . This is the method in CAMB.

## C.2 Numerical Integration of ODES

The bulk of the computation time involves ODE integration. ODE integration techniques are split between explicit and implicit methods. The difference is the numerical treatment of the first order derivatives. Explicit methods are also called “forward” methods since the values of the next time step are based entirely on the previous step. These methods tend to be easier to implement but sometimes require very small time steps to be stable. Implicit methods are also called “backward” methods since the next time step values are based on the previous step and the current step. This method requires non-linear solvers at each time step making them more complicated to implement than explicit methods. Implicit methods have the advantage of being stable therefore allowing for larger time steps. We only discuss explicit techniques.

The simplest explicit method of integrating ODES is the Euler’s method. First, the finite difference approximation to a first order derivative is

$$\frac{dy}{dt} \approx \frac{y(t+h) - y(t)}{h}$$

for some finite  $h$ . For the ODE

$$\frac{dy}{dt} = f(t, y)$$

The Euler method would approximate this as

$$y_{n+1} = y_n + f(t_n, y_n) * h$$

where  $y_n = y(t_n)$ ,  $t_n = t_0 + nh$ , and  $t_0$  is the initial time. In chapter 5, we discussed the generalization of Euler methods to multiple steps: the Runge-Kutta method. However, one may wish to have a flexible time step in order to ensure a proper error. This desire leads to RK pair methods. RK pair methods calculate RK coefficients twice: two sequential ordered pairs. These are then subtracted in order to reveal the error. This estimate of the error can be used to control the time steps. The CAMB code uses a RK 45 pair based on Verner's method[126]. A review of Verner's pair along with other pair methods can be seen in [127].

### C.3 Optimization and Nonlinear Solvers

Here, we discuss the Newton-Raphson and bisection methods that were used to solve for  $m_1$ . The bisection method finds the roots of a function by finding increasingly smaller intervals. Each iteration cuts the interval by half, then finds which interval still contains the root. The Newton-Raphson method uses the Taylor series approximation to find an approximation to the root. For each iteration, the root is estimated at  $y = x - f(x)/f'(x)$  where  $y$  is new point. This method requires knowledge of the first derivative.

Now, we discuss the optimization and nonlinear solver routines implemented within Matlab. The optimization routine is a direct search method.

In other words, this optimization algorithm requires no derivative calculation where other algorithms like Newton-Raphson do. The particular routine is the Nelder-Mead simplex search algorithm based on geometry[120] . For a single iteration, the points in the parameter space are set to the vertices of the simplex. The center of gravity of this simplex is used to calculate new iteration points by various geometric operations. For the nonlinear solver, we used the Levenberg-Marquardt algorithm[118][119]. This algorithm is a particular kind of least-squares curve fitter. This iterative process picks a new value based on the derivative much similar to a steepest decent algorithm.

## Bibliography

- [1] S. Galli *et al.*, Phys. Rev. **D82**, 123504 (2010), 1005.3808.
- [2] M. Tegmark and M. Zaldarriaga, Phys.Rev. **D66**, 103508 (2002), astro-ph/0207047.
- [3] <http://bccp.lbl.gov/personnel/anze.html>.
- [4] S. A. Thomas, F. B. Abdalla, and O. Lahav, Phys. Rev. Lett. **105**, 031301 (2010), 0911.5291.
- [5] D. Larson *et al.*, Astrophys. J. Suppl. **192**, 16 (2011), 1001.4635.
- [6] E. Komatsu *et al.* (WMAP), Astrophys. J. Suppl. **192**, 18 (2011), 1001.4538.
- [7] A. Lewis, A. Challinor, and A. Lasenby, Astrophys. J. **538**, 473 (2000), astro-ph/9911177.
- [8] H. Bondi, *Cosmology* (Cambridge University Press, 1952).
- [9] J. Yadav, S. Bharadwaj, B. Pandey, and T. R. Seshadri, Monthly Notices of the Royal Astronomical Society **364**(2), 601 (2005), ISSN 1365-2966, URL <http://dx.doi.org/10.1111/j.1365-2966.2005.09578.x>.

