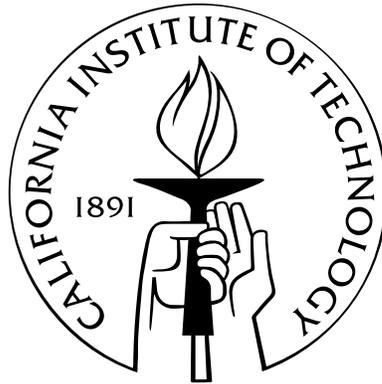


**Cosmological Consequences of Gravitation:**  
*Structure Formation and Gravitational Waves*

Thesis by  
Laura G. Book

In Partial Fulfillment of the Requirements  
for the Degree of  
Doctor of Philosophy



California Institute of Technology  
Pasadena, California

2012  
(Defended April 16, 2012)



I believe there is no philosophical high-road in science, with epistemological signposts. No, we are in a jungle and find our way by trial and error, building our road behind us as we proceed.

*—Max Born, 1943*

---

When we try to pick out anything by itself, we find it hitched to everything else in the universe.

*—John Muir, 1911*

# Acknowledgments

The work represented in this thesis was made possible by interactions with many people. I owe a great debt to my advisor, Dr. Andrew Benson, for interesting projects, many useful discussions, unstinting support, and an unusual level of understanding. I have also learned a great deal from Dr. Marc Kamionkowski, whom I thank for stimulating conversations as well as for providing many opportunities to grow as a cosmologist. I have greatly benefited from collaboration with many other talented scientists. I thank (in roughly chronological order) Drs. Annika Peter, Alyson Brooks, Fabio Governato, Éanna Flanagan, Tarun Souradeep, and Fabian Schmidt for their time, knowledge, and ideas, and for having helped to lead me along some very enjoyable lines of thought. I acknowledge the support of the NSF Graduate Research Fellowship Program, the Caltech Feynman Fellowship, and the Betty and Gordon Moore Foundation.

In my five years at Caltech, there have been a great number of friends that have helped me to relax, keep going, and keep having fun. The support of friends now scattered around the world from the University of Illinois and from University Laboratory High School in Urbana, IL, has also been indispensable. It is not the place here to list you all by name, but you recognize yourselves, and I thank you.

I would not be the scientist that I am today without the encouragement of my many excellent teachers at Uni High, and my mentors at the U of I. I specifically thank Drs. Charles Gammie and You-Hua Chu for showing me the beauty and fun of astrophysics research.

Finally, I have been very lucky to have a family that has constantly believed in me and supported me. I give my deep gratitude to my parents Ruth and Tom Book, my sister Lynn Book, and my grandparents Grace and William Schoedel and Marlene and Larry Book, for being there for me when I needed them. I also thank my in-laws for their kindness and support. Finally, my heartfelt thanks to my husband Esfandiar Alizadeh, for more than can be listed here.

# Abstract

This thesis contains work on four topics which fit into two broad areas of research: the quest to understand structure formation and through it the properties of the dark matter, and the search for primordial gravitational radiation. The first project details the effect of an accretion shock on the colors of satellites in galaxy clusters. A new model of ram pressure stripping including an accretion shock with variable radius is developed and implemented in the GALFORM semi-analytic model of galaxy formation. A comparison of this model with previous models and with observations indicates that current data is unable to discriminate between models, though future observations will be able to place stronger constraints on the role of ram pressure stripping in and around clusters.

Next, an analysis of the angular momentum evolution of dark matter particles in high-resolution N-body simulations of dark matter halos is presented. We find that individual particle angular momentum is not conserved, and also that the angular momentum of radial shells varies over the age of the Universe by up to factors of a few. These results have serious implications for the validity of current analytical models that assume angular momentum conservation.

Two methods for detecting the primordial gravitational wave (GW) background are then presented. Such a background, if detected, could greatly impact our understanding of the early universe. The first proposed method uses the apparent angular velocities of astrophysical objects induced by GWs, which may be detectable with upcoming astrometric missions such as the GAIA satellite. This work improves upon previous order-of-magnitude estimates, and presents a full calculation of the expected signal from a stochastic background of GWs.

The second method uses bipolar spherical harmonics decomposition, a formalism to characterize departures from statistical isotropy and Gaussianity, to quantify the expected lensing of the cosmic microwave background (CMB) and 21 cm radiation by GWs. The lensing of the CMB by GWs is found to not be detectable, but that of future 21 cm surveys could give a very high quality measurement of the primordial GW background.

# Contents

<b>Acknowledgments</b>	<b>iv</b>
<b>Abstract</b>	<b>v</b>
<b>List of Figures</b>	<b>ix</b>
<b>List of Tables</b>	<b>xi</b>
<b>1 General Introduction and Summary</b>	<b>1</b>
<b>2 The Role of Ram Pressure Stripping in the Quenching of Cluster Star Formation</b>	<b>4</b>
2.1 Introduction . . . . .	4
2.1.1 Analytical Prescriptions for Structure Formation . . . . .	6
2.1.2 Cosmological Implications of Galaxy Cluster Properties . . . . .	9
2.2 Simulations . . . . .	11
2.2.1 The GALFORM Semi-Analytic Model . . . . .	11
2.2.2 Implementation of Accretion Shocks . . . . .	13
2.3 Cluster Galaxy Properties with Different Models of Ram Pressure Stripping . . . . .	14
2.4 Comparison with Observations . . . . .	19
2.4.1 Comparison to CNOC1 Cluster Redshift Survey as Selected by Balogh et al. (2000) . . . . .	19
2.4.2 Comparison to Verdugo et al. (2008) Cluster Spectroscopy . . . . .	23
2.5 Discussion and Conclusions: Implications for Observations of Ram Pressure Stripping	24
<b>3 Angular Momentum Evolution in Dark Matter Halos</b>	<b>27</b>
3.1 Introduction . . . . .	27
3.1.1 Adiabatic Contraction Model . . . . .	29
3.2 Simulations . . . . .	32
3.2.1 Particle Subsets . . . . .	33
3.3 Halo Evolution . . . . .	34

3.3.1	Validity of the Adiabatic Approximation . . . . .	36
3.3.2	Angular Momentum Evolution . . . . .	37
3.3.3	Distribution of Angular Momentum . . . . .	41
3.4	Causes of Angular Momentum Evolution . . . . .	47
3.5	Discussion . . . . .	51
<b>4</b>	<b>Astrometric Effects of a Stochastic Gravitational Wave Background</b>	<b>53</b>
4.1	Introduction and Summary of Results . . . . .	53
4.1.1	The Stochastic Gravitational Wave Background . . . . .	53
4.1.2	High-Precision Astrometry . . . . .	54
4.1.3	Summary of Results . . . . .	56
4.2	Calculation of Astrometric Deflection in a Minkowski Background Spacetime . . . . .	59
4.2.1	Setting the Stage—Minkowski Calculation . . . . .	59
4.2.2	Photon Geodesic . . . . .	60
4.2.3	Boundary Conditions . . . . .	61
4.2.4	Perturbation to Observed Frequency . . . . .	63
4.2.5	Local Proper Reference Frame of Observer . . . . .	63
4.2.6	Observed Angular Deflection . . . . .	64
4.2.7	The Distant Source Limit . . . . .	66
4.3	Generalization to Cosmological Spacetimes . . . . .	67
4.3.1	Stationary Observers are Freely Falling . . . . .	67
4.3.2	Null Geodesic in the Conformal Metric . . . . .	67
4.3.3	Parallel Transport of Basis Vectors in FRW Background Spacetime . . . . .	68
4.3.4	Generalization of Angular Deflection Computation . . . . .	69
4.3.5	The Distant Source Limit . . . . .	69
4.4	Calculation of Angular Deflection Correlation Function . . . . .	70
4.4.1	Description of SGWB as a Random Process . . . . .	71
4.4.2	Power Spectrum of the Astrometric Deflections of the SGWB . . . . .	71
4.4.3	Basis Tensors and Their Symmetries . . . . .	72
4.4.4	Solving the General Integral . . . . .	73
4.4.5	Special Case: Coincidence . . . . .	75
4.5	Spectrum of Angular Deflection Fluctuations . . . . .	76
4.5.1	Overview . . . . .	76
4.5.2	Derivation . . . . .	78
4.5.2.1	EE correlation . . . . .	79
4.5.2.2	BB correlation . . . . .	81

<b>5</b>	<b>Bipolar Spherical Harmonic Decomposition</b>	<b>82</b>
5.1	Introduction . . . . .	82
5.1.1	Cosmic Microwave Background . . . . .	82
5.1.2	21 cm Radiation . . . . .	83
5.1.3	Weak Gravitational Lensing . . . . .	84
5.1.4	Previous Work with Bipolar Spherical Harmonics . . . . .	85
5.2	Review of Bipolar Spherical Harmonics . . . . .	86
5.2.1	Statistically Isotropic and Gaussian Maps . . . . .	86
5.2.2	Departures from Gaussianity/SI . . . . .	86
5.2.3	Odd-Parity Bipolar Spherical Harmonics . . . . .	87
5.3	Gravitational Lensing . . . . .	88
5.3.1	Gradient and Curl-Type Deflections . . . . .	88
5.3.2	Deflection Field from Metric Perturbations . . . . .	90
5.3.3	Lensing by Density (Scalar Metric) Perturbations . . . . .	92
5.3.4	Lensing by GWs (Tensor Metric Perturbations) . . . . .	94
5.3.5	Non-Cosmological Odd-Parity BiPoSH . . . . .	99
5.4	BiPoSHs as Probes of Parity Violation . . . . .	100
5.4.1	Correlation of Opposite-Parity Lensing Components . . . . .	100
5.4.2	Large-Angle CMB Polarization Spectra . . . . .	101
5.4.3	Parity-Violating Correlations from Chiral GWs . . . . .	103
5.4.4	Variance of $\phi$ -B and $\Omega$ -E Correlations . . . . .	105
5.4.5	Signal-to-Noise Ratio of Chiral GW Background Detection . . . . .	106
5.5	Lensing of 21 cm Fluctuations by Primordial Gravitational Waves . . . . .	108
5.6	Conclusions . . . . .	114
<b>A</b>	<b>Accretion Shock Calculation</b>	<b>116</b>
	<b>Bibliography</b>	<b>119</b>

# List of Figures

2.1	Stellar mass, gas mass, and gas fraction of cluster satellite galaxies with different models of ram pressure stripping . . . . .	15
2.2	Colors, star formation rates, and specific star formation rates for cluster satellite galaxies with different models of ram pressure stripping . . . . .	18
2.3	Comparison of GALFORM models with CNOC1 redshift survey . . . . .	20
2.4	Comparison of GALFORM models with (2) cluster spectroscopy . . . . .	22
3.1	The evolution of the specific halo angular momentum . . . . .	35
3.2	Energies and time scales of all particles within the virial radius at each timestep . . .	38
3.3	Energies and time scales of all particles within one tenth of the virial radius at each timestep . . . . .	39
3.4	Angular momentum evolution of Lagrangian mass shells . . . . .	42
3.5	Angular momentum evolution of Eulerian mass shells . . . . .	43
3.6	Angular momentum distribution at several redshifts for different particle subsets . . .	45
3.7	Evolution of angular momentum distribution in terms of Boltzmann fitting function .	46
3.8	Comparison of inner and outer halo torque timescales with halo dynamical time . . .	48
3.9	Total torque on individual particles compared with that due to the smooth halo and due to substructure . . . . .	49
4.1	Angular scale dependence of deflection power . . . . .	58
4.2	Angular separation dependence of the spatial deflection power . . . . .	76
5.1	Power spectrum of gradient-type modes of cosmic shear . . . . .	95
5.2	Power spectrum of curl-type modes of cosmic shear . . . . .	96
5.3	Cross-correlation of gradient-type modes of cosmic shear with magnetic-type CMB polarization . . . . .	103
5.4	Cross-correlation of curl-type modes of cosmic shear with electric-type CMB polarization	104
5.5	Power spectrum of curl-type lensing of 21 cm radiation at different redshifts compared with predicted noise . . . . .	109

5.6	Redshift of GW lenses responsible for curl-mode signal at different angular scales . . .	111
-----	--	-----

# List of Tables

3.1	Simulated galaxy properties . . . . .	33
3.2	Ratios of bin angular momentum at accretion and $z = 1$ ( $L_a$ and $L_1$ ) to that at $z = 0$	41
4.1	First 10 nonzero multipole coefficients $\alpha_l^{EE}$ of the deflection power . . . . .	59

## Chapter 1

# General Introduction and Summary

Theoretical cosmology consists, at its heart, of thinking about the Universe at large, and as such does not appear to differ fundamentally from other philosophical pursuits. Indeed, until fairly recently cosmology was just that, a branch of philosophy, a subject for intellectual debate detached from external data. However, cosmology has undergone a complete revolution, and now can be considered an exact science. All of this has come about due to advanced technology for observing the sky and recording the light that we see in different frequency bands, providing an unprecedented quantity and accuracy of data. It is thus an extremely privileged time to be a theoretical cosmologist, with a wealth of observations from which to piece together the history and fundamental laws of the Universe.

The standard picture of the Universe that we have uncovered, though far from complete and far from settled, provides a relatively elegant and straightforward place to start. In this picture, the Universe first went through a period of exponential growth, referred to as inflation. This prediction elegantly solves the problem of how the cosmic microwave background (CMB) can be so remarkably uniform over large scales that in the simplest model of big bang cosmology could not have been causally connected when it was emitted. Inflation solves this problem through the simple hypothesis that in the early Universe there existed a period in which the expansion was accelerated. We can see that the problem of the homogeneity of the CMB can be solved if the Universe expands by a factor of at least  $10^{28}$  during this time.

It is at these very early times that the perturbations that later become all of the structures in the Universe are laid down. In the inflationary model, these perturbations are quantum fluctuations which are blown up to astronomical scales during the period of accelerated expansion. These perturbations not only include the scalar perturbations that lead to structure formation, but also tensor perturbations, propagating perturbations in the metric known as gravitational waves. The properties of these initial perturbations depend on the physics at work during the inflationary era,

and therefore can be used as probes of these early times.

After the initial perturbations are laid down during and immediately after inflation, their large-scale evolution is governed completely by gravitation, and is determined by the constituents of the Universe: radiation, matter (dark and baryonic), and the dark energy, where dark indicates that these components do not interact via electromagnetism, and not coincidentally also indicates that we do not know what they are. After inflation, the energy density of the Universe, and thus its expansion, was dominated by relativistic components. The Universe expanded and cooled, moving eventually to the photon-baryon fluid phase, during which tightly coupled photons, nucleons, and electrons filled the Universe. This fluid exhibited acoustic oscillations due to the opposition between photon pressure and the gravitational pull of the dark matter.

Note that the background of gravitational radiation that was produced during inflation decoupled from matter only  $10^{-22}$  seconds after the end of inflation, while neutrinos decoupled at 2 seconds. After decoupling, these fields streamed freely, carrying their information about the early Universe to us today unimpeded. Two possible methods to detect the gravitational wave background are presented in Chapters 4 and 5. Meanwhile, at a redshift of around  $10^3$ , three very important processes occur. The energy density of radiation drops below that of matter, which decays less quickly with the expansion of the Universe. Then, the electrons and nucleons combine to form neutral atoms, and subsequently the photons, no longer coupled to matter, stream freely (modulo many other small effects), bringing us the CMB. In Chapter 5 we address one of these small effects, in which the gravitational wave background lenses the CMB.

After decoupling, gravitation is no longer impeded by the pressure of photons, and the perturbations in the baryon field stop oscillating and collapse. Also, matter perturbations that enter the horizon after matter-radiation equality collapse much faster than those that enter before, introducing a characteristic scale into the matter power spectrum. The process of collapse of the dark matter, while complicated, is governed entirely by gravitational forces, and many approximations (e.g., spherical top-hat collapse and the Press-Schechter model, see Section 2.1.1) have been developed to qualitatively describe this epoch and the statistics of the resulting gravitationally collapsed structures (dark matter halos). However, there are still a number of complexities, due for example to the presence of dark matter substructure and interaction with external structures, that must be modeled computationally. In Chapter 3 we present one such consideration, the evolution of the angular momentum of dark matter particles in halos.

Of course, we, our galaxy, and indeed all of the matter that we observe in the Universe is ordinary, baryonic matter, whose interactions are far more complicated than merely gravitation. Thus, explaining the structures that we see, such as galaxies and clusters of galaxies, based on the theory of dark matter halos that has been developed, must necessarily include the many complications of baryonic matter and its interactions (this refers to all of astrophysics, for example). This is an

extremely important field, since it allows a great deal of cosmological information to be gleaned from the properties of the galaxies and large-scale structures of the (relatively) nearby Universe. Progress in this field hinges upon simulations, due to the interplay between the many processes that determine structure formation and evolution. One very promising method is semi-analytic modeling of structure formation, in which baryonic physics is overlaid on a model of the dark matter structure of the Universe. This method has a great deal of power, since it can quickly determine the effect of different models of baryonic physics upon the observable properties of the galaxy field. In Chapter 2 we present such a study, looking into the effects of accretion shocks on clusters of galaxies.

As discussed above, this thesis contains four chapters, each one discussing a different aspect of theoretical cosmology. In the first two chapters, we discuss different aspects of the formation of gravitationally bound structures, and in the second two chapters we analyze two methods of detecting primordial gravitational waves. All of these topics are connected through the central role of gravitation, which of course is the main driving force of most of cosmology. The collapse of matter perturbations to form halos and galaxies is driven by gravitation, while gravitational radiation is clearly highly sensitive to the form of gravitation. The work contained in this thesis is primarily reproduced from published work, with permission from the publishers. New material is indicated at the beginning of each chapter.

## Chapter 2

# The Role of Ram Pressure Stripping in the Quenching of Cluster Star Formation<sup>1</sup>

### 2.1 Introduction

It is widely known that the star formation rate (SFR) of galaxies depends on the density of their environment, and specifically that galaxies within galaxy clusters tend to be forming stars at a lower rate than comparable field galaxies. Since star formation is fueled by gas, this leads to the conclusion that galaxies in clusters tend to have less gas from which to form stars than their counterparts in less-dense environments (3).

Several mechanisms have been proposed to explain the observed trend of star formation with environment. It is well known that spiral galaxies tend to be bluer, and therefore forming stars at a higher rate, than elliptical and S0 galaxies (the so-called morphology-density relation (4)) and that spiral galaxies tend to be rarer in the centers of galaxy clusters than early type galaxies (5; 6). It is clear, therefore, that morphological transformation, driven by interactions such as mergers or multiple weak gravitational encounters with other satellites (galaxy harassment, e.g., 7), likely plays a role in driving the observed quenching of star formation in clusters. However, this cannot completely explain the trend as it has been observed that even among galaxies with the same bulge-to-disk ratio, field galaxies are forming stars at higher rates than cluster galaxies (8).

Gas may be removed from galaxies passing through the dense intergalactic medium of a cluster by means of ram pressure stripping (RPS). As originally envisaged, in this mechanism the interstellar medium (ISM) of a satellite galaxy is lost due to interaction with the dense gas of the host halo through which it is moving (9; 10; 11; 12). Alternatively, a satellite moving through less-dense

---

<sup>1</sup>The work presented in this chapter was adapted from the paper “The Role of Ram Pressure Stripping in the Quenching of Cluster Star Formation,” L. G. Book and A. J. Benson, *Astrophys J* **716**, 1 (2010). Reproduced with permission, copyright (2010) by the American Astronomical Society. Sections 2.1.1 and 2.1.2, covering analytical models of structure formation and cosmological applications of galaxy clusters, respectively, are new material.

material may lose just a portion of its diffuse atmosphere of hot halo gas (colloquially known as starvation or strangulation, e.g., 13; 14). Finally, it is inevitable in a hierarchical Universe that many cluster galaxies were previously members of lower mass groups of galaxies, and environmental effects such as strangulation in these lower density environments may have begun the quenching of the star formation of these galaxies before they were accreted onto the cluster (15; 16; 17; 18).

It has traditionally been assumed that these environmental effects only occurred once satellites had fallen through the virial radius of a host halo. However, recent observational studies of satellite galaxy SFR versus cluster-centric radius have found that the SFR remains depressed relative to the field out to two to three times the virial radius of clusters (1; 2; 19). Three explanations have been suggested for the radial extent of environmental effects: that group-scale effects in the locality of the cluster are having a large impact on these cluster galaxies (15), that many satellite galaxies follow highly elliptical orbits which take them close to the cluster center, where they experience strong RPS, and then back out to large radius (1), and that the radius at which cluster environmental effects begins to become important is further out than the virial radius.

In fact, it is already known that using the virial radius as the location at which environmental processes related to the cluster begin is only an approximation. In current structure formation scenarios, dark matter halos are formed through gravitational collapse when an overdense patch of dark matter stops expanding with the Universe and collapses. The surrounding gas then falls into this potential well. If the scale of the halo is large enough so that the gravitational dynamical time of the halo is much less than the cooling time of the gas, then the gas will form an accretion shock, where the kinetic energy of the infalling gas is converted into thermal energy, heating the gas to the virial temperature of the halo (20). These are the so-called “virial shocks,” as they are expected to form near the virial radius of the halo.