- [10] A. Friedman, Zeitschrift fr Physik A Hadrons and Nuclei **10**, 377 (1922), ISSN 0939-7922, 10.1007/BF01332580, URL <http://dx.doi.org/10.1007/BF01332580>.
- [11] A. Friedmann, Zeitschrift fr Physik A Hadrons and Nuclei **21**, 326 (1924), ISSN 0939-7922, 10.1007/BF01328280, URL <http://dx.doi.org/10.1007/BF01328280>.
- [12] G. Lemaître, Annales de la Societe Scietifique de Bruxelles **47**, 49 (1927).
- [13] H. P. Robertson, Astrophysical Journal **82**, 284 (1935).
- [14] H. P. Robertson, Astrophysical Journal **83**, 187 (1936).
- [15] H. P. Robertson, Astrophysical Journal **83**, 257 (1936).
- [16] A. G. Walker, Proceedings of London Mathematical Society **42**, 90 (1937).
- [17] P. de Bernardis *et al.* (Boomerang), Nature **404**, 955 (2000), astro-ph/0004404.
- [18] A. Balbi *et al.*, Astrophys. J. **545**, L1 (2000), astro-ph/0005124.
- [19] D. N. Spergel *et al.* (WMAP), Astrophys. J. Suppl. **148**, 175 (2003), astro-ph/0302209.
- [20] C.-P. Ma and E. Bertschinger, The Astrophysical Journal **455**, 7 (1995).

- [21] V. F. Mukhanov, H. A. Feldman, and R. H. Brandenberger, Phys. Rept. **215**, 203 (1992).
- [22] S. Weinberg, *Cosmology* (Oxford University Press, USA, 2008), ISBN 0198526822, URL <http://www.amazon.com/exec/obidos/redirect?tag=citeulike07-20&path=ASIN/0198526822>.
- [23] G. F. Smoot, C. L. Bennett, A. Kogut, E. L. Wright, J. Aymon, N. W. Boggess, E. S. Cheng, G. de Amici, S. Gulkis, M. G. Hauser, G. Hinshaw, P. D. Jackson, *et al.*, Astrophysical Journal Letters **396**, L1 (1992).
- [24] S. Dodelson, *Modern Cosmology* (Academic Press, 2003).
- [25] E. M. Lifshitz, J. Phys. (USSR) **10**, 116 (1946).
- [26] P. J. Peebles and J. T. Yu, The Astrophysical Journal **162**, 815 (1970).
- [27] J. R. Bond and G. Efstathiou, The Astrophysical Journal **285**, L45 (1984).
- [28] U. Seljak and M. Zaldarriaga, Astrophys. J. **469**, 437 (1996), [astro-ph/9603033](http://arxiv.org/abs/astro-ph/9603033).
- [29] T. Padmanabhan, *Structure Formation in the Universe* (Cambridge University Press, 1993).
- [30] E. W. Kolb and M. S. Turner, *The Early Universe*, Frontiers in Physics (Addison-Wesley Publishing Company, 1990).

- [31] V. Mukhanov, *Physical Foundations of Cosmology* (Cambridge University Press, 2005).
- [32] E. Hubble, Proceedings of the National Academy of Science **15**, 168 (1929).
- [33] M. A. C. Perryman, L. Lindegren, J. Kovalevsky, E. Hoeg, U. Bastian, P. L. Bernacca, M. Crézé, F. Donati, M. Grenon, F. van Leeuwen, H. van der Marel, F. Mignard, *et al.*, aap **323**, L49 (1997).
- [34] B. Bassett and R. Hlozek, in *Dark Energy: Observational and Theoretical Approaches*, Edited by Pilar Ruiz-Lapuente. Published: Cambridge, UK ; New York by Cambridge University Press, 2010. ISBN 9780521518888 (*hardback*), p.246, edited by Ruiz-Lapuente, P. (Cambridge University Press, 2010), pp. 246–+.
- [35] J. A. Cardelli, G. C. Clayton, and J. S. Mathis, ApJ **345**, 245 (1989).
- [36] B. T. Draine, Annual Review of Astronomy & Astrophysics **41**, 241 (2003), arXiv:astro-ph/0304489.
- [37] A. G. Riess *et al.* (Supernova Search Team), Astron. J. **116**, 1009 (1998), astro-ph/9805201.
- [38] S. Perlmutter *et al.* (Supernova Cosmology Project), Astrophys. J. **517**, 565 (1999), astro-ph/9812133.
- [39] L. A. L. da Silva, Astrophysics and Space Science **202**, 215 (1993).