Accretion shocks were predicted in a cosmological scenario by (21), who found self-similar, spherically symmetric analytic solutions of collisional gas falling into a density perturbation in an Einstein-de Sitter universe. More recently, models of large-scale, spherical accretion shocks have been developed by (22) and (23), and three-dimensional cosmological N-body and hydrodynamical simulations have confirmed the existence of such shocks (24; 25). Further, (25) has predicted that the extent of such shocks should be observable in the Sunyaev-Zel’dovich effect of clusters with next-generation radio telescopes such as ALMA. While virial shocks may be unstable, and so not survive, around low mass halos, they are expected to be an inevitable result of structure formation for halos above a few times  $10^{11} M_{\odot}$  (26; 27), including groups and clusters of galaxies. When applied in the GALFORM semi-analytic structure formation code, the criterion for shock stability of (26) gives the similar result that stable shocks form only in halos more massive than around  $10^{12} M_{\odot}$  (28).

To determine the cosmological importance of accretion shocks, we apply this well-developed theory to galaxy formation, and, in particular, to the environmental effects on galaxies in clusters.

Semi-analytic models of structure formation are a very powerful tool to investigate the impact of this additional physics on structure formation, as their relatively light computational requirements allow us to probe the effects of accretion shocks on large samples of galaxy clusters, therefore complementing the hydrodynamical simulations mentioned above. Such semi-analytic models have proven to be very successful probes of galaxy formation; they have well reproduced the luminosity function of galaxies locally (29; 30; 31), and at high redshift (32). The addition of astrophysical effects such as supernova feedback, a photoionizing background, and environmental effects have substantially improved the fit at low luminosities (33; 34; 35), while the recent addition of feedback and heating effects to counter large-scale cooling flows, such as heat conduction and feedback from active-galactic nuclei, have produced faithful matches to the high-luminosity tail of the luminosity function (36; 37; 38). The colors of satellite galaxies have also recently been matched by the addition of a more-detailed model of RPS (39; see also 40; 41).

In this work, we consider the physical and observational consequences of the onset of environmental effects at the radius of the accretion shock rather than around the virial radius as has been previously assumed. To study this problem we apply RPS including accretion shocks to the clusters in the Millennium Simulation (42) using the GALFORM semi-analytic model of structure formation. In Section 2.1.1 we review spherical top hat collapse and the Press-Schechter model, and in Section 2.1.2 we review the ways in which galaxy clusters are important for cosmology. In Section 2.2.1 we describe the GALFORM model that we use as the basis of our calculations, while in Section 2.2.2 we briefly describe our calculation of the location of the accretion shock and its implementation in the semi-analytic model, giving a more-detailed description of our calculation in the Appendix. In Section 2.3 we present the cluster galaxy properties produced by our model, and compare them with other models of RPS. In Section 2.4 we compare our model with recent observational results of the SFRs of cluster galaxies, and in Section 2.5 we discuss the conclusions that we draw from these results for the observability of environmental effects in clusters.

### 2.1.1 Analytical Prescriptions for Structure Formation

Inflation predicts that the initial perturbations in the matter field were Gaussian, making the statistics of their distribution very simple. Then they underwent gravitational collapse, which removed them from the regime of linear perturbation theory and Gaussianity. However, this picture is still relatively simple, since we have so far ignored baryonic matter and any non-gravitational forces. Thus, many attempts have been made to explain the distribution of collapsed dark matter structures analytically.

The simplest way to understand the nonlinear gravitational collapse of a matter overdensity is called “spherical top hat collapse.” As the name suggests, in this model we begin by assuming that the universe is completely homogeneous, other than a slightly overdense, spherically symmetric

region. Birkhoff's theorem states that, in such a region, nothing external to the sphere can have an impact on the evolution of internal matter. Thus, we can consider the spherical region in isolation, resulting in equations for its evolution much like those for the evolution of the Universe as a whole.

Using Newtonian gravity, which is a good approximation for the relatively small densities and distances considered here, we can see that the acceleration of a particle on the edge of the overdensity, relative to the center, is

$$\frac{d^2r}{dt^2} = -\frac{GM}{r^2}, \quad (2.1)$$

where  $M$  is the total mass of the overdensity, and  $r$  is its radius. This differential equation has a parametric solution

$$\begin{aligned} r(\theta) &= r_{\max}(1 - \cos \theta) \\ t(\theta) &= \sqrt{\frac{r_{\max}^3}{GM}} (\theta - \sin \theta). \end{aligned} \quad (2.2)$$

Here,  $\theta$  ranges from 0 to  $2\pi$ . For convenience, let us define  $A \equiv \sqrt{\frac{r_{\max}^3}{GM}}$ . We can expand in a Taylor series

$$\begin{aligned} r(\theta) &\sim r_{\max} \left( \frac{1}{2!}\theta^2 - \frac{1}{4!}\theta^4 + \frac{1}{6!}\theta^6 \right) \\ t(\theta) &\sim A \left( \frac{1}{3!}\theta^3 - \frac{1}{5!}\theta^5 + \frac{1}{7!}\theta^7 \right). \end{aligned} \quad (2.3)$$

We wish to write the radius as a function of time. The first-order behavior is  $r(t) \sim r_{\max}/2(6t/A)^{2/3}$ . The next order gives

$$r(t) \sim \frac{r_{\max}}{2} \left( \frac{6t}{A} \right)^{2/3} \left( 1 - \frac{1}{20} \left( \frac{6t}{A} \right)^{2/3} \right). \quad (2.4)$$

This gives a first-order fractional change in the radius of

$$\frac{\delta r}{r} = -\frac{1}{20} \left( \frac{6t}{A} \right)^{2/3}, \quad (2.5)$$

leading to a fractional change in the density, i.e., an overdensity

$$\delta = -3\frac{\delta r}{r} = \frac{3}{20} \left( \frac{6t}{A} \right)^{2/3}. \quad (2.6)$$

The goal of this derivation is to determine the value of the overdensity, in linear theory, at which the spherical halo collapses. This occurs in the linear model derived above when the radius shrinks back to zero at  $\theta = 2\pi$ . The overdensity at this point, extrapolated (incorrectly) from the linear theory is  $\delta_c \approx 1.69$ . Using this value, we can approximate the number of collapsed halos by simply counting

the number of places in the linear density field with overdensities in excess of  $\delta_c$ , a calculation that is even tractable analytically, with simplifications.

Now that we have understood the collapse of an (admittedly highly simplified) overdensity, we would like to apply it to find the number density of halos as a function of their mass. To do so, we will use the simple Press-Schechter (PS) argument. The basis of the idea is as follows: the density fluctuations of the Universe can be written as  $\delta(\vec{x}) = \frac{\rho(\vec{x}) - \bar{\rho}}{\bar{\rho}}$ , where  $\rho(\vec{x})$  is the matter density at position  $\vec{x}$  and  $\bar{\rho}$  is the average density of the universe at the time being considered. As mentioned above, inflation predicts that this field will initially be Gaussian distributed (putting aside considerations that  $\delta$  cannot become less than  $-1$ , while it can and definitely does go beyond 1. The PS argument states that, if we smooth the density field over a mass scale  $M$  (corresponding to radii  $R = \left(\frac{3M}{4\pi\bar{\rho}}\right)^{1/3}$ ), every point at which  $\delta$  exceeds the linear collapse threshold  $\delta_c = 1.69$  (calculated above) resides in a halo with mass  $M$  (or greater). The assumption of Gaussianity implies that the probability for  $\delta$  to obtain any given value at a specific point in space is

$$P(\delta) = \frac{1}{\sigma(M)\sqrt{2\pi}} e^{-\delta^2/2\sigma(M)^2}, \quad (2.7)$$

where  $\sigma(M)$  is the variance in the density field for regions with typical mass  $M$ . Thus, the probability that the overdensity at a given position is above the critical linear density for collapse is

$$p(\delta > \delta_c) = \frac{1}{\sigma(M)\sqrt{2\pi}} \int_{\delta_c}^{\infty} d\delta e^{-\delta^2/2\sigma(M)^2}. \quad (2.8)$$

Then, the PS argument says that the fraction of the mass in the Universe contained within bound halos with mass greater than  $M$  is given by

$$F(M) = 2p(\delta > \delta_c) = \frac{2}{\sigma(M)\sqrt{2\pi}} \int_{\delta_c}^{\infty} d\delta e^{-\delta^2/2\sigma(M)^2}. \quad (2.9)$$

Here, the factor of two was added ad hoc to fix the problem that only the matter in overdense regions is assumed to end up in bound structures in this model, and can be derived exactly in the excursion set approach. From here, we can obtain the number density of halos with mass in the range  $M + dM$ ,  $n(M) dM$ , by multiplying the fraction of the universe contained in such halos  $dF(M)/dM$  by the number density that objects of this mass would have if they contained all of the mass in the Universe,  $\bar{\rho}/M$

$$n(M)dM = -\frac{dF(M)}{dM} dM \frac{\bar{\rho}}{M}, \quad (2.10)$$

where the minus sign accounts for the fact that  $\sigma(M)$  is a monotonically decreasing function of  $M$ , and therefore  $F(M)$  is also monotonically decreasing with  $M$ . Plugging in the fraction of halos with

mass greater than  $M$ , we find the result

$$n(M) = -\sqrt{\frac{2}{\pi}} \frac{\bar{\rho}}{M^2} \frac{\delta_c}{\sigma(M)} \frac{d \ln \sigma(M)}{d \ln M} e^{-\delta_c^2/2\sigma(M)^2}. \quad (2.11)$$

Thus, in this simple theory we predict that the number density of halos will fall off exponentially fast at masses higher than a critical mass  $M_*$ , for which  $\sigma(M_*) = \delta_c$ . This mass today is around  $10^{13} h^{-1} M_\odot$ , corresponding to poor galaxy clusters. Therefore, clusters with masses greater than this characteristic scale probe the exponential tail of the distribution.

This simple argument has been applied to obtain halo merger rates (43; 44; 45), and expanded to fit N-body simulations (46). Despite the amazingly simple derivation of these results, they are the backbone of current structure formation research.

### 2.1.2 Cosmological Implications of Galaxy Cluster Properties

In the simple PS model presented above, we saw that the number of halos above the critical mass  $M_*$  falls off exponentially fast. This is the key to one of the main uses of galaxy clusters in cosmology, the measurement of the number count of clusters as a function of mass or of redshift. Since these objects are generally above the critical mass, they are very sensitive to the cosmological parameters that influence the number density through  $\bar{\rho}$ ,  $\sigma(M)$  and  $\rho_c$ , including the matter density  $\Omega_M$ , the baryon density  $\Omega_b$ , the dark energy density  $\Omega_\Lambda$ , the scalar spectral index  $n_s$  and the power spectrum normalization  $\Delta_R(k_0)$ . Note that other effects, such as non-Gaussianity (47; 48; 49; 50) or departures from general relativity (51; 52), also affect the number densities of bound structures, and thus can also be constrained using cluster number counts. There are several current and future surveys aiming to measure cluster number counts and therefore directly constrain these parameters (53; 54; 55; 56; 57; 58; 59; 60; 61).

The main difficulty with the above approach is the determination of cluster masses—the exponential sensitivity of the number density of clusters to their mass means that a small mass error can translate into an enormous error in cosmological parameter estimation. There are several approaches currently used to estimate cluster mass, including x-ray luminosity, cluster kinematics, the Sunyaev-Zel'dovich effect and weak lensing.

Let us first consider the x-ray luminosity of clusters as a probe of their mass. In hydrostatic equilibrium, the pressure  $P$  of gas in a cluster is directly related to the cluster's gravitational potential  $\Phi$

$$\nabla P = \rho_g \nabla \Phi, \quad (2.12)$$

where  $\rho_g$  is the gas density. The gas temperature can then be linearly related to the pressure using the ideal gas law. We can therefore see that a very deep potential well, such as that caused by a

massive cluster, will host very hot gas, and that, moreover, the cluster mass can be directly obtained from the gas temperature and density profiles. This method of measuring the mass of clusters from x-ray luminosity has been well studied both in theory (62; 63) and observation (64; 65; 66; 67).

Similarly, the velocities of satellite galaxies in clusters are directly related to the gravitational potential of the cluster. For example, for a spherically symmetric cluster whose satellites have an isotropic velocity distribution, the gravitational potential  $\Phi(r)$  and radial velocity dispersion  $\sigma_r$  (the rms radial velocity) can be related via

$$\frac{d}{dr} (\rho(r)\sigma_r^2) = \rho(r)\frac{d\Phi}{dr}, \quad (2.13)$$

where  $\rho(r)$  is the (total) matter density at radius  $r$ . In this way, we can relate the observable velocity dispersion to the gravitational potential, providing another measurement of cluster density profiles and thus masses (68; 69; 70; 71).

A massive cluster, containing hot gas as described above, will also inverse Compton scatter the relatively cold CMB photons, shifting lower-frequency photons to higher frequency and resulting in cold or warm spots in the CMB (depending on the wavelength of the measurement) which are correlated with cluster positions. This process is called the thermal Sunyaev-Zel'dovich effect (72; 73), and the shape of the effect with frequency is very well understood. The amplitude of the effect is directly proportional to an integral of the pressure of the gas along the line of sight, which is then proportional to the temperature times the electron density integrated along the line of sight (from the ideal gas law). Thus, if the temperature of the cluster gas is understood, we can extract the electron density, and from there infer its mass by applying models of the composition and ionization state of the cluster gas. This effect has been observed in the CMB by a number of missions, including the Planck satellite, the South Pole Telescope, and the Atacama Cosmology Telescope (55; 58; 59; 60; 61; 74; 75).

All of the methods discussed above to measure cluster masses depend on assumptions about the dynamical state of the cluster. However, we can get rid of these assumptions (or check them) by using the gravitational effect of a cluster on passing photons to directly measure its mass. This is the field of lensing, also discussed in Section 5.1.3. In this method, the shapes of background galaxies can be measured and related directly to the gravitational potential of the cluster (and any other mass along the line of sight). Specifically, let us consider a photon emitted from a galaxy a comoving distance  $\chi$  away from us, with transverse coordinates  $(\chi\vec{\theta}_S)$ , and look at the effect of lensing upon the observed angular position of the origin of this photon  $\vec{\theta}$ . We define  $A_{ij} \equiv \frac{\partial\theta_S^i}{\partial\theta^j}$ , the change in shape of background galaxies, and relate it to the foreground gravitational potential evaluated along

the trajectory of the photon (76)

$$A_{ij} - \delta_{ij} = 2 \int_0^x d\chi' \Phi_{,ij}(\chi') \chi' \left(1 - \frac{\chi'}{\chi}\right). \quad (2.14)$$

This quantity can then be related to the induced ellipticities of background galaxies (in the case of weak lensing), providing a complementary probe of cluster masses to the previous methods, since it relies on completely different assumptions and thus is sensitive to different systematics. These distortions in the shapes of background galaxies are clearly measurable with current telescopes (77; 78; 79; 80). These methods agree fairly well (75), and together provide us with fairly robust cluster masses.

Beyond cluster number counts, there are many other ways of using galaxy clusters in cosmology. For example, since clusters are extremely large systems, whose mass has been collected over a huge area of around 1000 Mpc<sup>3</sup>, it is theorized that the ratio of baryonic to dark matter in clusters should be representative of that of Universe as a whole (81; 82; 83; 84). Also, clusters are important probes of the formation of structure, and their study can provide information about the processes relevant in galaxy formation. This then allows us to better understand galaxies and their properties for use in cosmology.

## 2.2 Simulations

### 2.2.1 The GALFORM Semi-Analytic Model

We use as the basis of our work the GALFORM semi-analytic model of galaxy formation as described by (37), to which the reader is referred for a detailed description of the model formulation. As described in (37, see also 85), we apply this model to the detailed merger histories of the entire volume of the Millennium Simulation (42). Our model uses the same cosmological parameters as that simulation:  $\Omega_m = 0.25$ ,  $\Omega_b = 0.045$ ,  $\Lambda = 0.75$ , and  $h = 0.73$  at  $z = 0$ .

As we are working with N-body merger trees, there are situations in which halos decrease in mass with time (85), as a result of unbound particles incorrectly being tagged as halo members, for example. The original implementation of our semi-analytic model was not well equipped to deal with mass loss in halos, so these N-body merger trees were artificially forced to conserve mass. However, in this work we utilize the merger trees without requiring mass conservation, as described by (86), as this is a fairer representation of the true behavior of the N-body simulation and is important for this work in which we utilize halo mass growth rates to compute virial shock radii. We find that relaxing the requirement of mass conservation results in an increase in the mean stellar mass content of galaxies of around 0.3 dex. This change is larger than expected from the work of (85), due to the sensitive nature of the AGN feedback included in our current model (but which was not present in

that of 85).

We therefore found it necessary to adjust a single parameter of our model relative to that of (37) to retain a good fit to the observed local  $b_J$ - and K-band galaxy luminosity functions. Specifically we adjust the parameter  $\alpha_{\text{cool}}$ , described by (37), which determines the halo mass scale above which AGN feedback becomes effective. We find that increasing  $\alpha_{\text{cool}}$  from 0.58 to 0.9 reduces the lower mass limit for AGN heating to become effective from a few times  $10^{11}M_{\odot}$  in the Bower model to a few times  $10^{10}M_{\odot}$  in this work, and therefore reduces the mean stellar mass of galaxies. This change tends to cancel out the change of stellar masses caused by not conserving mass in merger trees (as was done in (37), and brings our model back into agreement with the local luminosity functions.

The prescription for the treatment of reheated gas in satellite galaxies that we adopt is the same as that of (39), to which the reader is directed for complete description. In brief, gas in a satellite galaxy that is reheated by supernovae or AGN feedback is transferred to the hot halo of the satellite, from which it is transferred to the host halo as the satellite is ram pressure stripped.

In the (37) model, a galaxy that is identified as part of a friends-of-friends group of a more-massive halo is considered its satellite. These tend to roughly correspond to satellites within the virial radius of the halo, although the prescription is by no means spherically symmetric. The satellite has all of its hot gas instantaneously stripped away, leaving the cold gas in the ISM of the galaxy but removing the source of gas to be accreted onto the galaxy and form stars. Although this model has been quite successful in reproducing the luminosity function and star formation history of galaxies and their evolution, it fails to reproduce the colors of satellite galaxies, tending to predict them to be redder than is observed.

The challenge to reproduce the correct colors of satellite galaxies was taken up by (39), who implemented a more-nuanced approach to the RPS of satellite galaxies based on the hydrodynamic simulations of (14). Keeping the definition of satellite galaxies as those belonging to the friends-of-friends group of a more-massive halo, in their model each satellite is assigned an orbit assuming the velocity distributions determined by (87), and from this they calculate the maximum ram pressure exerted on the satellite by the host halo and galaxy, which occurs at the pericenter of its orbit. The radius at which this maximum ram pressure is equal to the gravitational restoring force per unit area of the satellite is termed the stripping radius, and all of the hot halo gas beyond this radius is stripped at the moment that the satellite crosses the virial radius of the host halo. This calculation was repeated (possibly resulting in more gas being removed from the satellite) every time its host halo doubled in mass since the previous ram pressure calculation. This model provides a less-extreme implementation of RPS, and manages to match the colors of satellite galaxies by allowing them to accrete hot gas and remain blue for a longer time after being accreted.

However, the ram pressure model of (39) continues to use the virial radius as the location around which a satellite begins to feel ram pressure from its host galaxy. In fact, ram pressure forces begin

to be felt by the infalling galaxy when it passes through the accretion shock, at which the cluster gas temperature, density and pressure discontinuously increase. The accretion shock can be up to twice as far from the host galaxy as the virial radius, potentially significantly altering the effect of a massive host halo on nearby galaxies. Additionally, the model of (39) uses randomly assigned orbital parameters for satellites to compute their orbit and, therefore, the ram pressure force that they experience.

## 2.2.2 Implementation of Accretion Shocks

In this work we calculate the location of the accretion shock of halos and use this as the radius at which the RPS of satellite galaxies begins, thereby more completely modeling the environmental effects of a host halo on its satellites. We calculate the radius of the accretion shock with a model based on the calculations of (88, hereafter V03), with a few assumptions relaxed to obtain a more-accurate accretion radius in a wider range of situations. We use the method of (39) to implement the RPS of a galaxy once it comes within this radius of the cluster, but use the actual orbit of the satellite (taken directly from the N-body simulation) to compute the ram pressure force experienced at each timestep of our calculation.

The method of accretion shock calculation of V03 is an approximate solution with many simplifying assumptions. For example, it assumes smooth, spherical accretion, an assumption that is known to be quite incorrect in the context of hierarchical galaxy formation. For this reason, there is uncertainty as to the accuracy of the predictions of this model for the hydrostatic structure of cluster gas. For example, cosmological hydrodynamic simulations have shown that the hot gas halo of a cluster can extend well beyond the location of the accretion shock (89), pointing to far more complicated physics than is included in the simple V03 model. This discrepancy is likely to be small, as the relative velocity between infalling satellite galaxies and accreting gas, which determines the ram pressure along with the density, will tend to be low outside the shock radius since they are both falling into the host halo and feel little pressure, while within the shock radius the accreted gas becomes nearly stationary creating a large relative velocity with the satellites. However, as it has been shown in numerical simulations of the hierarchical dark matter and gas evolution of galaxy clusters that the profiles of the outer regions of clusters tend to agree with smooth accretion models (90), we use this simplified model in this work.