- [40] F. Hoyle and W. A. Fowler, ApJ **132**, 565 (1960).
- [41] J. Whelan and I. Iben, Jr., ApJ **186**, 1007 (1973).
- [42] S. Chandrasekhar, MNRAS **95**, 207 (1935).
- [43] S. A. Colgate and C. McKee, ApJ **157**, 623 (1969).
- [44] D. A. Howell *et al.*, Nature **443**, 308 (2006), astro-ph/0609616.
- [45] D. C. Leonard, ApJ **670**, 1275 (2007), 0710.3166.
- [46] M. M. Phillips, ApJL **413**, L105 (1993).
- [47] J. Guy, P. Astier, S. Nobili, N. Regnault, and R. Pain, A&A **443**(3), 781 (2005), URL <http://dx.doi.org/10.1051/0004-6361:20053025>.
- [48] S. Jha, A. G. Riess, and R. P. Kirshner, Astrophys. J. **659**, 122 (2007), astro-ph/0612666.
- [49] C. R. Burns *et al.* (2010), 1010.4040.
- [50] W. M. Wood-Vasey *et al.*, Astrophys. J. **689**, 377 (2008), 0711.2068.
- [51] M. Hicken *et al.*, Astrophys. J. **700**, 331 (2009), 0901.4787.
- [52] A. V. Filippenko, in *White dwarfs: cosmological and galactic probes*, edited by E. M. Sion, S. Vennes, & H. L. Shipman (2005), vol. 332 of *Astrophysics and Space Science Library*, pp. 97–133, arXiv:astro-ph/0410609.

- [53] H. A. Smith, *RR Lyrae stars* (Cambridge University Press, 1995).
- [54] A. S. Eddington, *The Observatory* **40**, 290 (1917).
- [55] J. R. Percy, *Understanding variable stars* (Cambridge University Press, 2007).
- [56] H. S. Leavitt, *Annals of Harvard College Observatory* **60**, 87 (1908).
- [57] H. S. Leavitt and E. C. Pickering, *Harvard College Observatory Circular* **173**, 1 (1912).
- [58] R. B. Tully and J. R. Fisher, *Astronomy and Astrophysics* **54**, 661 (1977).
- [59] J. Tonry and D. P. Schneider, *Astronomical Journal* **96**, 807 (1988).
- [60] W. L. Freedman, B. F. Madore, B. K. Gibson, L. Ferrarese, D. D. Kelson, S. Sakai, J. R. Mould, J. Robert C. Kennicutt, H. C. Ford, J. A. Graham, J. P. Huchra, S. M. G. Hughes, *et al.*, *The Astrophysical Journal* **553**(1), 47 (2001), URL <http://stacks.iop.org/0004-637X/553/i=1/a=47>.
- [61] A. G. Riess *et al.*, *Astrophys. J.* **699**, 539 (2009), 0905.0695.
- [62] C. Henkel *et al.* (2005), astro-ph/0503070.
- [63] W. L. Freedman and B. F. Madore, *Annual Review of Astronomy and Astrophysics* **48**, 673 (2010), 1004.1856.
- [64] G. F. Benedict *et al.*, *Astron. J.* **133**, 1810 (2007), astro-ph/0612465.