The V03 calculation of the accretion shock radius also assumes that the accretion shock is always perfectly strong, so that the Mach number of the shock approaches infinity. We relax the strong-shock assumption to allow the accretion shock to have any strength, which we find to be justified as even some massive clusters have accretion shocks with Mach numbers  $\mathcal{M} \sim 10$ . We also correct for the neglect of an integration constant in the hydrostatic profile derived for the clusters, which we found to be non-negligible.

Briefly, the calculation of the accretion shock proceeds as follows: using the shock jump conditions and the assumption that the total accreted gas mass must be contained within the accretion shock, we obtain a simple hydrostatic model of the cluster gas. Using this model we derive an equation for the accretion radius in terms of the halo mass, mass accretion rate and halo concentration, as well as the Mach number of the accretion shock. By also assuming that the cluster gas has adiabatically contracted from a temperature of  $T_{\text{IGM}} = 3000K$  in the intergalactic medium, we simultaneously solve for the Mach number of the accretion shock and its position.

The calculation on which our model is based is described in Appendix A of V03, and a more-detailed description of our generalized version of the calculation can be found in the appendix.

## 2.3 Cluster Galaxy Properties with Different Models of Ram Pressure Stripping

We first compare the properties of galaxies in our model, in which the stripping of gas begins at the radius of the accretion shock, with the similar model of (39) in which these effects begins at the virial radius, and with the model of (37) in which all of a satellite galaxy’s hot gas, not just the maximum amount as determined by the parameters of its orbit within the host halo, is stripped away at the virial radius. We will call these models the Shocks, Font, and Bower models, respectively.

In Figure 2.1 we show the resulting stellar mass, cold gas mass and gas fraction profiles of cluster galaxies, excluding the central galaxy, averaged within cluster-centric radial bins. The values of the central galaxies are plotted, for comparison, as points at zero radius. Here we define galaxy clusters as halos whose mass is greater than  $10^{14}h^{-1}M_{\odot}$ , and we include all galaxies in each radial bin in the averages. It should be noted that the choice of which galaxies to include makes a large difference in the properties that are observed, as can be seen in the significantly different trends obtained in Figures 2.3 and 2.4 where we have selected a different set of galaxies to compare to observations.

We compare our Shocks model (red solid lines), with the Font (green solid lines) and Bower (blue solid lines) models. It should be noted that there are other parameter differences between these models: the Shocks model has had a parameter associated with AGN feedback adjusted from its Bower model value to match the local galaxy  $b_J$ - and  $K$ -band luminosity functions, while the Font model has a different value of the metal yield to match the zero point colors of the red and blue sequences. Thus, the difference between the Bower model and the Shocks and Font models is due both to their differing treatments of RPS, and to their different physical parameters. To assess the relative contributions of each of these components, we also consider the Shocks, Font, and Bower models without RPS, thus isolating the effects of the parameter changes (red, green and blue dotted lines). We also show the position of the average accretion shock radius of these clusters as a black dashed vertical line, and the properties of the central galaxies as crosses. When these values are very

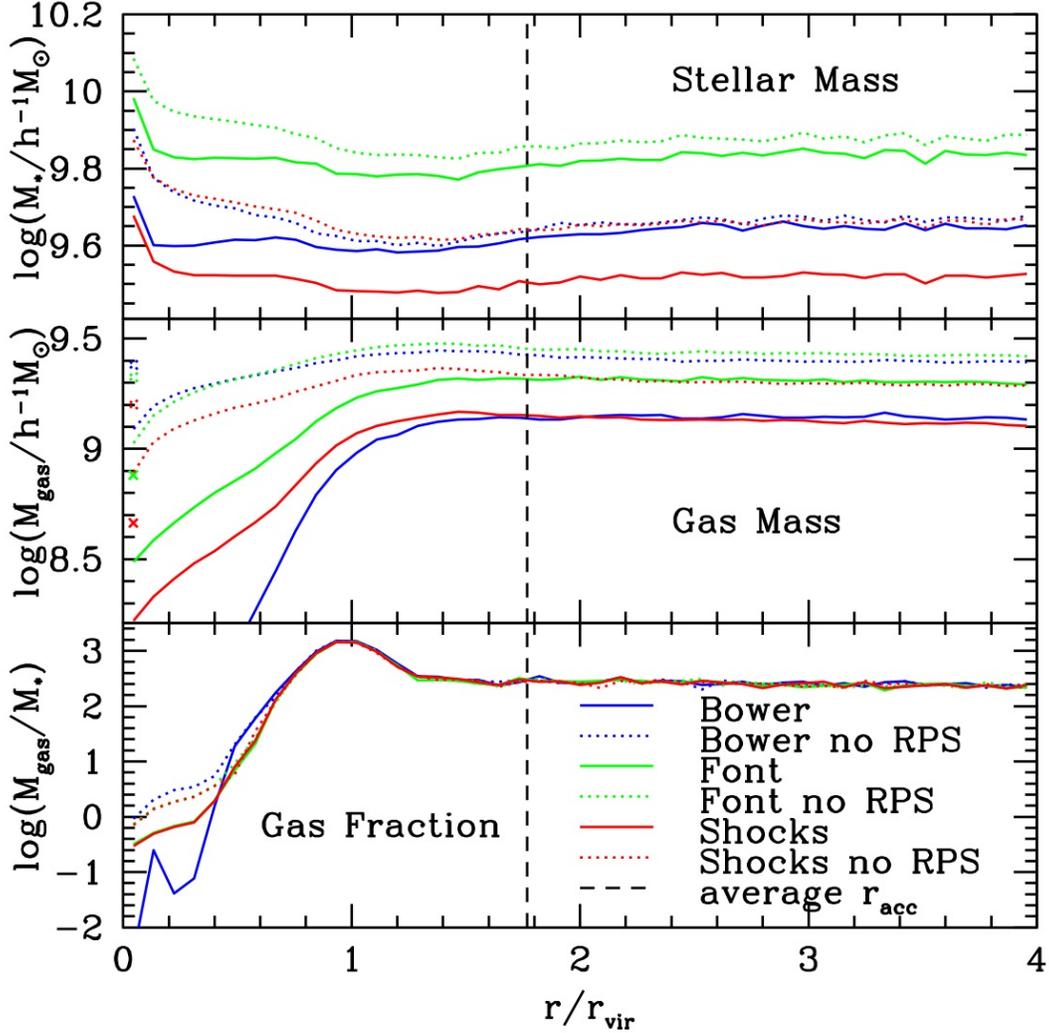


Figure 2.1: Stellar mass ( $M_*$ , top), cold gas mass ( $M_{gas}$ , middle), and the gas fraction,  $M_{gas}/M_*$  (bottom), averaged for satellite galaxies in radial bins, as a function of the cluster-centric radius divided by the virial radius of the cluster. The black dashed line shows the average position of the accretion shock in these clusters, while the blue, green, and red solid lines show the results for the (37), (39), and Shocks models. The blue, green, and red dotted lines show the Bower, Font, and Shocks models without ram pressure stripping, respectively. The crosses at zero radius indicate the average properties of the central galaxies.

different from those of the satellite galaxies, we leave them out of the plot as the satellite galaxy trends are more interesting for the purposes of this work.

In the top panel of Figure 2.1, we see that in all of the models there is a peak in galactic stellar mass in the centers of clusters; as the central galaxy has been excluded, this shows an increase in stellar mass of the innermost satellites in all models. This is reasonable, as we expect more-massive satellite galaxies to sink deeper into the potential well of the cluster. Further, other than a slight rise in stellar mass towards the center of clusters in the Font and Shocks models, the average stellar

mass of galaxies remains more-or-less constant with radius, with the Shocks model having the lowest average level of stellar mass, the Bower model slightly more, and the Font model having the most stellar mass of all. The flatness of these curves indicates that most of the stars of satellite galaxies were formed before they merged with the cluster. Note that the central galaxies in all of the models have more than an order of magnitude more stellar mass than the satellite galaxies, as would be expected since they are at the center of the potential well of the cluster and are not subject to RPS.

The distribution of cold gas mass, in the middle panel of Figure 2.1, shows a different trend. The Bower model, with its more-extreme removal of gas in clusters, shows the least cold gas in satellite galaxies, with a very distinct drop in gas mass at around 1.5 times the virial radius. This drop in cold gas mass is quite steep, with its slope determined by the timescale of satellite orbits as compared to the timescale on which cold gas is made into stars in the satellite galaxies. Interestingly, the Shocks model shows a very similar trend with slightly more gas mass inside the radius of the accretion shock, as would be expected. That the Bower model also shows a drop at around this radius, which happens to be the average accretion shock radius, is interesting, and points to the “preprocessing” of satellite galaxies, in which satellite galaxies experience weaker environmental effects as members of smaller groups of galaxies before merging with the cluster. The Bower model is likely to exhibit stronger group effects than the Shocks and Font models, since its RPS efficiency is always very high, while the RPS of the Shocks and Font models depend on the density of the halo intracluster medium (ICM). Thus, we see that these two very different models of environmental effects give a similar qualitative prediction for the radial dependence of satellite gas mass, indicating that if such a trend in gas mass is observed, we cannot distinguish between a model whose RPS starts at the accretion radius and a much stronger RPS model which starts at the virial radius.

The Font model predicts a higher level of gas mass at all radii, as is reasonable due to its less-harsh RPS implementation, while as expected the models without ram pressure exhibit significantly higher gas mass at all radii than the other models, since it is RPS that is mainly responsible for the sharp drop in gas mass with decreasing cluster-centric radius. It is reasonable that there is a slight drop-off in gas mass towards the center of the cluster even in these models, since in these models the accretion of new gas from the intergalactic medium (IGM) onto the hot gaseous halo of the satellites is suppressed though none of their halo gas is removed. We see that the central galaxies have higher gas mass than satellites, as they are not affected by RPS and also gain gas through mergers. The higher level of gas mass in central galaxies in the models without RPS is likely due to merging satellites, which have not had gas removed by RPS and therefore give more gas mass to the central galaxy.

Finally, in the bottom panel of Figure 2.1 we can see the average ratio of cold gas to stellar mass in satellite galaxies. It is quite striking that, despite the differences in gas mass and stellar mass when averaged separately, the average of their ratio is nearly the same in all of these models, with

the only significant difference occurring between the Bower model and the Font and Shocks models at less than half of the virial radius. The general behavior of this quantity, in all of the models, is a gradual increase in gas fraction moving into the cluster until the virial radius, at which the gas mass declines precipitously. This peak is likely due to the biased sample of galaxies located close to massive clusters. As they are in general more massive, they may well still be cooling gas and thus forming stars at rates higher than further away from the cluster.

Further inwards, the Shocks and Font models reach a higher, roughly flat, central gas fraction than the Bower model, as expected due to their less-extreme removal of gas in cluster members. We also see that the gas fraction reaches only a slightly larger value in the runs without any RPS, indicating that at least the subtle ram pressure of the Shocks and Font models makes little difference in the gas fraction of cluster galaxies, while the implementation of a Bower-type complete removal of gas makes a far larger difference. Note, finally, that the spike in the Bower model gas fraction very close to the center is due to a single galaxy in our sample, which formed in the final timestep and so exhibits an unusually low stellar mass and therefore a very high gas fraction. Its location deep within the cluster is surprising given its very recent formation and it may represent a flaw in the halo detection or tree building algorithms. We leave it in the sample in any case for completeness. Finally, the gas fractions of central galaxies in these models are all well below those of satellites, which is reasonable as cluster central galaxies tend to be ellipticals which have little cold gas although they may have large hot gas halos.

In Figure 2.2, we show observationally measurable quantities related to the implementation of RPS: the B–V color, SFR, and specific SFR of satellite galaxies in clusters. Each of these quantities, as in the previous figure, is calculated for each galaxy and then averaged within cluster-centric radial bins.

In the top panel of Figure 2.2 we can see that in all of the models the satellite galaxies become increasingly redder moving towards the center of the cluster, as expected, with this reddening beginning around the virial radius and having the same general shape in all models. We see that the Shocks model has bluer colors than the other models, and that the Font model has slightly redder colors than the Bower model. The Font model, with the most gas, has the most dust extinction, and it is this effect which makes it redder than the Bower model. As the separation between the Shocks and Font models is preserved when RPS is turned off, we see that these average color differences result from the star formation and AGN feedback parameter adjustments which were required to bring the models into agreement with the local luminosity function. As expected, without RPS we see much bluer colors in all of the models. The central galaxies are all much redder than their satellites, which follows from our knowledge that these galaxies tend to be large elliptical galaxies.

In the middle panel of Figure 2.2, we see the SFR variation with cluster-centric radius. The SFR versus radius for the Shocks and Font models decreases moderately inside the virial radius,

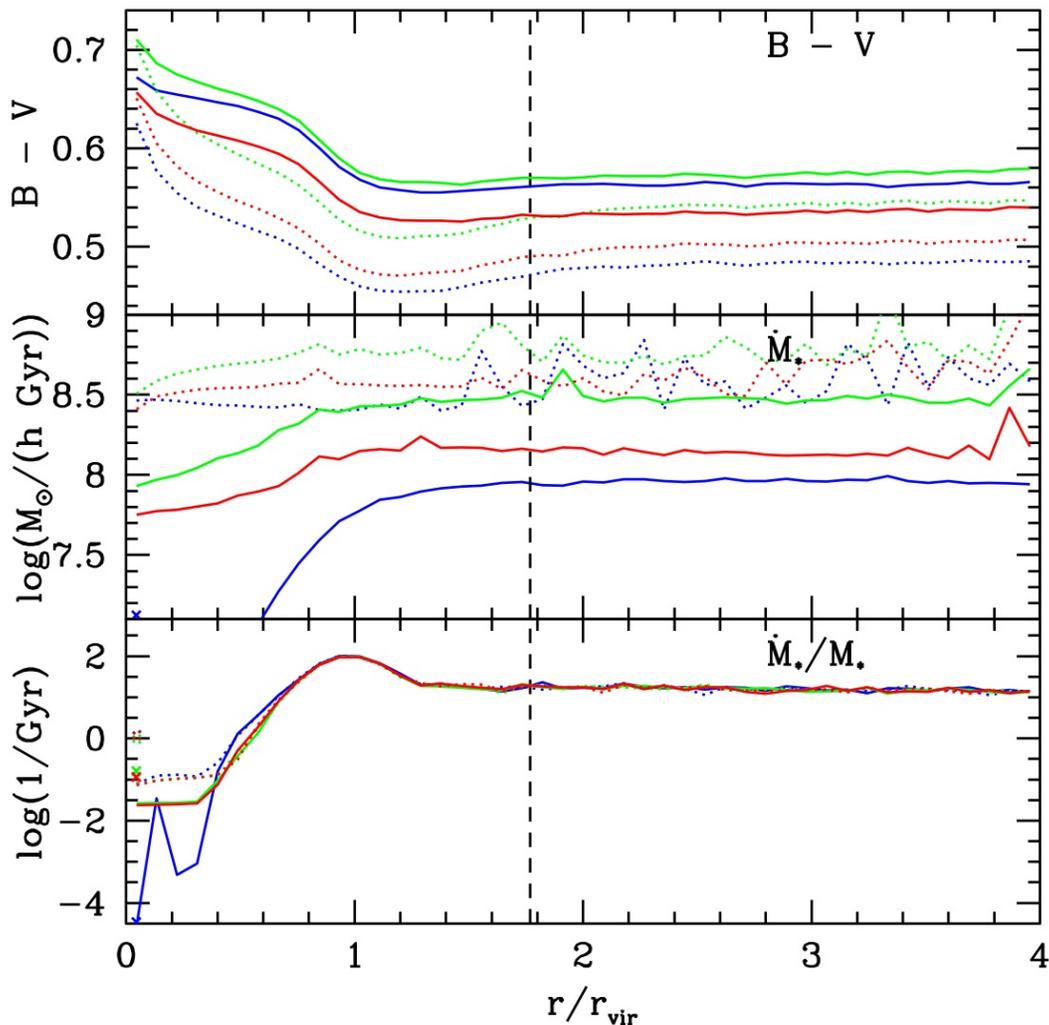


Figure 2.2: B-V colors (top), star formation rate (middle), and specific star formation rate (bottom), averaged for satellite galaxies in radial bins, as a function of the cluster-centric radius divided by the virial radius of the cluster. The lines and models are the same as in Figure 2.1.

in a similar manner to the gas mass, though not quite as steep. We see a distinct decrease in star formation in the Bower model towards the center of the cluster, due to its harsher implementation of RPS and thus lower gas mass. The Font model has more star formation at all radii with respect to the other models, while the Bower model always has the lowest SFR. This is reasonable as it mirrors the order of strength of RPS in these models. Also as expected, the models without RPS show higher rates of star formation than the other models. The central galaxies in these models have much larger SFRs than their satellite galaxies, which is reasonable given their larger gas mass.

In the lower panel of Figure 2.2, we see the average specific SFR, that is the average of the ratio of the SFR and the total stellar mass of satellite galaxies. The trends in this plot are very similar

to those of the gas fraction plotted in the lower panel of Figure 2.1. This is no coincidence; in fact the SFR in GALFORM is determined such that

$$\frac{\dot{M}_*}{M_*} = \frac{\epsilon}{\tau_{\text{disk}}} (1 - R) \frac{M_{\text{gas}}}{M_*}, \quad (2.15)$$

where  $\epsilon$  is the star formation efficiency,  $\tau_{\text{disk}}$  is the disk timescale, and  $R$  is the fraction of mass going into stars that is recycled back into the interstellar medium. As with the gas fraction, a single outlier galaxy in the Bower model is causing the spike at low radius, and can be ignored. We see that the central galaxies in all models have significantly higher specific SFR, as expected. Also, the specific SFR is directly proportional to the gas fraction, and as  $\epsilon$  and  $R$  are constants and  $\tau_{\text{disk}}$  does not depend strongly on environment, we see similar behavior in the specific SFR as a function of cluster-centric radius as we saw in the gas fraction. As with the gas fraction, it is very similar in all models and shows a very slightly rising level moving in towards the virial radius, at which there is a peak and then a sharp decline towards the center of the cluster. As with the gas fraction, the models without RPS reach a higher central value of the specific SFR, the Font and Shocks models are very similar, and the Bower model shows a much lower central specific SFR. The central galaxies in these clusters all show a much lower specific SFR than the satellites; this is directly related to their low gas fraction, which is understood as these galaxies tend to be large ellipticals.

## 2.4 Comparison with Observations

We compare the results of our GALFORM model including accretion shocks and the more-extreme RPS model of (37) to recent observations and simulations. In Section 2.4.1, we compare with the simulations of (1) and their observational dataset selected from the CNOC1 cluster redshift survey (91). In Section 2.4.2 we compare the Shocks model to the cluster spectroscopy of (2). In both papers, cosmological parameters  $\Lambda = 0.7$ ,  $\Omega_0 = 0.3$ , and  $h = 0.7$  are used. Despite the fact that GALFORM uses a different set of cosmological parameters as listed in Section 2.2.1, we conduct our comparison analyses using the same parameters as the observations so as to better reproduce their analysis. In both comparisons, we plot the properties only of satellites in halos with masses greater than  $10^{14}h^{-1}M_\odot$  to more accurately mimic the selection of massive clusters in these two samples.

### 2.4.1 Comparison to CNOC1 Cluster Redshift Survey as Selected by Balogh et al. (2000)

The cluster redshift survey CNOC1 (91) provides spectra for member galaxies of 15 x-ray luminous clusters. We compare our simulations to the sample of twelve of these clusters presented in (1), which were selected to lie within a redshift range  $0.19 < z < 0.45$  and have well-defined cluster

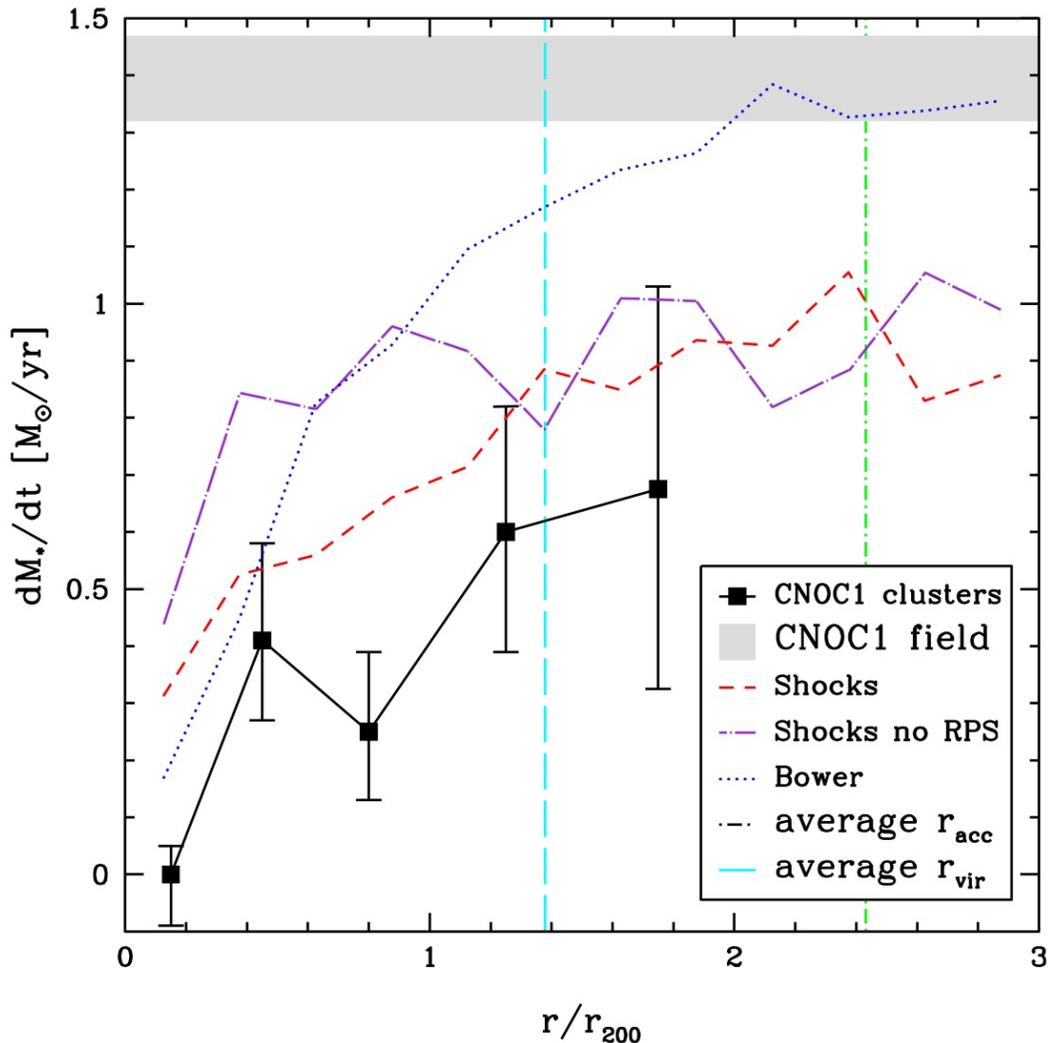


Figure 2.3: Comparison of the cluster redshift survey CNOC1 data as obtained by (1), with the Millennium results using three different implementations of RPS. The CNOC1 data is shown as black squares and solid lines, while the extreme ram pressure stripping model of (37) is shown in blue dotted lines, the more-nuanced accretion shocks model of this chapter is shown in red dashed lines, and this same Shocks model without ram pressure stripping is shown in purple dot-dashed lines. The vertical lines show the average virial radius (cyan long-dashed) and accretion shock radius (green short dot-dashed) for the clusters considered in our semi-analytic model. Note that the x-axis of this plot is the projected cluster-centric radius divided by  $r_{200}$ , as defined in the text.

centers. Cluster members were selected based on velocity and magnitude cuts. The cluster-centric radii of galaxies were measured relative to  $r_{200}$ , the radius at which the interior density is 200 times the critical density. See (1) for a complete description of the selection criteria.

To ensure accurate comparison to the data, we mimic the observational techniques as described in Section 2 of (1) in our analysis of the Millennium/GALFORM models. We analyze clusters at redshift  $z = 0.3$ , and cluster members are selected based on projected radial position, magnitude, and

velocity. Specifically, we select galaxies whose projected cluster-centric radius is less than  $2 r_{200}$ , where  $r_{200}$  is calculated assuming an NFW-type density profile, with a Gunn r-band magnitude greater than  $-18.8 + 5 \log h$  at  $z=0.3$ , and with velocity within  $3 \sigma$  of the average velocity of all previously selected galaxies, including the effect of Hubble expansion. The average virial radius (as defined by GALFORM using the overdensity of a spherical top-hat collapse model for this cosmology) of the clusters we “observe” in this manner is  $1.37 r_{200}$ . As in (1), we remove the central cluster galaxies from the sample.

We calculate the SFR as was done with the observations, by using the equivalent width of the OII line as a direct indicator of SFR. (1) find a prescription to determine SFR from observed line widths and luminosities that matches the relationship between these quantities that they see in their semi-analytic simulations. In their prescription, the SFR is the product of this equivalent width with the rest-frame B-band luminosity, an extinction factor, and a normalization constant chosen by comparison with their simulations. In Figure 2.3 we show the comparison of the CNOC1 data of (1) with the results of taking our semi-analytically determined cluster galaxy luminosities and line widths and plugging them into this prescription. We note that this prescription tends to under-predict the SFRs of our galaxies as compared to the rates directly obtained from GALFORM, and that this effect is strongest for models with less RPS, more galactic gas and dust and thus more extinction. Thus, the SFRs inferred from this method agree fairly well with the directly computed rates for the Bower model, but are much lower than the directly computed rates of the Shocks model. This shows that prescriptions to determine SFR from observed quantities are quite model dependent, and in this case disagree with the relationship we see in GALFORM.

The resulting curves are compared with the CNOC1 data in Figure 2.3. Here we plot three models of ram-pressure stripping: the (37) model, our new Shocks model, and this same Shocks model with the RPS turned off (blue dotted, red dashed and purple long-dot-dashed lines, respectively). We also show, with cyan long-dashed and green short-dot-dashed lines, the average virial and accretion shock radii of the clusters considered. The radius plotted on the x-axis is the projected cluster-centric radius. Note that the SFRs shown in Figure 2.3 are different from those in Figure 2.2 due to their differing definitions of radius, different selection of galaxies, the different units and the linear scale in Figure 2.3 as compared to the logarithmic scale in Figure 2.2. When taking these into account, the SFRs in the two plots are similar.

We can see that all three models predict similar radial trends in SFR. The Bower model, with more-extreme RPS, is quite close to the data in the innermost regions, but has a much steeper rise to larger radii than is observed, thus overproducing the SFR in the outer regions. The Shocks model rises more gradually in the outer regions of the cluster, bringing both of these models into better agreement with the data than the Bower model at large radii. However, we can see that, given the large error bars on the data, neither the Shocks model (with RPS) or the Bower model is

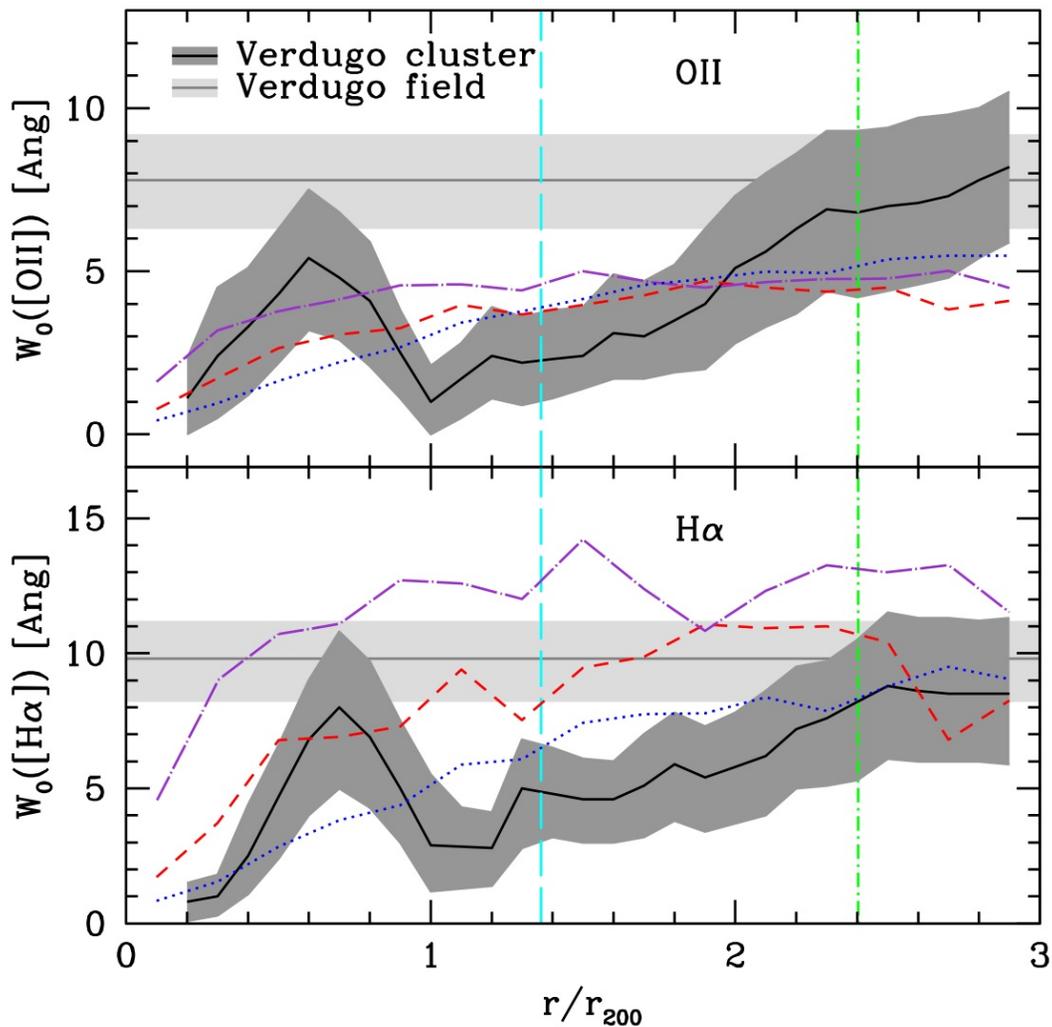


Figure 2.4: Comparison of the (2) cluster spectroscopy with our semi-analytic results using three different implementations of ram pressure stripping. The x-axis is the projected cluster-centric radius divided by  $r_{200}$ , which is defined differently than in the previous section, as explained in the text. The solid black lines with dark grey shading show the observed cluster results and their errors, while the solid grey lines surrounded by light grey show the observed results and errors for field galaxies. The models shown here are the same as in Figure 2.3.

a significantly better fit to the observations. This interesting result implies that, even if only RPS is acting to cause these gradients in SFR, from data such as these alone we cannot distinguish the details of the active RPS mechanism. Nevertheless, our Shocks model is marginally the best match to the data.

### 2.4.2 Comparison to Verdugo et al. (2008) Cluster Spectroscopy

We also compare with the sample of six galaxy clusters spectroscopically observed by (2). The star formation activity of cluster galaxies as a function of cluster-centric radius was investigated using the equivalent widths of the H $\alpha$  and OII emission lines, and a significant drop in star formation in cluster galaxies relative to the field was observed. For the details of the observations, the reader is referred to (2).

To mimic the observational selection and analysis, we determine the mean cluster redshift and the cluster velocity dispersion using the bi-weight location and scale estimators of (92), and select cluster members as galaxies whose redshift is within  $3\sigma$  of the mean cluster redshift. We use the projected cluster radius normalized to  $r_{200}$ , where  $r_{200}$  is here defined to be

$$r_{200} = \frac{\sqrt{3}}{10} \frac{\sigma}{H(z)}. \quad (2.16)$$

We further remove from our cluster galaxy sample those galaxies whose apparent I-band magnitude is above the survey spectroscopic limit  $I = 19.5$ , those whose line-of-sight velocity relative to the cluster center places them clearly outside of the cluster, and those whose absolute I-band magnitude is greater than the limit of  $M_I = -21.4$ , which was imposed on the observed data to treat clusters at different redshifts with the same luminosity restrictions.

The comparison of the observational results of (2) with our ram pressure implementations is shown in Figure 2.4. The solid black lines with dark grey shading show the observed cluster results and their errors, while the solid grey lines surrounded by light grey show the observed results and errors for field galaxies. To this we compare our models; as in Figure 2.3, we show the Bower model (blue dotted), and the Shocks model with (red dashed) and without (purple dot-dashed) RPS. We also show the average virial radius (cyan long-dashed) and accretion radius (green short dot-dashed) for the Millennium clusters considered.

In Figure 2.4 we can see that, as in Figure 2.3, the average virial radius of these clusters is around  $1.4 r_{200}$ , while the average accretion shock radius is at  $2.5 r_{200}$ . This is just at the edge of the radial extent of the observations, and indicates that if our model of the accretion shock is correct, then the current observations are not probing the cluster beyond the radius of environmental effects.

The top panel of Figure 2.4 shows the radial dependence of the OII equivalent width. We see that the three models predict very similar, nearly straight-line radial profiles, all of which are plausible given the size of the observational error bars. Therefore, these cluster galaxy OII data cannot reliably determine whether any form of RPS is causing the observed decline in SFR. However, all of the models tend to predict a shallower slope than is observed.

In the lower panel of Figure 2.4, we see the radial profile of the H $\alpha$  line equivalent width. In this case, both the Bower and Shocks models are plausible with the given error bars, while the Bower

model predicts a lower overall equivalent width and fits the data slightly better.

It can be seen that the data from both (2) and (1) have a local star formation peak at around  $0.5 r_{200}$ . The statistical significance of these peaks, due to the size of the errors, is questionable, and their correlation can not be immediately understood to be due to the same phenomenon, as  $r_{200}$  was defined differently in these two papers. However, if these peaks do in fact represent a general feature of cluster members' star formation rates, then this effect is the product of physics which has not been included in the current model.

## 2.5 Discussion and Conclusions: Implications for Observations of Ram Pressure Stripping

We have addressed the question of the nature of the environmental effects felt by satellite galaxies in galaxy clusters by implementing three different prescriptions of RPS onto the Millennium Simulation, using the semi-analytic galaxy formation model of GALFORM. The three models implemented are the complete RPS of (37), in which all of the hot gaseous halo of a galaxy is removed when it is first identified as a member of the friends-of-friends group of a more-massive halo, the more-nuanced model of (39), in which hot halo gas is only stripped up to a maximum stripping radius determined by the orbit of the satellite, and our new model which incorporates the nuanced Font RPS model but differs from their model by using the radius of the cluster accretion shock as the location at which these effects begin and by utilizing the actual satellite orbit measured directly from the N-body simulation.

Considering the results of applying these models to the detailed merger histories provided by the Millennium Simulation, we see that, in general, all of the models show similar trends in stellar and gas mass with cluster-centric radius, with differences in normalization due to the different ram pressure models and due to the different parameters adopted in the models to obtain a close fit to the properties of the local galaxy population. Further, the gas fraction and therefore the specific SFR is very similar in all of the models considered, with differences only apparent in the regions less than half of a virial radius from the center of the cluster. We see that, in particular, the Shocks and Bower models predict very similar radial average gas mass profiles, with a sharp down-turn at 1.5 virial radii, the average radius of the accretion shock in these clusters. This indicates that if we were to observe the gas mass trend in cluster galaxies, we would not be able to distinguish these two very different implementations of RPS.

As expected, we find that the B-V colors of satellite galaxies in all of the models become redder towards the center of the cluster, and that in all of the models we considered these effects begin just outside the virial radius of the cluster. The onset of this reddening well outside of the virial radius, as well as the decrease in gas mass beyond than the virial radius, indicate either the preprocessing of

satellite galaxies before their accretion onto clusters or the presence of satellites on highly elliptical orbits which have already passed through the central regions of the cluster.

We have compared the results of our accretion shock model of RPS and that of (37) with the observational results of (2) and those of (91) as selected by (1). We see in these comparisons that both of these models are consistent with the observations, given their uncertainties. Thus, we conclude that current observational data on the radial SFR gradient in clusters do not strongly discriminate between different models of cluster environmental effects. However, our models clearly indicate that the presence of RPS has a strong effect on radial trends within clusters and we expect that it will be possible to constrain the details of these models given future observations with larger samples of cluster galaxies.

Recent work on the semi-analytical modeling of satellite galaxy stripping effects in clusters was carried out by (41). As in this work, they have modified previous models, in which all halo gas was stripped as soon as a galaxy became a satellite in a larger halo (93), to include a more-nuanced form of gas stripping. They model cluster-environment effects entirely with tidal stripping, and presume that gas is lost from satellite galaxies in proportion to the loss of dark matter. The baryonic physics of RPS is neglected in the model of (41), while in the model of this work we neglect tidal stripping. Of course, in cosmological structure formation both of these mechanisms have an effect on the stripping of gas from satellite galaxies. However, using the virial scaling relations, we find that the ratio of the forces of ram pressure and tidal stripping is given by

$$\left. \frac{F_{RPS}}{F_{TS}} \right|_{r_{vir}} = A \left( \frac{M_{host}}{M_{sat}} \right)^{2/3}, \quad (2.17)$$

where  $A$  is a constant of order unity, so the force of RPS is larger than that of tidal stripping, for a satellite at the virial radius, by a factor proportional to the ratio of the mass of the host halo to the mass of the satellite to the two-thirds power. By this argument then, at least at the virial radius RPS exerts a larger force on gas in a satellite halo, and so we can be confident that we have included the dominant physical effect in the present work. The results of (41) are compatible with our results, and they also find that several of their models are plausible given the error bars of the data with which they compare.

In conclusion radial trends in galaxy properties around clusters can now be accurately predicted by the GALFORM model of galaxy formation. Measurements of these trends therefore have the potential to place strong constraints on the processes of mass accretion and star formation, both of which are key components of our picture of galaxy formation. We have not discussed the distribution of cluster gas in detail in this work, but it is clearly a key ingredient in any model invoking RPS as a driver of cluster galaxy evolution. Recently, (94) described a more-advanced calculation of cluster gas physics within GALFORM which aimed to match the x-ray properties of clusters. Future work

in this subject should clearly explore both cluster galaxy and x-ray properties in tandem to ensure realistic modeling of the cluster physics. This, coupled with larger samples of cluster galaxies would greatly improve the statistical power of this method as an important constraint on galaxy formation physics.

## Chapter 3

# Angular Momentum Evolution in Dark Matter Halos<sup>1</sup>

### 3.1 Introduction

Recent years have seen much progress in our understanding of the growth of dark-matter halos from initial density fluctuations and the characterization of their properties. For example, it has been possible to find analytic results to describe the merger trees of dark matter-only systems (95; 96; 97; 98; 99; 100). Additionally, high-resolution N-body simulations containing only cold (i.e., non-relativistic at the epoch of kinetic decoupling) dark matter have been performed (42; 101; 102; 103; 104), leading to detailed knowledge of the structure of cold dark-matter halos, such as the fact that halos have an approximately universal radial density profile (105; 106; 107) and are generally triaxial in shape (108).

There is currently a great deal of interest in the angular momentum of particles in dark matter halos, since it is this angular momentum, when transferred to baryons, that creates the disks of galaxies. It is also interesting to look at this topic in the light of the adiabatic contraction (AC) model, which is used to model how the condensation of baryons during the formation of a galaxy affects the mass profile of its host halo (e.g., 109; 29; 38; 110). In this model, it is assumed that the gravitational potential of the system changes very slowly, so that it can be approximated as adiabatic. Further, two simplifying assumptions are usually employed to calculate the density increase resulting from the growth of galaxies at centers of dark-matter halos using AC. First, the gravitational potential of the system (including contributions from both the dark-matter halo and the galaxy) is assumed to be spherically symmetric and the orbits of particles are presumed to be circular, such that the angular momentum of particles is conserved.

---

<sup>1</sup>The work presented in this chapter was adapted from the paper “Angular momentum evolution in dark-matter haloes,” L. G. Book, A. Brooks, A. H. G. Peter, A. J. Benson, and F. Governato, *MNRAS* **411**, 1963 (2011). Reproduced with permission, copyright (2011) by the Royal Astronomical Society. Discussion of the details of the adiabatic contraction method and the N-body simulation have been added to Sections 3.1.1 and 3.2, respectively.

The back reaction of the evolution of those galaxies on the dark matter is poorly understood. The properties of the galaxy population and of the dark-matter response appear to depend quite strongly on the star-formation prescription and strength of feedback processes, among other things, but a physically motivated and vetted mapping between baryonic physics and the evolution in the dark-matter phase-space density is lacking (29; 111; 112; 113; 114; 38; 115; 116; 110; 117; 118).

The development of an accurate model of this back reaction is required to accurately compare theoretical predictions of galaxy formation, both in the context of canonical cold-dark-matter and alternative dark-matter cosmologies with observations. A number of observed galaxy properties, such as the rotation curves of disk galaxies and the associated Tully-Fisher relation (119), depend on the gravitational potential of both baryons and dark matter. In particular, the Tully-Fisher relation is frequently used as a constraint on galaxy evolution processes in models of galaxy formation (29; 120; 121; 36; 122; 123; 38; 124; 110). Moreover, different dark-matter candidates are expected have noticeably different distributions in dark-matter halos in the absence of baryons (e.g., 125; 126; 127; 128); it is not clear how those distributions will change as a result of baryonic physics. Observed density profiles sometimes appear consistent with cold-dark-matter predictions in the absence of baryons, and sometimes do not (129; 130; 131; 132; 133; 134; 135; 136). In the absence of a physically vetted predictive model for the impact on baryonic physics processes on dark-matter distributions in halos, variations on the adiabatic contraction (AC) model are often applied to compare galaxy evolution and dark-matter theories with observations (e.g., 109; 29; 38; 110).