- [65] M. Hicken *et al.*, *Astrophys. J.* **700**, 1097 (2009), 0901.4804.
- [66] A. C. Layden, R. B. Hanson, S. L. Hawley, A. R. Klemola, and C. J. Hanley, *Astron. J.* **112**, 2110 (1996), astro-ph/9608108.
- [67] A. Dressler, *The Astrophysical Journal* **317**, 1 (1987).
- [68] M. Bonamente, M. K. Joy, S. J. LaRoque, J. E. Carlstrom, E. D. Reese, and K. S. Dawson, *The Astrophysical Journal* **647**, 25 (2006), arXiv:astro-ph/0512349.
- [69] S. H. Suyu *et al.*, *Astrophys. J.* **711**, 201 (2010), 0910.2773.
- [70] G. Gamow, *Physical Review* **74**, 505 (1948).
- [71] G. Gamow, *Nature* **162**, 680 (1948).
- [72] R. A. Alpher and R. C. Herman, *Physical Review* **74**, 1737 (1948).
- [73] A. A. Penzias and R. W. Wilson, *Astrophysical Journal* **142**, 419 (1965).
- [74] C. L. Bennett, A. J. Banday, K. M. Gorski, G. Hinshaw, P. Jackson, P. Keegstra, A. Kogut, G. F. Smoot, D. T. Wilkinson, and E. L. Wright, *Astrophysical Journal Letters* **464**, L1+ (1996), arXiv:astro-ph/9601067.
- [75] G. Hinshaw *et al.* (WMAP), *Astrophys. J. Suppl.* **170**, 288 (2007), astro-ph/0603451.
- [76] D. Baumann (2009), 0907.5424.

- [77] G. Casella and R. L. Berger, *Statistical Inference* (the Wadsworth Group, 2002).
- [78] A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin, *Bayesian Data Analysis* (Chapman & Hall/CRC, 2004).
- [79] N. Jarosik *et al.* (WMAP), *Astrophys. J. Suppl.* **170**, 263 (2007), astro-ph/0603452.
- [80] L. Page *et al.* (WMAP), *Astrophys. J. Suppl.* **148**, 39 (2003), astro-ph/0302214.
- [81] M. Tegmark, A. Taylor, and A. Heavens, *Astrophys. J.* **480**, 22 (1997), astro-ph/9603021.
- [82] D. J. Eisenstein, W. Hu, and M. Tegmark, *Astrophysical Journal* **518**, 2 (1999).
- [83] L. Perotto, J. Lesgourgues, S. Hannestad, H. Tu, and Y. Y. Y. Wong, *JCAP* **0610**, 013 (2006), astro-ph/0606227.
- [84] D. Munshi, P. Valageas, L. Van Waerbeke, and A. Heavens, *Phys. Rept.* **462**, 67 (2008), astro-ph/0612667.
- [85] D. H. Weinberg, R. Dave, N. Katz, and J. A. Kollmeier, *AIP Conf. Proc.* **666**, 157 (2003), astro-ph/0301186.
- [86] J. Huchra, M. Davis, D. Latham, and J. Tonry, *Astrophysical Journal Supplement Series* **52**, 89 (1983).

- [87] W. J. Percival *et al.* (The 2dFGRS), *Mon. Not. Roy. Astron. Soc.* **327**, 1297 (2001), astro-ph/0105252.
- [88] D. J. Eisenstein *et al.* (SDSS) (2011), 1101.1529.
- [89] Z. Ivezić *et al.* (2008), 0805.2366.
- [90] T. Abbott *et al.* (Dark Energy Survey) (2005), astro-ph/0510346.
- [91] G. J. Hill *et al.*, *ASP Conf. Ser.* **399**, 115 (2008), 0806.0183.
- [92] Y. B. Zel'Dovich, *Astron. Astrophys.* **5**, 84 (1970).
- [93] A. J. S. Hamilton (1997), astro-ph/9708102.
- [94] N. Kaiser, *Astrophysical Journal Letters* **284**, L9 (1984).
- [95] V. Springel *et al.*, *Nature* **435**, 629 (2005), astro-ph/0504097.
- [96] K. Heitmann, M. White, C. Wagner, S. Habib, and D. Higdon, *Astrophys. J.* **715**, 104 (2010), 0812.1052.
- [97] R. E. Smith *et al.* (The Virgo Consortium), *Mon. Not. Roy. Astron. Soc.* **341**, 1311 (2003), astro-ph/0207664.
- [98] S. Cole *et al.* (The 2dFGRS), *Mon. Not. Roy. Astron. Soc.* **362**, 505 (2005), astro-ph/0501174.
- [99] B. A. Reid *et al.*, *Mon. Not. Roy. Astron. Soc.* **404**, 60 (2010), 0907.1659.