In this work, we examine two key assumptions of the AC model in the context of dark-matter-only simulations of galaxy-mass halos: the adiabaticity of the evolution of the gravitational potential, and the angular momentum distribution of dark-matter particles halos. Specifically, we will investigate the extent to which angular momentum is invariant for individual particles and subsets of particles, as is assumed in the spherically symmetric model that is generally applied (and which will be true even if particle orbits are non-circular). Invariance may be broken for a number of reasons, including a break down of the adiabatic assumption, the non-sphericity of the halo potential and torques from the external mass distribution. Even if the angular momentum of individual particles is not conserved due to the triaxiality of the halo density profile, AC might be applicable if the *distribution* of angular momenta of all particles were invariant with time. If the angular momentum distribution varies with time then the simple model of AC cannot work precisely for even the simplest galaxy-evolution models.

Note that, as we are working with dark-matter-only simulations, we set only lower limits on the level to which AC is not applicable in halos with both dark matter and baryons, as the inclusion of baryonic physics is likely to exacerbate these effects. The observed angular momentum distribution of baryons in galaxies has been shown to deviate significantly from that expected based on simulations (137; 138; 139), so we do not expect the angular momentum distribution that we measure

to be representative of the baryonic distribution of angular momenta. However, the extent of non-conservation of angular momentum in dark-matter-only simulations is likely to be less than that when baryons are added, since baryons and dark matter can exchange angular momentum. Hence, in this chapter, we examine the accuracy of the adiabatic assumption and the level to which the angular momentum distribution changes with time in dark-matter-only simulations, as this places a limit on the possible accuracy of the AC model. Our intention is not to provide a precise quantification of this limit, but merely to highlight its existence and provide an approximate measure of its magnitude.

This chapter is organized as follows: in Section 3.2, we describe the simulations and halos that we analyze, and we describe the particle subsets that we use in Section 3.2.1. We show how the average angular momentum of halo particles with respect to the center of mass is evolving in Section 3.3, and analyze the extent to which the adiabatic approximation is valid in the halos in Section 3.3.1. In Section 3.3.2 we present the evolution of the average angular momentum of halo particles with respect to the center of the halo, and we present the evolution of the angular momentum distribution in Section 3.3.3. Finally, we examine the causes of this evolution in Section 3.4, and discuss our results in Section 3.5.

### 3.1.1 Adiabatic Contraction Model

Here, we provide a review of the AC model as it is commonly applied, with some derivations of the important results. We follow the discussion of Binney and Tremaine (2008) (140).

We first consider a system with a stationary (time-independent) potential, and consider the Hamiltonian for a test particle as a function of generalized coordinates  $\theta_i$  ( $i = 1, 2, 3$ ) and conjugate momenta  $J_i$ ,  $H = H(\theta_i, J_i)$ . If we presume that all three momenta are conserved integrals of the motion, Hamilton's equations give us

$$\begin{aligned}\frac{\partial H}{\partial \theta_i} &= -\frac{dJ_i}{dt} = 0 \\ \frac{\partial H}{\partial J_i} &= \frac{d\theta_i}{dt} = \Omega_i(\mathbf{J}),\end{aligned}\tag{3.1}$$

where  $\Omega_i(\mathbf{J})$  are constants, not dependent on time. We can thus write the time evolution of  $\theta_i$  as

$$\theta_i(t) = \theta_i(0) + \Omega_i t.\tag{3.2}$$

This implies, then, that if we can find three integrals of motion in this potential, then the conjugate coordinates behave in this very simple way. Note that, even if only one or two integrals of motion can be found, this argument still holds for their conjugate coordinates, greatly simplifying the analytical consideration of orbits in such systems.

Now, let us consider in Cartesian space the orbits described by the coordinates  $\theta_i$ . We can see from Eq. (3.2) that the  $\theta_i$  increase monotonically with time. However, these coordinates must be able to represent bound orbits, for which the corresponding Cartesian coordinates  $x_i(t)$  are bounded. This then implies that the  $x_i(t)$  must be periodic functions of the  $\theta_i$ ; let us consider one specific orbit, and renormalize the  $\theta_i$  such that the particle returns to the same position (in  $x_i$ ) when each  $\theta_i \rightarrow \theta_i + 2\pi$ . (This renormalization has no effect on the previous argument, since it simply renormalizes the conjugate momenta by the same amount, without affecting their constancy over the orbit.) We can then expand the orbit in a Fourier series

$$x_i(\theta, \mathbf{J}) = \sum_{\mathbf{n}} \mathbf{X}_{\mathbf{n}}(\mathbf{J}) e^{i\mathbf{n}_i\theta_i}, \quad (3.3)$$

where  $\mathbf{n}$  is a vector of integers. Thus, we can write the orbit as a series with the frequencies multiples of the fundamental frequencies  $\Omega_i$ .

The  $J_i$  thus presented are called the “actions” of the system, while the  $\theta_i$  are the “action angles”. The reason why we are interested in these quantities is their property that, if the potential is changed slowly enough, the actions of individual orbits remain unchanged. This result gives a great deal of power to the theorist, since it allows for a direct understanding of the final state of all of a system’s particles given the initial and final states of the potential. The more-precise condition of a “slow enough” change in the potential is that dynamical time of the change in the potential is much longer than the orbital periods of the particles. Under these conditions, it can be proved with more detail than is needed here that the actions of each individual orbit are unchanged.

Now let us consider a spherically symmetric system, and derive its actions. To do so, we consider two sets of coordinates, spatial ones  $x_i$  and conjugate momenta  $p_i$ , and the action angle variables considered above,  $\theta_i$  and  $J_i$ . We can find a function  $S(x_i, J_i)$  (the “generating function” of the coordinate transformation) such that

$$\begin{aligned} \frac{\partial S(x_i, J_i)}{\partial J_i} &= \theta_i \\ \frac{\partial S(x_i, J_i)}{\partial x_i} &= p_i. \end{aligned} \quad (3.4)$$

We know that the Hamiltonian, evaluated along a particle’s orbit, will give that particle’s energy. Assuming no dissipative properties, this energy is conserved along the orbit. Using the above generating function to write the Hamiltonian as a function of  $x_i$  and  $J_i$ , we have

$$H\left(x_i, \frac{\partial S}{\partial x_i}\right) = E. \quad (3.5)$$

This is known as the Hamilton-Jacobi equation. We can find the actions of a given potential using

the above equation and

$$J_i = \frac{1}{2\pi} \int_{C_i} \frac{\partial S}{\partial x_j} dx_j, \quad (3.6)$$

where  $C_i$  is the path determined by varying  $\theta_i$  from 0 to  $2\pi$  while holding the other action angles fixed. This equation comes from the recognition that (for renormalized  $\theta_i$  as discussed above)

$$J_i = \frac{1}{2\pi} \int \int_{\text{interior of } C_i} d\theta_i dJ_i, \quad (3.7)$$

and since this integral is a Poincaré invariant, it is the same for any set of conjugate coordinates and momenta, so we can equally well write

$$J_i = \frac{1}{2\pi} \int \int_{\text{interior of } C_i} dx_j dp^j. \quad (3.8)$$

Finally, we note that we can rewrite Eq. (3.8) as a line integral using Green's theorem

$$J_i = \frac{1}{2\pi} \int_{C_i} p^j dx_j. \quad (3.9)$$

We finally use Eq. (3.4) to obtain Eq. (3.6).

From these equations, assuming a spherically symmetric potential, we can obtain the forms for the conserved actions. I will skip the algebra here (look at (140) for the details); the result is that the actions for a potential  $\Phi(r)$  are

$$\begin{aligned} J_\phi &= L_z \\ J_\theta &= L - |L_z| \\ J_r &= \frac{1}{\pi} \int_{r_{\min}}^{r_{\max}} dr \sqrt{2E - 2\Phi(r) - \frac{L^2}{r^2}}. \end{aligned} \quad (3.10)$$

In Eq. (3.10),  $\theta$  is the polar angle of the spatial coordinate system, not an action angle, and  $L$  is the angular momentum of the particle.

Now, let us give a brief description of the AC model as it relates to the above results. If our potential is indeed integrable such that we can find three conserved actions, then we can define a distribution function (DF)  $f = f(J_i)$  such that  $f(J_i)d^3J$  that the orbit of a given particle can be described by actions in the volume  $d^3J$ . Note that, since the actions are invariant under adiabatic changes of the potential, as long as this constraint is justified the DF is independent of time. This then gives the AC model in its purest form: given the initial DF, AC tells us that the final DF is the same. However, transforming from a DF to a spatial distribution is non-trivial, and so simple approximations to the true AC model are often used instead. The two most common simplifying assumptions are that the gravitational potential (from both the dark matter and the galaxy) is

spherically symmetric, such that the magnitude of the angular momentum and the radial action are the relevant action variables as we saw above; and that all orbits are circular such that angular momentum is the only nonzero and non-infinite-period angle variable (141). If we assume circular orbits, then the total angular momentum per particle mass can be written as  $|L|/m = \sqrt{GrM(r)}$ , where  $M(r)$  is the mass contained within the orbit of the particle. This is the origin of the commonly used form of the AC model in which the quantity  $rM(r)$  is held constant, leading to a very simple expression for the evolution of halo density with baryon condensation. Occasionally, the assumption of circular orbits is swapped in favor of choosing a variant of the radial action as a conserved quantity (141; 111).

One can thus analytically calculate the final mass profile of the dark matter given the growth of the galaxy. Although these simple models have a mixed track record of matching observations and hydrodynamic simulations of galaxy evolution, they are currently the only predictive models for the effects of galaxy evolution on the dark-matter profile in halos (134; 135; 136; 129; 131; 130; 124; 112; 113; 114; 115; 118; 116; 117).

## 3.2 Simulations

The dark-matter halos used in this study were simulated with the code PKDGrav (142). PKDGrav is a parallel N-body code, which uses a tree data storage structure optimized for massively parallel gravitational simulations. It improves efficiency by using multipole expansions to approximate the gravitational force, and using the local expansion to treat the force due to distant masses. Its name derives from the data structure, k-d trees (“k-dimensional” trees), upon which the code’s more-specialized spatial binary trees are based. Periodic boundary conditions are approximated using the Ewald summation (143).

We used the *Wilkinson Microwave Anisotropy Probe* (WMAP) three-year cosmology ( $\Omega_m = 0.24$ ,  $\Omega_\Lambda = 0.76$ ,  $H_0 = 73$  km/s,  $\sigma_8 = 0.77$ ; 145). Our four halos were originally chosen from a low-resolution volume of 50 Mpc on a side, and selected to span a range of merger histories and spin values at roughly the mass of the Milky Way halo. Each halo was then re-simulated using the volume renormalization technique (146). This approach creates successively finer resolution layers around the halo of interest, allowing for high resolution on one halo while maintaining the large-scale structure (from the original 50 Mpc box) at lower resolution. Importantly, the large-scale structure can deliver tidal torques and angular momentum to the halo.

Table 3.1 lists properties of each of the simulated halos. Three of the four halos have been presented at similar resolution, but including gas through smoothed particle hydrodynamics (SPH), in previous papers (e.g., 147; 148; 149; 150; 151). As examined in (149) and (151), halo h258 has an approximately binary merger that occurs at  $z \sim 1$ , but exhibits a very quiescent evolution afterward.

Table 3.1: Simulated galaxy properties

simulation	$M_{\text{vir}} [M_{\odot}]$	$\lambda^a$	$z_{\text{LMM}}^b$	N within $R_{\text{vir}}^c$
H239	$9.3 \times 10^{11}$	0.01	1.25	$7.6 \times 10^5$
H258	$8.2 \times 10^{11}$	0.03	1.25	$6.7 \times 10^5$
H277	$7.2 \times 10^{11}$	0.03	2.5	$5.9 \times 10^5$
H285	$7.4 \times 10^{11}$	0.02	3.75	$6.1 \times 10^5$

<sup>a</sup> Global spin parameter as defined in (158)

<sup>b</sup> Redshift of last major merger

<sup>c</sup> Number of dark matter particles within the virial radius at  $z = 0$

(150) showed that h277 has a fairly quiescent merger history back to  $z \sim 3$ , while h285 experiences a large number of minor mergers all the way to redshift 0, despite not having a major merger since high redshift. Halo h239 is presented here for the first time. It has a continually active merger history, both major and minor, until  $z \sim 0.5$ .

The mass resolution of the particles that make up these halos is  $1.2 \times 10^6 M_{\odot}$ , with a spline force softening of 350 pc. At each output time step ( $\sim 80$  Myr), high-resolution halos with more than 64 particles (above which the mass function converges, 152; 153) in the volume are identified using AHF (AMIGA’s Halo Finder, 154; 155). AHF adopts results from (156), calculating the overdensity assuming a spherical top hot collapse, under the assumption that the halo has just virialized. The definition for  $\delta_{\text{vir}}$  differs from its value as defined in (157) by the factor  $\Omega(z) * (1 + \delta_{\text{vir}})$ . Thus  $\Delta_c$ , the value for which  $\rho_{\text{vir}} = \Delta_c \rho_{\text{crit}}$ , is  $\sim 100$  at  $z = 0$ . We follow the main halo through time by identifying the most massive progenitor at high redshift.

### 3.2.1 Particle Subsets

There are several different subsets of particles whose properties we analyze here. We take two different approaches to selecting particles; in the first case, we choose particles based on their radius relative to the halo center at  $z = 0$  and follow these same particles back through the simulation (Lagrangian selection), while in the second case we choose particles based on their radius at each time step (Eulerian selection). We use the Lagrangian method to highlight the evolution over the course of the simulation of those particles that will be at a certain radius at  $z = 0$ . The Eulerian approach is complementary, as it shows how the particles at a certain radius at each timestep are related. This is relevant in the context of galaxy evolution, as the baryons condense to a specific region of physical space.

We also use two different methods of averaging the angular momenta of particles, one adding them as vectors, and a second simply adding their magnitudes. These methods also highlight different features of the angular momentum distribution. The vector addition of angular momenta allows us

to see the extent to which a given set of particles have their angular momenta aligned, as a decrease in angular momentum may represent a mixing of angular momentum directions as well as a change in magnitude. Thus, the vector-averaged angular momentum of a radial shell of particles is essentially telling us about the evolution of the “spin” of that shell. In contrast, adding the magnitudes allows us to single out only the change in the magnitude of particle angular momentum, and therefore tells us about changes in the particle orbits. It is this quantity which is relevant for the distribution function and adiabatic invariance.

### 3.3 Halo Evolution

Before we describe the evolution of the angular momentum distribution in the simulations outlined in Section 3.2, we show the results of two simple checks. First, to determine that any evolution in the angular momentum distribution is real and not a numerical artifact, we simulated a  $10^{12} M_{\odot}$ , spherically symmetric halo with equilibrium initial conditions, which in the absence of a spurious numerical effect would maintain a static distribution of angular momenta. These initial conditions were used in (161) in the context of decaying dark matter<sup>2</sup>. The orbits of  $10^6 M_{\odot}$  particles were integrated for 10 Gyr using PKDGrav (142). We found no secular drift in the angular momentum distribution of particles in the halo, just as we expected for a spherically symmetric halo in equilibrium. Further, we found that the angular momentum distribution in radial shells is invariant with time. Thus, we determined that the angular momentum evolution that we see in the cosmological case is not due to numerical effects.

As a further check, we compared the evolution of the total center-of-mass angular momentum of our cosmological models with that found by other authors. In Figure 3.1, we plot the specific halo angular momentum *with respect to the center of mass of the system* of all particles that are within the virial radius of the halo at  $z = 0$  (i.e., the Lagrangian region corresponding to the  $z = 0$  halo) for halos (a) 239, (b) 258, (c) 277, and (d) 285, as a function of the cosmological scale factor  $a$ . Specifically, in the top panel of each panel we show the specific angular momentum  $L$ , as defined above in the center-of-mass frame, divided by its value  $L_0$  at  $z = 0$ . Tidal torque theory, as was first calculated by (159) and was observed in simulations by (160), implies that the specific angular momentum of a halo grows as  $a^{3/2}$  until the halo virializes, at which point it becomes roughly constant. In the lower panels, we show the  $L/L_0$  now divided by  $a^{3/2}$  to explicitly remove this expected dependence. We find that in most of our halos the angular momentum defined with respect to the center of mass of the system behaves in the manner expected from tidal torque theory. Halo 239 also follows this trend but has significantly bumpier evolution than the other halos, possibly as a result of its continuously active merger history.

---

<sup>2</sup>No decay term was included in this work, however.

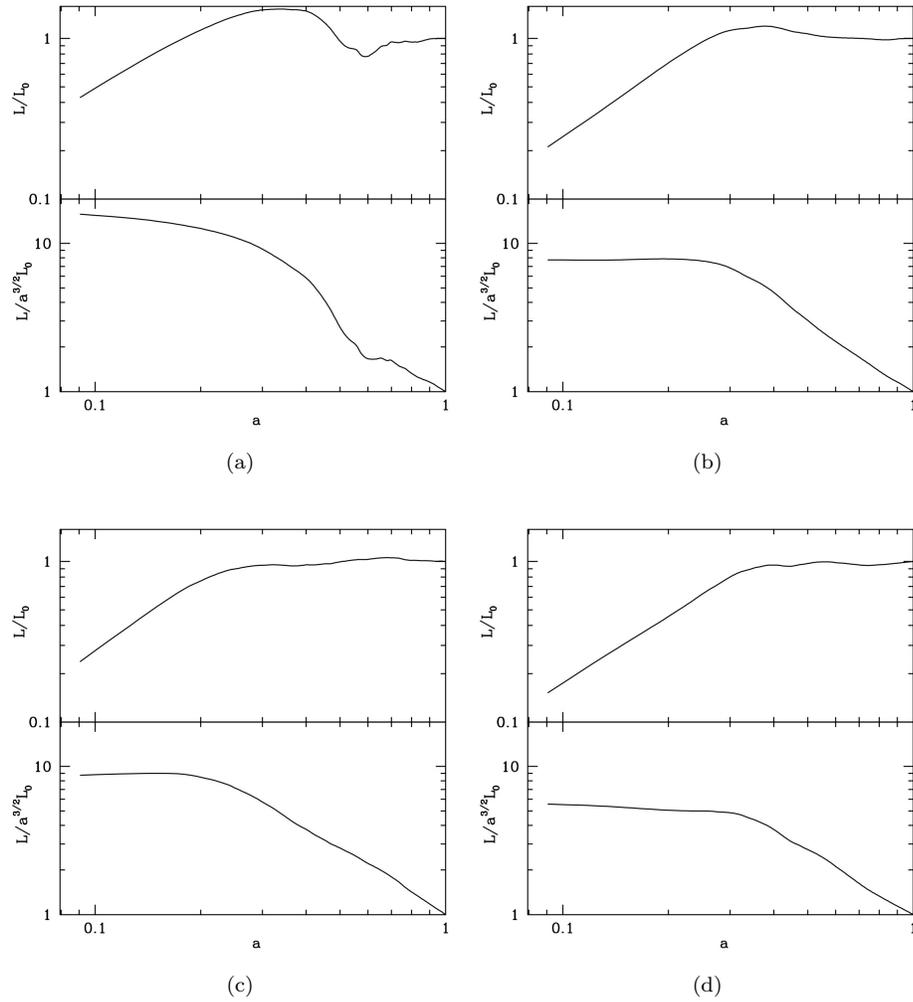


Figure 3.1: The evolution of the specific halo angular momentum, *with respect to the center of mass of the system*, of all particles that are within the virial radius of the halo at  $z = 0$  for halos (a) 239, (b) 258, (c) 277, and (d) 285. All halos follow a basic trend that agrees with the results of tidal torque theory (159; 160).

### 3.3.1 Validity of the Adiabatic Approximation

To determine how well the prescription of adiabatic contraction is followed in simulations, we must first determine to what extent the potential of a halo evolves adiabatically. The adiabatic contraction model is only applicable if the typical time scale for order unity changes to the halo gravitational potential be long compared to the dynamical time of a typical particle in the halo. If this is not the case it would be an indication that the adiabatic contraction approximation may not be used<sup>3</sup>.

To answer this question, we examined the energies and time scales of all particles within the virial radius of each halo. We show these quantities for halo 239 in Figures 3.2 and 3.3. The top panel in Figure 3.2 shows the dynamical time of the entire halo, defined as the virial radius divided by the virial velocity, with the time scale for the potential of the entire halo to change by order unity. The potential change time scale is defined as  $t_{\text{pot}} = |\Phi / (d\Phi/dt)|$ , and has been smoothed on the dynamical time of the halo. It can be seen that, in general, the dynamical time is at least an order of magnitude smaller than the time scale for the gravitational potential to change by order unity, implying that the adiabatic approximation is reasonably accurate in these regions. However, when a halo experiences a merger, the potential change time scale is reduced and the evolution of the potential is non-adiabatic. Such mergers occur here at redshifts 2.2 and 1.2, with mass ratios of 1.1 and 1.9, respectively. The lower panel of Figure 3.2 shows the evolution of the potential and kinetic energy of all particles within the virial radius at each timestep, and shows that the halo reaches a stable quasi-equilibrium around a redshift of  $z = 1$ . Overall, we see that the halo is roughly adiabatically evolving, particularly after this redshift.

As galaxies generally form in the inner regions of their host halos, we would also like to determine whether the adiabatic approximation is valid in the innermost regions of halos. This can be seen in Figure 3.3, which shows the same quantities as Figure 3.2 but for only those halo particles that are within the Eulerian-selected region  $r(z) < 0.1 r_{\text{vir}}(z = 0)$ . The potential change time scale of the inner halo is generally much closer to the dynamical time of this region, implying that the adiabatic approximation is less valid for the inner halo. We can see from the evolution of the kinetic and potential energies that the inner region reaches a state of stable quasi-equilibrium around  $z = 1$ , similar to the outer regions. However, we notice that halo 239 has a more-rapid change in angular momentum than the other halos (see Figure 3.1). Thus, we would expect the other halos to have larger angular momentum change timescales, and for the adiabatic approximation to be more correct in those cases.

These figures show that the approximation that the halo is evolving adiabatically is roughly accurate for the halos studied herein, although the dynamical time is often only an order of magnitude or less smaller than the potential change time scale. The entire halo reaches a quasi-equilibrium by

---

<sup>3</sup>This does not, of course, address the issue of whether changes in the potential remain adiabatic once baryonic physics are included.

around  $z = 1$  here, while halos with more-quiet merger histories reach an equilibrium earlier. However, in the inner regions of the halo where a galaxy might form, the adiabatic approximation is not as good, with the potential time scale usually above but frequently close to the dynamical time scale. Thus we must apply the assumption of adiabaticity with caution in the inner regions of the halo, and to some extent even in the outer regions, especially as work in the context of decaying dark matter has shown that AC works well only if the time scale for changes in the potential is much longer than the dynamical time (161).

### 3.3.2 Angular Momentum Evolution

We next examine the extent to which angular momentum is conserved in our dark-matter-only halos. Here, we consider only angular momentum with respect to the center of the halo, and not the center of mass (as in, e.g., the comparison with 160, at the beginning of this section). The center of the halo is determined by AHF, which uses an adaptive mesh to search for overdensities in the simulation. Once it identifies a halo, it defines the center of the halo as the center of mass of the particles on the highest refinement grid. We show the average particle angular momentum as it evolves over the course of the simulation for various groups of particles, chosen to emphasize particular physical characteristics of the system. In all cases, we bin particles radially so that the innermost ten percent of the halo particles are placed into bin 0, the next ten percent into bin 1, and so forth.

We present this information in Figs. 3.4 and 3.5. In Figure 3.4 we show the evolution of the vector-averaged angular momentum of particles chosen to lie within radial bins interior to the virial radius at  $z = 0$  (Lagrangian radial selection, solid black lines), compared to the evolution of the vector-averaged angular momentum of all particles within the virial radius at  $z = 0$  (blue dashed lines), for halos (a) 239, (b) 258, (c) 277, and (d) 285. The angular momenta in these plots were vector-added and then divided by the particle number to obtain the average, i.e., the specific angular momentum. Particles were placed into 10 radial bins at  $z = 0$ , and we show the 0th (innermost with respect to the halo center), 5th (middle), and 9th (outermost) bins in the top, middle and bottom panels of each plot. We also show the redshift at which the average radius of particles in this radial bin passes within the virial radius of the halo (green dot-dashed vertical lines), and the ratios of the angular momentum at  $z = 1$  and at the redshift of their accretion onto the halo to illustrate quantitatively the extent of angular momentum loss. In this plot, we can see that generally, the particles that end up in the inner regions of these halos at  $z = 0$  lose a larger fraction of their angular momentum than those that end up in the outer regions, and the angular momentum of these  $z = 0$  inner particles also fluctuates more rapidly. The blue dashed line, which shows the evolution of the average angular momentum of all particles within the virial radius at  $z = 0$ , shows the average behavior of the particles in all bins, and, like the binned behavior, always decreases.

As mentioned earlier, the angular momentum of particles selected based on their Eulerian radius

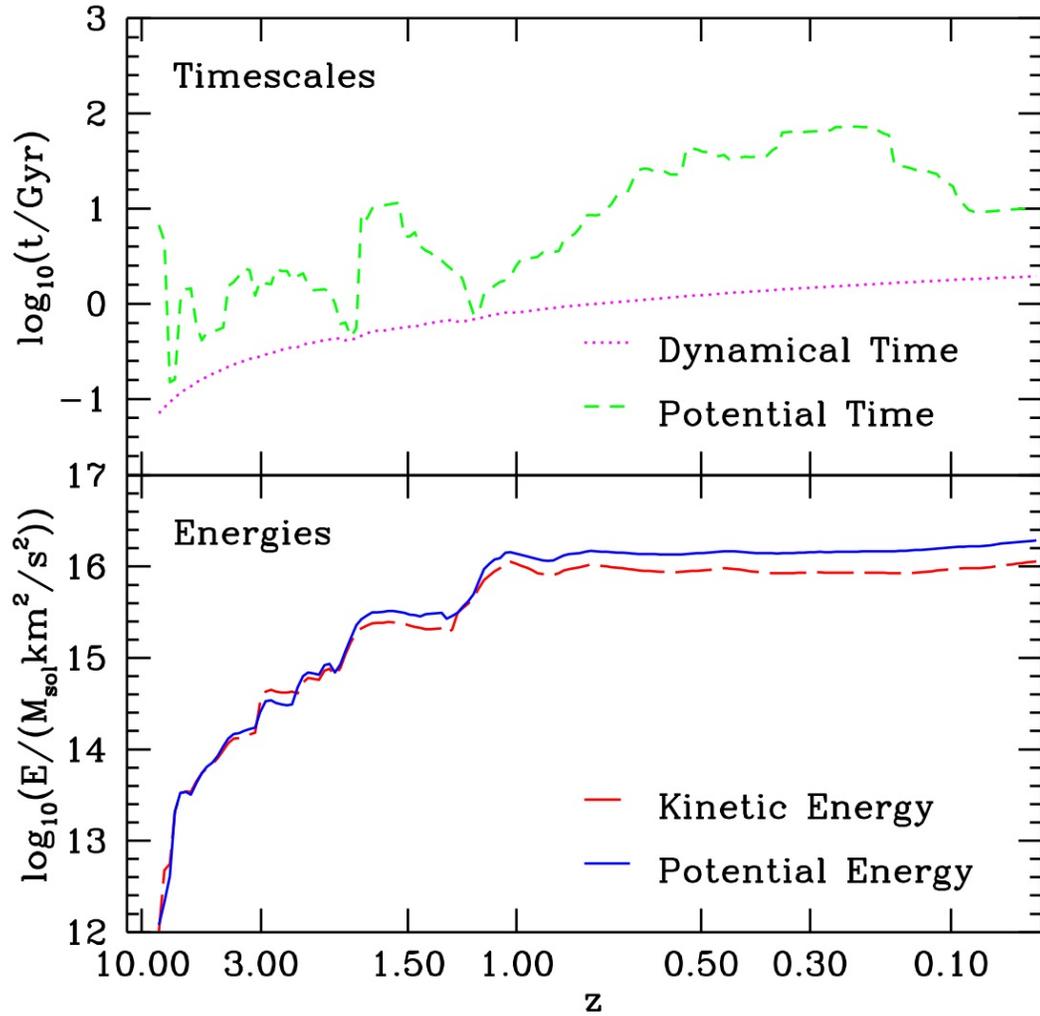


Figure 3.2: Energies and time scales of all particles within the virial radius at each timestep. The halo shown is halo 239. The top panel shows the dynamical time of the entire halo vs. the time scale for changes to the gravitational potential, which has been smoothed over the dynamical time of the halo. The bottom panel shows the evolution of the potential and kinetic energy of all particles within the virial radius at each time step.

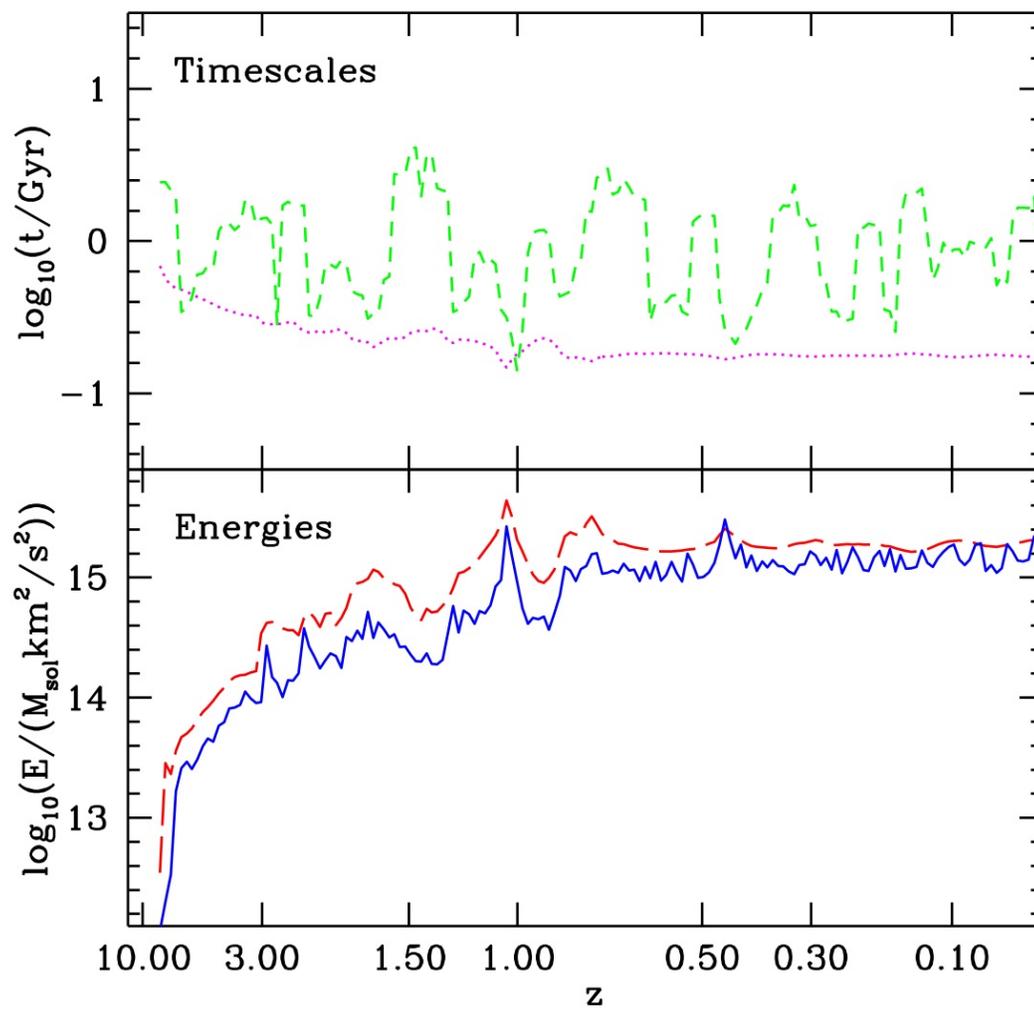


Figure 3.3: This figure shows the same quantities as Figure 3.2, but for only those halo particles that are at each time step within a physical radius  $0.1r_{vir}(z=0)$  of the center of the halo.

is an interesting compliment to the Lagrangian case as it shows how those particles that are in a given radial bin at each timestep evolve in angular momentum, and thus more clearly illustrates the evolution of the radial angular momentum profile of the halo. By contrast, the angular momentum evolution of radially Lagrangian-selected particles shows how the angular momentum of those particles which end up in certain radial bins have changed over time.

The quantities shown in Figure 3.5 are similar to those in Figure 3.4, except that the particles here (specifically, all particles that are within the virial radius at  $z = 0$ ) were binned radially at each time step in an Eulerian fashion, so that the “bin 0” line shows the specific angular momentum of the ten percent of particles that are closest to the center of the halo at each time step. The magenta long-dashed line shows the specific angular momentum of all particles that are within the virial radius at each time step. This quantity goes to zero at the beginning of the simulation as the virial radius is zero at the first time step and increases from there. This line tends to increase because the virial radius increases with time. This is different from the binned values which include all particles that are within the virial radius at  $z = 0$  and thus always contain a nonzero number of particles. All of the angular momenta in this plot were vector added as in the previous figure.

In Figure 3.5, there is no clear pattern as to which bin loses a larger fraction of its angular momentum, although the inner bin generally has less angular momentum. This shows that, in general, particles throughout the halo lose a comparable fraction of their angular momentum, but those in the inner regions at each timestep tend to have less angular momentum than those in the outer regions. However, we saw in Figure 3.4 that when the particles were binned in a Lagrangian fashion, the particles in the inner regions lost a larger fraction of their angular momentum as well as having lower angular momentum in general. Noting that both Figures 3.5 and 3.4 must and do converge to the same value at  $z = 0$ , this implies that the particles that end up in the innermost bin at  $z = 0$  both begin with a larger angular momentum than those particles that are in the innermost bin at the beginning of the simulation, and subsequently lose this angular momentum at a faster rate than those particles that are at each timestep at the center of the halo. This behaviour is to be expected, as the particles at the center of the halo at each timestep will tend to have a low angular momentum, and it is reasonable to expect that this value will tend to drop rather slowly since we are choosing new, low-angular-momentum sets of particles at each timestep in the innermost bin.

We also examined the evolution of the angular momentum in radial shells when averaged using the magnitudes of particle angular momenta rather than vector-averaging. We have omitted plots of the magnitude-averaged cases as they are similar to Figs. 3.4 and 3.5, and have instead included the relevant information about their evolution in Table 3.2. As we noted in Section 3.2.1, magnitude-averaged angular momenta tell us about the evolution of particle orbits, thus providing complementary information to the vector-averaged angular momenta that tell us about the evolution of the spin of radial shells. Further, it is the magnitude of angular momentum which is most relevant

Table 3.2: Ratios of the average binned angular momentum at  $z = 1$  and at accretion of the radial bin to the value of the bin’s average angular momentum at  $z = 0$  for both Lagrangian and Eulerian selected radial bins, differing from Figures 3.4 and 3.5 in that the angular momenta are magnitude-averaged rather than vector-averaged

	Lagrangian						Eulerian					
	Bin 0		Bin 5		Bin 9		Bin 0		Bin 5		Bin 9	
	$L_1/L_0$	$L_a/L_0$	$L_1/L_0$	$L_a/L_0$	$L_1/L_0$	$L_a/L_0$	$L_1/L_0$	$L_a/L_0$	$L_1/L_0$	$L_a/L_0$	$L_1/L_0$	$L_a/L_0$
H239	2.24	3.45	1.44	1.27	1.41	1.07	1.29	1.22	1.2	1.2	2.52	1.05
H258	2.42	5.96	1.21	1.23	1.3	1.03	1.31	1.31	1.08	1.29	1.81	0.99
H277	1.63	3.31	1.31	1.45	1.38	1.06	1.18	1.64	1.19	1.08	1.68	1.01
H285	1.22	3.55	1.11	1.1	1.05	0.99	0.95	1.36	1.04	1.12	1.28	1.02

for the distribution function and adiabatic invariance.

The magnitude-averaged angular momenta evolve much more smoothly than the vector-averaged ones, indicating that cancellation between particles in the vector-averaged case causes more variability than in the magnitude-averaged case. We show in Table 3.2 the values of  $L[1]/L[0]$  and  $L[z_{acc}]/L[0]$  for these cases. The particles that end up in the inner regions tend to have a higher angular momentum to start with than those particles that are in the innermost bin at the beginning of the simulation, and subsequently tend to lose this angular momentum more rapidly, though the effect is not as pronounced as in Figure 3.4. The more pronounced nature of this effect in Figure 3.4 indicates that the inner particles also tend to have their directions scrambled more (i.e., are more thoroughly virialized) than those in outer regions.

The mean magnitude-averaged angular momentum in each radial bin also decreases from the time of accretion to the present, while the vector-averaged angular momenta in Figure 3.5 in general decrease more since their accretion and have lower values. In both cases, the inner regions lose a larger fraction of their angular momentum than the outer regions. It should be noted that the trend of inner particles losing more angular momentum than outer particles was also found by (160). The generally larger decrease over time in the vector-averaged angular momenta implies that the scrambling, or decoherence, of angular momentum direction in the halo tends to increase over time.

Thus, we have found that those particles that end up in the inner regions of halos lose a larger fraction of their angular momentum than other halo particles, and that all of the particles in the halo tend to have the direction of their angular momenta scrambled progressively more over time. Further, we see that those particles that end up in the inner regions of the halo at  $z = 0$  tend to start out with a higher angular momentum than those particles that are in the innermost bin at each timestep, and that they tend to lose this angular momentum more quickly.

### 3.3.3 Distribution of Angular Momentum

Next, we consider changes to the distribution of the magnitudes of the particle angular momenta, not just the average as shown in the last section. In Figure 3.6, we show the distributions of angular mo-

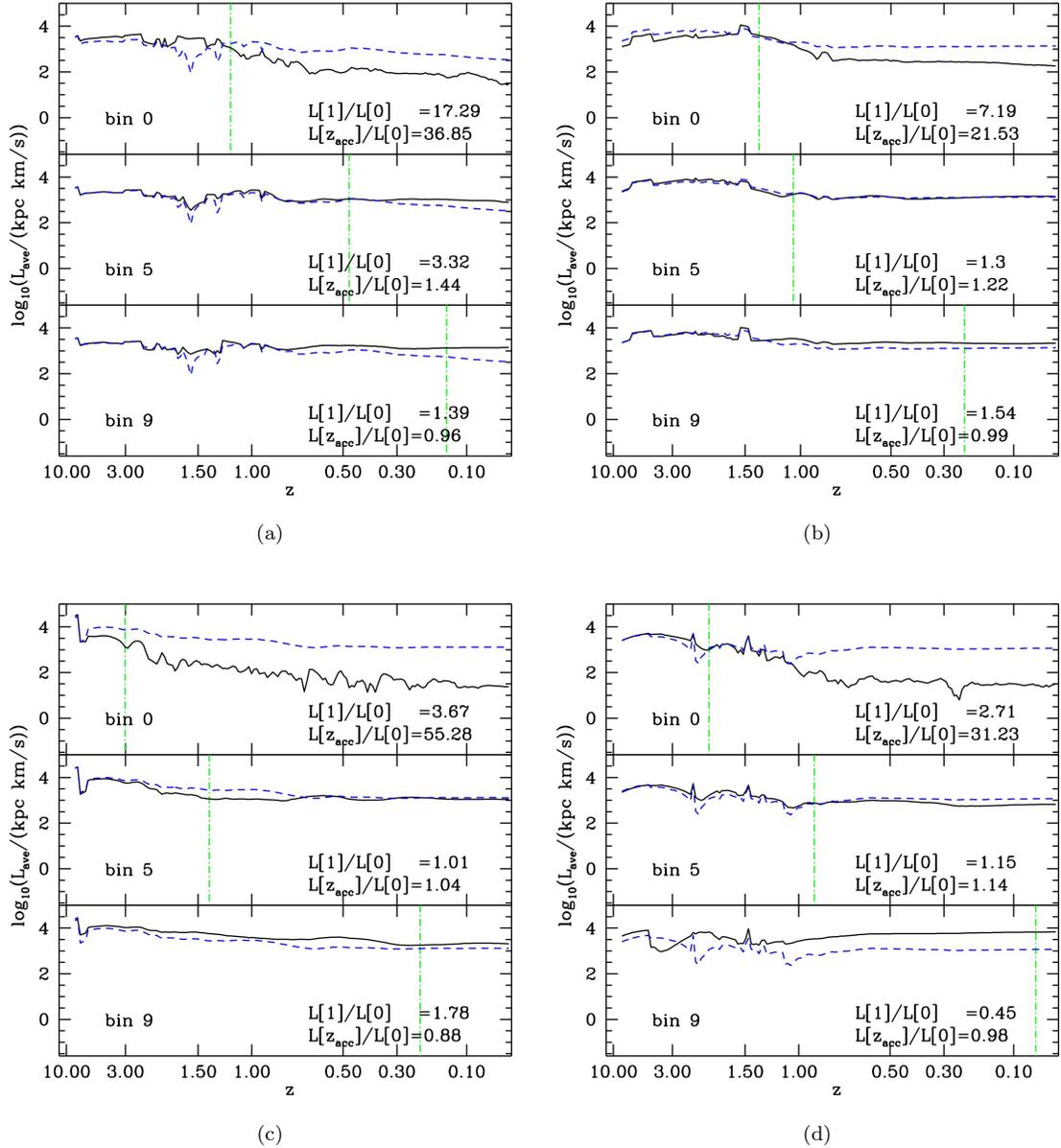


Figure 3.4: The evolution of the average angular momentum of particles chosen in a Lagrangian fashion to lie within the 0th, 5th, and 9th of ten radial bins (solid black lines) compared to the evolution of the average angular momentum of all halo particles that are within the virial radius at  $z = 0$  (blue dashed lines), for halos (a) 239, (b) 258, (c) 277, and (d) 285. The angular momenta in these plots were vector-averaged, and we show the redshift at which each bin is accreted onto the halo (green dot-dashed vertical lines). In each panel is given the ratios of the angular momentum at  $z = 1$  and at the redshift of their accretion,  $z_{\text{acc}}$ , onto the halo to the value at  $z = 0$ .