- [100] B. A. Reid, D. N. Spergel, and P. Bode, *Astrophys. J.* **702**, 249 (2009), 0811.1025.
- [101] D. J. Eisenstein, H.-J. Seo, and M. White, *The Astrophysical Journal* **664**, 660 (2007), arXiv:astro-ph/0604361.
- [102] S. S. Gershtein and Y. B. Zel'Dovich, *Soviet Journal of Experimental and Theoretical Physics Letters* **4**, 120 (1966).
- [103] K. Sato and M. Kobayashi, *Progress of Theoretical Physics* **58**, 1775 (1977).
- [104] J. R. Bond, G. Efstathiou, and J. Silk, *Physical Review Letters* **45**, 1980 (1980).
- [105] S. Dodelson, E. Gates, and A. Stebbins, *Astrophys. J.* **467**, 10 (1996), astro-ph/9509147.
- [106] S. Bashinsky and U. Seljak, *Phys. Rev.* **D69**, 083002 (2004), astro-ph/0310198.
- [107] M. Shoji and E. Komatsu, *Phys. Rev.* **D81**, 123516 (2010), 1003.0942.
- [108] S. Hannestad and J. Brandbyge, *JCAP* **1003**, 020 (2010), 0910.4578.
- [109] S. Agarwal and H. A. Feldman (2010), 1006.0689.
- [110] S. Saito, M. Takada, and A. Taruya, *Phys. Rev.* **D83**, 043529 (2011), 1006.4845.

- [111] S. Hannestad, A. Mirizzi, G. G. Raffelt, and Y. Y. Y. Wong, JCAP **1008**, 001 (2010), 1004.0695.
- [112] F. Villaescusa-Navarro, J. Miralda-Escude, C. Pena-Garay, and V. Quilis (2011), 1104.4770.
- [113] C. Jose, S. Samui, K. Subramanian, and R. Srianand (2011), 1104.3714.
- [114] F. Marulli, C. Carbone, M. Viel, L. Moscardini, and A. Cimatti (2011), 1103.0278.
- [115] J. Lesgourgues and S. Pastor, Phys. Rept. **429**, 307 (2006), astro-ph/0603494.
- [116] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes 3rd Edition: The Art of Scientific Computing* (Cambridge University Press, New York, NY, USA, 2007), 3 ed., ISBN 0521880688, 9780521880688.
- [117] <http://lambda.gsfc.nasa.gov>.
- [118] K. Levenbger, Quarterly Applied Mathematics **2**, 164 (1944).
- [119] D. Marquardt, SIAM Journal Applied Mathematics **11**, 431 (1963).
- [120] J. C. Lagarias, J. C. Lagarias, J. A. Reeds, J. A. Reeds, M. H. Wright, M. H. Wright, P. E. Wright, and P. E. Wright, SIAM Journal of Optimization **9**, 112 (1996).

- [121] K. Nakamura *et al.* (Particle Data Group), J.Phys.G **G37**, 075021 (2010).
- [122] (2006), astro-ph/0604069.
- [123] J. Bock *et al.* (2009), 0906.1188.
- [124] D. J. Schlegel *et al.* (2009), 0904.0468.
- [125] N. Gehrels (2010), 1008.4936.
- [126] J. H. Verner, SIAM Journal on Numerical Analysis **15**, 772 .
- [127] J. H. Verner, SIAM Journal on Numerical Analysis **27**, 1332 .
- [128] M. Tegmark *et al.* (SDSS), Phys. Rev. **D74**, 123507 (2006), astro-ph/0608632.
- [129] D. J. Eisenstein *et al.* (SDSS), Astron. J. **122**, 2267 (2001), astro-ph/0108153.