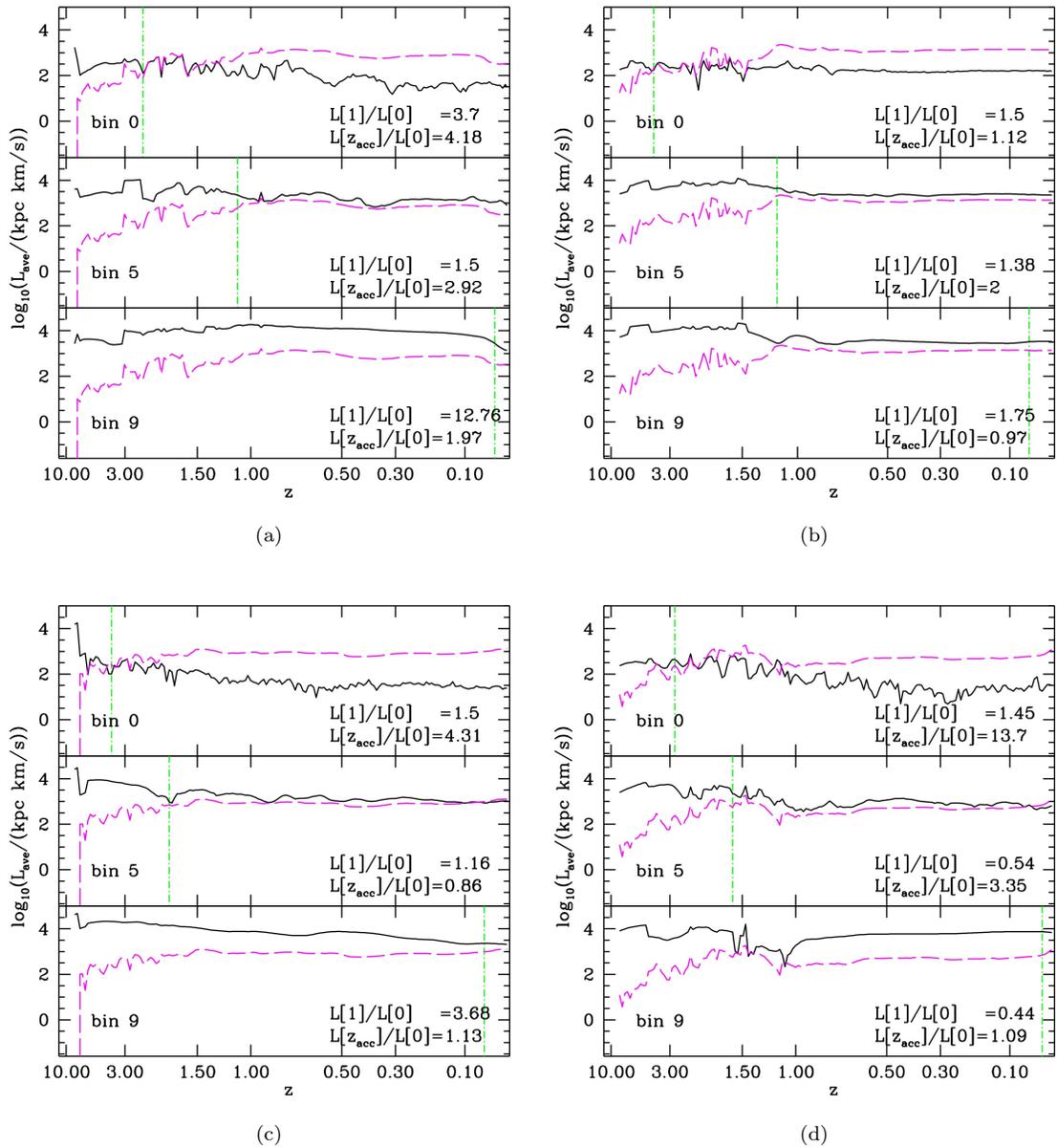


Figure 3.5: Similar to Figure 3.4, except that the particles here are binned radially in an Eulerian fashion. The magenta long-dashed line shows the specific angular momentum of all halo particles within the virial radius at the given redshift.

menta of all halo particles and inner halo particles for all four halos. The top panel shows the angular momentum distribution of all particles within the virial radius chosen in an Eulerian fashion. Here, the number of particles, as well as their mean angular momentum, increases with time. In the middle and bottom panels, we show the evolution of “inner” halo particles chosen in an Eulerian fashion, defined as those particles within  $0.1 r_{\text{vir}}(z = 0)$  (middle) and the innermost 90000 particles corresponding to the innermost  $10^{11} M_{\odot}$  (bottom). We also show in these panels the Boltzmann parametric fit  $N = A (\log_{10}(L/(\text{kpc km/s})) - \log_{10}(L_0))^2 \exp(-(\log_{10}(L/(\text{kpc km/s})) - \log_{10}(L_0))^2/2a^2)$ . The number of particles in the middle panel increases with time, while in the bottom panel it remains constant.

The inner particles have a lower average angular momentum than all halo particles, as is to be expected. The evolution of the distribution of angular momenta of all halo particles is qualitatively similar in all four halos, with the average angular momentum of halo particles increasing with time as higher angular momentum particles are accreted.

Looking at the middle and lower panels, we can see that for halo 277 the angular momentum distribution stays roughly constant in the inner regions, with few extra particles entering  $0.1 r_{\text{vir}}$  at each time step. There is, however, a gradual decline in average angular momentum with time. This smooth evolution is related to the relatively quiescent evolution of halo 277. By contrast, we find that the innermost  $10^{11} M_{\odot}$  of halo 258 have a significantly lower average angular momentum after  $z = 1$  than before, reflecting a major merger around this redshift. In general, all of these angular momentum profiles have a similar shape, which we find is fit well by the above fitting function.

In all the halos, though, there are changes of the angular momentum distribution with redshift. This implies that the angular momentum of a particle is strictly not an adiabatic invariant (or, perhaps, the time scale for changes is not sufficiently slow for the adiabatic approximation to hold). This means that assuming that the distribution function is constant when expressed in terms of the angular momentum during periods when a galaxy is growing is formally incorrect. In practice, this lack of precise conservation simply sets a limit on how accurate the AC approach can ever be.

As a test, we show the fit parameters to the Boltzmann distribution as a function of redshift. In Figure 3.7 we show the evolution of the parameters of the Boltzmann fitting function  $N = A(\log_{10}(L/(\text{kpc km/s})) - \log_{10}(L_0))^2 \exp(-(\log_{10}(L/(\text{kpc km/s})) - \log_{10}(L_0))^2/2a^2)$ , as shown in Figure 3.6, for each of the four halos we consider. Solid lines correspond to the particles selected to lie within  $0.1 r_{\text{vir}}(z = 0)$  in an Eulerian fashion, as in the middle panel of Figure 3.6, while the dotted lines correspond to the innermost  $10^{11} M_{\odot}$  as shown in the bottom panel of Figure 3.6. The x-axis shows the redshift. Note that the discrete nature of the curves is a byproduct of the discrete sampling of fit parameters used to determine the minimum  $\chi^2$  fit. The trends in these plots are qualitatively related to the merger history of the halos. We know that halo 277 (lower left) is quite quiescent since around  $z = 2$ , which is reflected in the mostly smooth evolution of its fit parameters. Halo 258

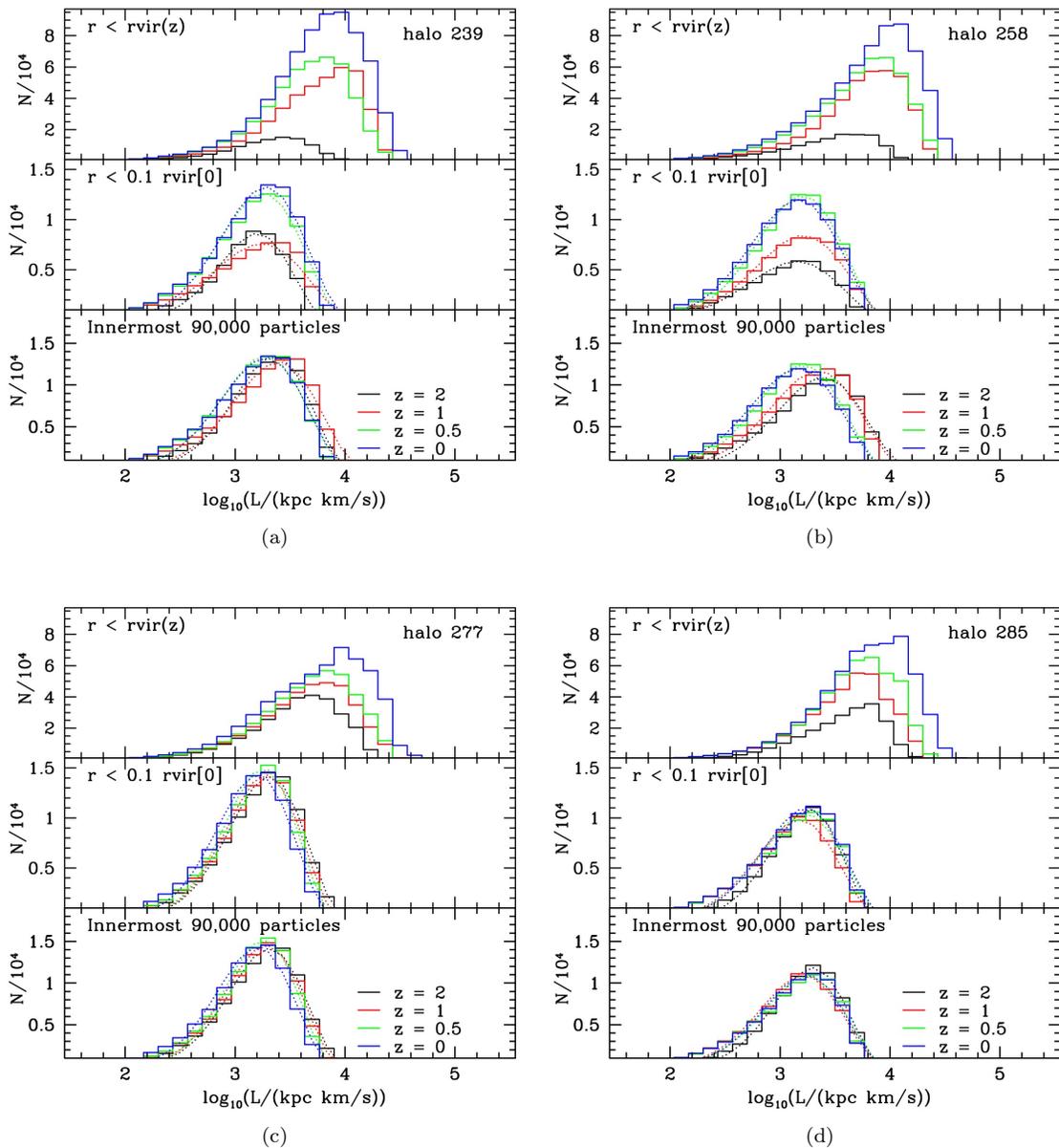


Figure 3.6: The evolution of the angular momentum of halo dark-matter particles for halos (a) 239, (b) 258, (c) 277, and (d) 285. In each figure, the top panel shows the angular-momentum distribution of all particles within the virial radius selected in an Eulerian fashion. In the middle and bottom panels, we present the evolution of “inner” halo particles, also selected in an Eulerian fashion, chosen radially (middle) and by mass (bottom). We also show in these panels the fit  $N = A (\log_{10}(L/(\text{kpc km/s})) - \log_{10}(L_0))^2 \exp(-(\log_{10}(L/(\text{kpc km/s})) - \log_{10}(L_0))^2/2a^2)$  as dotted lines.

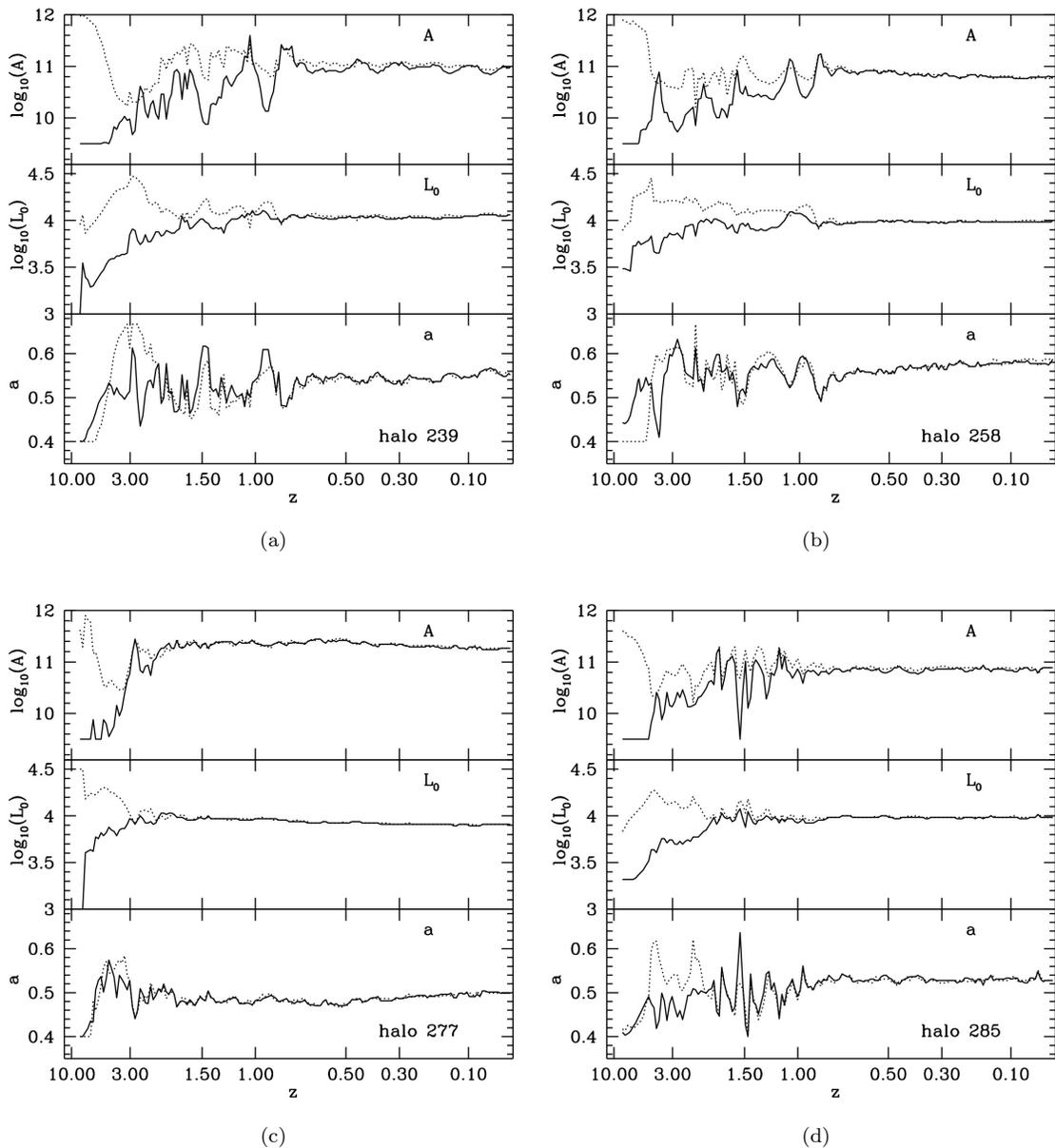


Figure 3.7: Evolution of the parameters of  $N = A (\log_{10}(L/(\text{kpc km/s}) - \log_{10}(L_0))^2 \exp(-(\log_{10}(L/(\text{kpc km/s}) - \log_{10}(L_0))^2/2a^2))$ , the Boltzmann fitting function in Figure 3.6, shown here for each of the four halos, where solid lines correspond to the particles selected to lie within  $0.1 r_{vir}(z=0)$  as in the middle panel of Figure 3.6, while the dotted lines correspond to the innermost  $10^{11} M_{\odot}$  as shown in the bottom panel of Figure 3.6. The  $x$ -axis shows the redshift. The trends in these plots are qualitatively related to the merger history of the halos.

(top right), whose last major merger occurred around  $z = 1$ , shows roughly smooth evolution after this point. Halos 239 and 285 (top left and bottom right), which have more-active merger histories up to the present time, have more variation in their parameter evolution. As yet, though, we do not have a quantitative description of the fit parameters as a function of halo-evolution properties, nor do we know if a quantitative fit is possible. Investigations into such a description would likely require a far larger statistical study than we have presented here.

### 3.4 Causes of Angular Momentum Evolution

We have seen that there is significant evolution over the history of these halos of the average angular momentum (Section 3.3.2) and of the angular momentum distribution (Section 3.3.3). We now investigate the origin of the torques that are causing this non-conservation of angular momentum. There are several possibilities for the origin of these torques, including external structure, non-sphericity of the halo itself, or infalling substructures that we would expect to play a role since we saw in the previous section that merger events are correlated with changes in the distribution of angular momenta.

To investigate the relative importance of these mechanisms to the inner halo, we show in Figure 3.8 the time scale associated with torques on the inner halo due to particles that are members of the host halo at  $z = 0$  (blue dotted lines) and due to particles which are members of clumped substructures which will be within the virial radius at  $z = 0$  (red dashed lines). We define the inner halo as those particles within the Eulerian-selected region  $r(z) < 0.1 r_{\text{vir}}(z = 0)$ . We compare this with the time scale for the angular momentum to change due to all torques, defined as the ratio of the total angular momentum of the inner region to its time derivative  $L_{\text{in}}/[dL_{\text{in}}/dt]$  (black solid lines), and the inner halo dynamical time (magenta dot-dashed lines). Note that a longer time scale here implies a weaker torque, and conversely a small time scale implies a strong torque that can change the angular momentum relatively quickly.

From these figures we can see that, in general, the time scale for the angular momentum to change by order unity due to all external and internal torques is at least an order of magnitude less than that for the angular momentum to change due to torques only by other particles in the halo or by clumped substructures, implying that the torques due to non-sphericity of the host halo and clumpy substructure are far weaker than torques due to external structure in the angular momentum evolution of the inner halo. Thus, external torques are the dominant source of angular momentum evolution of the inner halo as a whole. We also observe that all of these time scales become large relative to the inner halo dynamical time during quiescent merger phases, for example after redshift  $z \sim 1$  for halo 258, while the angular momentum change time scale becomes comparable to the dynamical time during mergers both major and minor. This is reasonable, as it implies

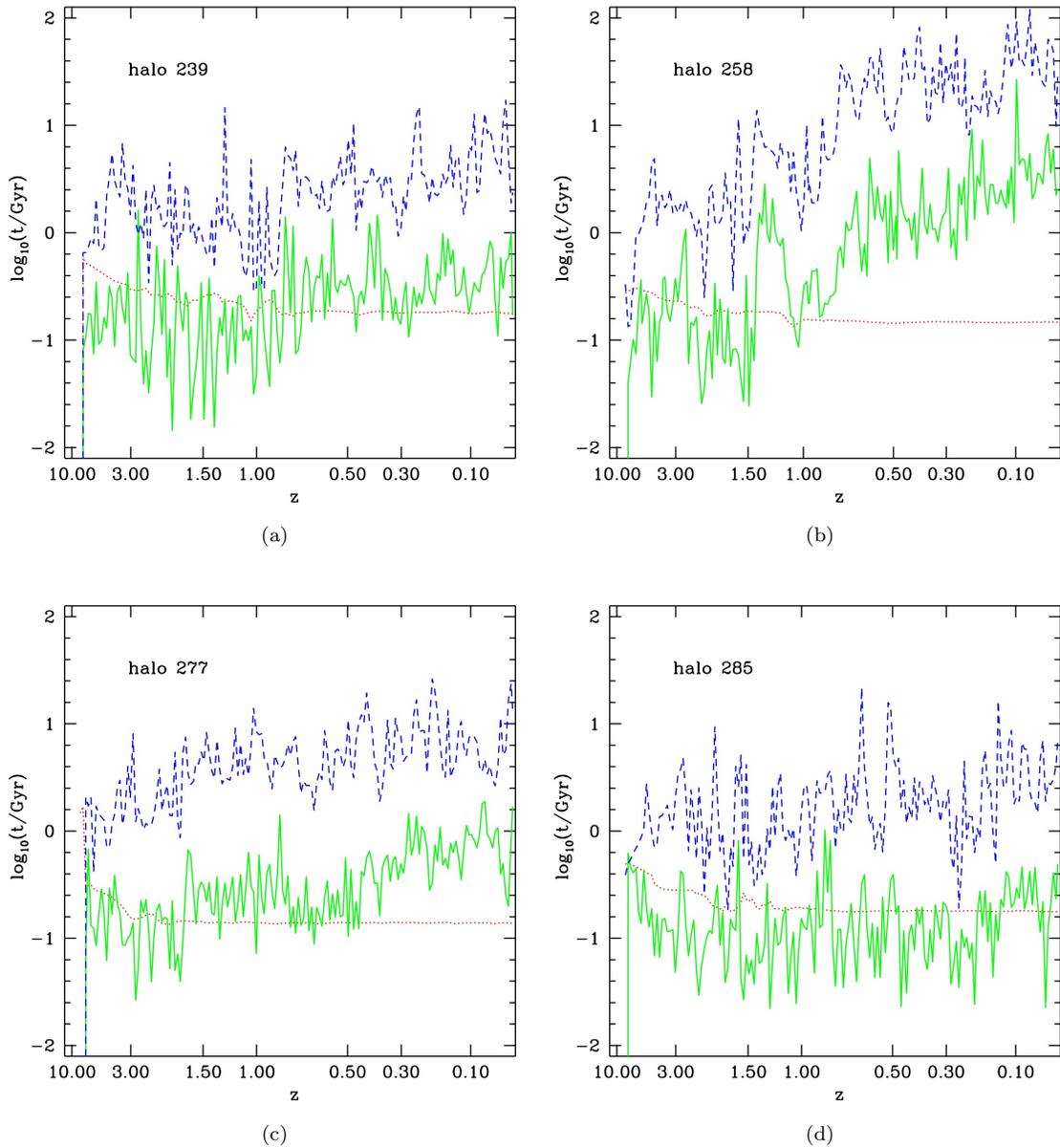


Figure 3.8: Time scales associated with torques on the inner halo due to outer halo particles (blue dotted lines), clumpy substructure particles (red dashed lines), and all torques (black solid lines), for halos (a) 239, (b) 258, (c) 277, and (d) 285. Also plotted is the inner halo dynamical time (magenta dot-dashed lines).

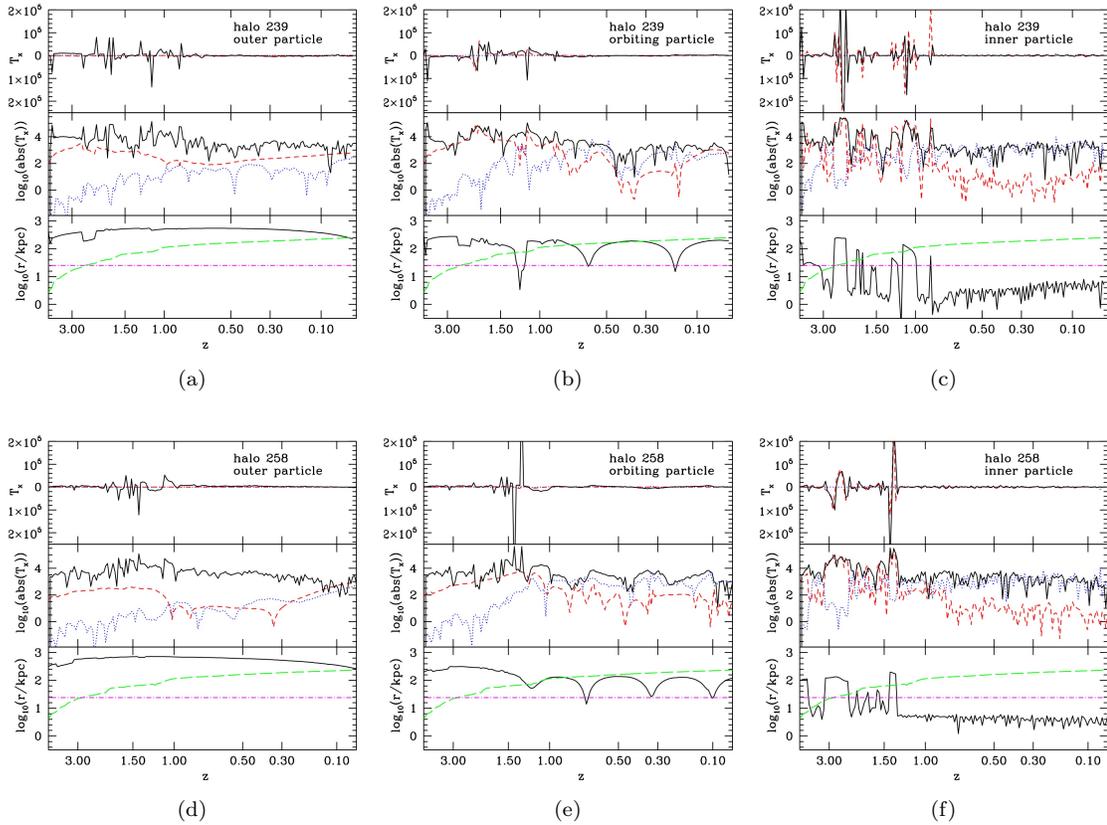


Figure 3.9: Total torque on selected halo particles (solid black lines) compared with the torque due to halo particles (blue dotted lines) and particles which are members of clumped substructures that will be within the virial radius of the halo at  $z = 0$  (red dashed lines). The torque is shown on a linear scale in the top panel, and its absolute value is shown on a log scale in the middle panel of each figure. The units of the torque in the top two panels are  $\text{km kpc s}^{-1} \text{Gyr}^{-1}$ . In the bottom panel, we show the radius of this particle (black solid lines) and compare it with the virial radius of the halo (green long-dashed lines), and the line  $0.10r_{\text{vir}}(z = 0)$  (magenta dot-dashed lines). The top three plots show three particles in halo 239, while the bottom three are from halo 258. For each halo we have chosen one “outer”, one “orbiting,” and one “inner” particle to show here.

that the angular momentum changes slowly during periods of quiescent evolution. Note that the angular momentum change time scale of halo 277 is closer to the dynamical time scale during its quiescent evolution after  $z = 3$ , while the angular momentum change time scale of halo 258 becomes significantly larger than its dynamical time scale when it is quiescent after  $z \sim 1$ . This indicates that the speed of the evolution of angular momentum during quiescent phases does vary between halos, probably due to the environment of the halo which is producing external torques.

We address the role played by these torques on individual particles in Figure 3.9, in which we show the total torque  $dL/dt$  on selected particles (solid black lines), and compare this with the torque on these particles due to particles that are currently bound to the main halo and no other substructure (blue dotted lines) and particles which are members of clumped substructures at the given redshift that will be within the virial radius of the halo at  $z = 0$  (red dashed lines). These torques are shown on a linear scale in the top panel, and their absolute value on a log scale is shown in the middle panel of each figure. The difference between the sum of the torques due to halo particles and clumped particles and the total torque is the torque due to external structure, from particles that are not members of the halo at  $z = 0$ . The units of the torque in the top two panels are  $\text{km kpc s}^{-1} \text{Gyr}^{-1}$ . In the bottom panel, we show the radius of each individual particle (black solid lines) and compare it with the virial radius of the halo (green long-dashed lines), and the line  $0.10r_{\text{vir}}(z = 0)$  (magenta dot-dashed lines) below which particles are classified as inner halo particles. The top three plots show three particles in halo 239, while the bottom three are from halo 258. For each halo we have randomly chosen one particle that spends most of its time outside the virial radius (“outer particle”), one that orbits the halo, entering the virial radius on its closest approach (“orbiting particle”), and a particle that spends much of its time in the inner regions of the halo (“inner particle”).

In general, we see that the torque on the individual particles that we have chosen is dominated by smaller subhalos before the particle is accreted, but that this torque falls below that due to halo particles after the particle is accreted onto the halo. However, in Figure 3.8 we saw that the angular momentum of the inner halo as a whole is dominated by external torques, with mergers also tending to lower the inner halo angular momentum. When taking the individual particle behavior shown in Figure 3.9 together with the behavior of the entire inner halo in Figure 3.8, we see that the importance of torques due to halo and clumped particles in the evolution of some of the individual particles goes away when considering the inner halo as a whole. Also, external structure, which plays a relatively minor role in the evolution of some of the individual particle angular momenta, is a dominant effect on the evolution of the total angular momentum of the inner halo. We may interpret this to say that external structure tends to torque each particle in the entire halo coherently, while torques due to other halo particles and clumpy substructures can be larger on individual particles but are incoherent when considering the entire inner halo. If this is the case, then the torques due to halo

particles will tend to cancel out when averaged over many individual particles in the inner halo, explaining the relative unimportance of these torques for the entire inner halo, and why external structure comes to dominate the evolution of the inner halo.

### 3.5 Discussion

We have analyzed whether the assumptions of the AC model are valid in the case of simulated dark matter halos. Specifically, we have looked into whether such halos can be characterized as adiabatically evolving, and whether their angular momentum or angular momentum distributions are conserved as we would expect from the assumption of spherical symmetry in the model of AC most commonly applied to dark matter halos. We have found that the assumption of adiabatic evolution of halos is not completely correct, that the angular momentum of regions of the halo is not an adiabatic invariant as is assumed in the AC model, and that its distribution also varies. We have investigated the sources of these torques, and found that external torques are the main source of torque on the halo as a whole.

In greater detail, our major results are as follows:

1. Halo particles follow the trend of center-of-mass angular momentum expected from tidal torque theory.
2. In general, the gravitational potential of our halos changes on time scales larger than the dynamical time in the halos, which is a prerequisite for adiabatic invariance. For the halo as a whole, the potential change time scale is generally around an order of magnitude larger than the dynamical time scale, while for the inner halo the potential change time scale is only a factor of a few larger than the dynamical time scale.
3. Angular momentum is lost from all particles as halos virialize, more from those particles that end up in the center of halos. We find that both the vector- and magnitude-averaged angular momentum in fixed radial bins about the halo center decreases with time, by a few tens of percent to factors of a few. The vector-averaged quantities usually decrease more than the magnitude-averaged ones, implying that the directions of angular momenta in radial bins become progressively more misaligned over time.
4. The distribution of angular momentum magnitudes is well fit by a simple Boltzmann fitting function. Trends in the evolution of these distributions qualitatively reflect the merger history of the halo.
5. External torques dominate the angular momentum evolution of the inner halo, while substructure and halo non-sphericity torques can be important for the angular momentum evolution of

individual particles. The dominant role of external torques in changing the angular momentum of halo particles agrees with the results of (162), who found the evolution of halos during baryonic condensation to be mostly reversible when external torques were not included.

We find that halo particles are losing angular momentum even in these dark-matter-only simulations, and that the net angular momentum loss of the inner regions of the halo is due mainly to external tidal torques. The amount of this decrease depends on what region we choose and how we add the angular momenta, but the ratio of angular momenta at  $z = 1$  and  $z = 0$  can vary from a few tenths to a few. The distribution of angular momenta in both the halos as a whole and in the innermost parts in which a galaxy would live are not time-invariant, which means that the spherically symmetric form of adiabatic contraction that is typically applied to dark-matter halo profiles is not strictly valid.

While we find indications that the evolution in the angular-momentum distribution is correlated with the halo accretion histories, we have not found a quantitative description of these changes. Larger statistical studies are required to determine if a quantitative relation between the angular momentum distribution and halo accretion history can be established.

Many previous studies of AC have found it to overpredict the effect of baryon condensation on dark matter density in the centers of halos under the assumption that angular momentum is conserved. Our finding that angular momentum is lost from all particles in the halo over time likely exacerbates this problem, as it implies that even more mass would be concentrated in the center due to dark matter only interactions. The interaction of the angular momentum loss of dark matter particles observed here with baryonic physics is an interesting direction for future research.

That angular momentum does not behave as a perfect adiabatic invariant is not surprising since, for example, halos are non-spherical and evolve with time. Our results serve to highlight the magnitude of this issue and serve as a caution to applications of the adiabatic contraction approximation—there is a limit to the precision which we can reasonably expect it to provide. The fact that recent simulations show that no single model of adiabatic contraction works well in all cases (e.g., (112)) suggests that this limit may have been reached. Overall, our results imply that a fundamental limit to the applicability of current adiabatic contraction models which should be kept in mind when applying these approximations to the effects of galaxy formation on dark matter halos.

## Chapter 4

# Astrometric Effects of a Stochastic Gravitational Wave Background <sup>1</sup>

### 4.1 Introduction and Summary of Results

#### 4.1.1 The Stochastic Gravitational Wave Background

There is great interest in detecting or constraining the strength of stochastic gravitational waves (GWs) that may have been produced by a variety of processes in the early Universe, including inflation. The strength of the waves is parameterized by their energy density per unit logarithmic frequency divided by the critical energy density,  $\Omega_{\text{gw}}(f)$ . Current observational upper limits include (i) the constraint  $\Omega_{\text{gw}} \lesssim 10^{-13}(f/10^{-16} \text{ Hz})^{-2}$  for  $10^{-17} \text{ Hz} \lesssim f \lesssim 10^{-16} \text{ Hz}$  from large angular scale fluctuations in the cosmic microwave background temperature (163); (ii) the cosmological nucleosynthesis and cosmic microwave background constraint  $\int d \ln f \Omega_{\text{gw}}(f) \lesssim 10^{-5}$ , where the integral is over frequencies  $f \gtrsim 10^{-15} \text{ Hz}$  (164); (iii) the pulsar timing limit  $\Omega_{\text{gw}} \lesssim 10^{-8}$  at  $10^{-9} \text{ Hz} \lesssim f \lesssim 10^{-8} \text{ Hz}$  (165); (iv) the current LIGO/VIRGO upper limit  $\Omega_{\text{gw}} \lesssim 7 \times 10^{-6}$  at  $f \sim 100 \text{ Hz}$  (166); and (v) the limit  $\int d \ln f \Omega_{\text{gw}} \lesssim 10^{-1}$  for  $10^{-17} \text{ Hz} \lesssim f \lesssim 10^{-9} \text{ Hz}$  from VLBI radio astrometry of quasars.

Many new techniques also promise future measurements of these primordial GWs. Firstly, it has been shown that such a GW background would leave a detectable signature in the polarization of the cosmic microwave background (CMB; 167; 168), which will be measured by many current and future observational efforts (169; 170; 171; 172; 173; 174; 175; 176; 177; 178). The planned space-based interferometer eLISA/NGO will also set limits on the primordial stochastic gravitational wave background (SGWB) (179; 180). Similarly, DECIGO is a space-based interferometer mission designed primarily to detect the primordial SGWB (181). Advanced LIGO (182) will soon improve

---

<sup>1</sup>The work presented in this chapter was reproduced from the paper “Astrometric effects of a stochastic gravitational wave background,” L. G. Book and Éanna É. Flanagan, PRD **83**, 024024 (2011). Reproduced with permission, copyright (2011) by the American Physical Society.

the limits set by earth-based interferometers, while the European Einstein Telescope has the potential to significantly enhance the scientific return of these efforts (183). Finally, Seto and Cooray have suggested that measurements of the anisotropy of time variations of redshifts of distant sources could provide constraints of order  $\Omega_{\text{gw}} \lesssim 10^{-5}$  at  $f \sim 10^{-12}$  Hz (184). For more details on GWs, the search for them, and the SGWB, see the review articles (163; 185; 186).

### 4.1.2 High-Precision Astrometry

The possibility of using high-precision astrometry to detect GWs has been considered by many authors (187; 188; 189; 190; 191; 192; 193; 194; 195; 196; 197; 198; 199). There was an early suggestion by Fakir (190) that GW bursts from localized sources could be detectable by the angular deflection  $\Delta\theta$  to light rays that they would produce. Fakir claimed that  $\Delta\theta \propto 1/b$ , where  $b$  is the impact parameter. This claim was shown later to be false, and in fact the deflection scales as  $1/b^3$  (189; 191). Therefore the prospects for using astrometry to detect waves from localized sources are not promising (199).

However, the situation is different for a SGWB, as first discussed by Braginsky et al. (188). For a light ray propagating through a SGWB, one might expect the direction of the ray to undergo a random walk, with the deflection angle growing as the square root of distance. However, this is not the case; the deflection angle is always of order the strain amplitude  $h_{\text{rms}}$  of the GWs, and does not grow with distance <sup>1</sup> (188; 194; 195). Specifically, a SGWB will cause apparent angular deflections which are correlated over the sky and which vary randomly with time, with an rms deflection  $\delta_{\text{rms}}(f)$  per unit logarithmic frequency interval of (see Eq. 4.3 below)

$$\delta_{\text{rms}}(f) \sim h_{\text{rms}}(f) \sim \frac{H_0}{f} \sqrt{\Omega_{\text{gw}}(f)}. \quad (4.1)$$

Suppose now that we monitor the position of  $N$  sources in the sky, with an angular accuracy of  $\Delta\theta$ , over a time  $T$ . For a single source, one could detect an angular velocity (proper motion) of order  $\sim \Delta\theta/T$ , and for  $N$  sources, a correlated angular velocity of order  $\sim \Delta\theta/(T\sqrt{N})$  should be detectable, assuming that the errors on individual sources are uncorrelated. The rms angular velocity from (4.1) is  $\omega_{\text{rms}}(f) \sim f \delta_{\text{rms}}(f) \sim H_0 \sqrt{\Omega_{\text{gw}}(f)}$ , and it follows that one should obtain an upper limit on  $\Omega_{\text{gw}}$  of order (198)

$$\Omega_{\text{gw}}(f) \lesssim \frac{\Delta\theta^2}{NT^2H_0^2}. \quad (4.2)$$

This bound will apply at a frequency of order  $f \sim 1/T$ . It will also apply at lower frequencies (198) since the angular velocity fluctuations are white (equal contributions from all frequency scales),

<sup>1</sup>It is sometimes claimed in the literature that the deflection angle depends only on the GWs near the source and observer. In fact, this is not true, as we discuss in Appendix A. A similar claim about the frequency shift that is the target of pulsar timing searches for GWs is also false in general.

assuming a flat GW spectrum  $\Omega_{\text{gw}} = \text{const.}$  The quantity that will be constrained by observations is roughly this total  $\Omega_{\text{gw}}$ ,  $\int_{f \lesssim T^{-1}} d \ln f \Omega_{\text{gw}}(f)$ .

The advent of microarcsecond astrometry has started to make the prospects for constraining GW backgrounds more interesting. The future astrometry mission GAIA (Global Astrometric Interferometer for Astrophysics) is expected to measure positions, parallaxes and annual proper motions to better than  $20 \mu\text{as}$  for more than 50 million stars brighter than  $V \sim 16$  mag and 500,000 quasars brighter than  $V \sim 20$  mag (187). Similarly the Space Interferometry Mission (SIM) is expected to achieve angular accuracies of order  $10 \mu\text{as}$ . Estimates of the sensitivities of these missions to a SGWB, at the  $\Omega_{\text{gw}} \sim 10^{-3} - 10^{-6}$  level, are given in Refs. (187; 196; 197).

VLBI radio interferometry is another method that can be used to detect the astrometric effects of a SGWB on distant sources. This method detects the same pattern as that discussed in this chapter for visible astrometry, and differs from astrometry using the GAIA satellite in its longer duration (tens of years versus a few years for GAIA), and in the smaller number of sources, on the order of hundreds, that have currently been measured using this method. In the radio, the planned Square Kilometer Array (SKA) is also expected to be able to localize sources to within  $\sim 10 \mu\text{as}$  (200). Jaffe has estimated that with  $10^6$  QSO sources, the SKA could achieve a sensitivity of order  $\Omega_{\text{gw}} \sim 10^{-6}$  (193).

The astrometric signals due to a SGWB expected for a single object are quite small, on the order of  $0.1 \mu\text{as yr}^{-1}$ , much smaller than the typical intrinsic proper motion of a star in our galaxy. We therefore propose to use quasars as our sources, since their extragalactic distances cause their expected intrinsic proper motions to be smaller than those expected from a SGWB (187). The construction of a non-rotating reference frame using quasars in astrometric studies will remove the  $l = 1$  dipole component of the measured quasar proper motions, but will leave intact the  $l = 2$  and higher multipoles which are expected to be excited by GWs.

Using the estimate  $N \sim 10^6$  (GAIA),  $\Delta\theta \sim 10 \mu\text{as}$ ,  $T \sim 1$  yr gives from Eq. (4.2) the estimate

$$\Omega_{\text{gw}} \lesssim 10^{-6}$$

at  $f \lesssim 10^{-8}$  Hz for astrometry. This is an interesting sensitivity level, roughly comparable with that obtainable with pulsar timing (165).

Astrometry has already been applied to obtain upper limits on  $\Omega_{\text{gw}}$  using a number of different observations. First, Gwinn et al. analyzed limits on quasar proper motions obtained from VLBI astrometry, and obtained the upper limit  $\Omega_{\text{gw}} \lesssim 10^{-1}$  for  $10^{-17}$  Hz  $\lesssim f \lesssim 10^{-9}$  Hz (192). This limit was recently updated by Titov, Lambert and Gontier (201). Finally, Linder analyzed observed galaxy correlation functions to obtain the limit  $\Omega_{\text{gw}} \lesssim 10^{-3}$  for  $10^{-16}$  Hz  $\lesssim f \lesssim 10^{-10}$  Hz (202).

All of these analyses used a relatively simple model of the effect of gravitational waves on proper

motions. In this chapter we give a detailed computation of the spectrum of angular fluctuations produced by a stochastic background, including the relative strengths of E- and B-type multipoles for each order  $l$ .

### 4.1.3 Summary of Results

For a source in the direction  $\mathbf{n}$ , the effect of the GW background is to produce an apparent angular deflection  $\delta\mathbf{n}(\mathbf{n}, t)$ . We first find a general formula for the angular deflection of a photon, for an arbitrary GW signal  $h_{ij}$ , emitted by a source that can be at a cosmological distance. This deflection is derived in Sections 4.2 and 4.3 below, and is given by [see Eq. (4.55)]

$$\delta n^i = \frac{1}{2} \left\{ n^j h_{ij}(0) - n^i n^j n^k h_{jk}(0) - \frac{\omega_0}{\zeta_s} (\delta^{ik} - n^i n^k) n^j \cdot \left[ -2 \int_0^{\zeta_s} d\zeta' \int_0^{\zeta'} d\zeta'' h_{jk,0}(\zeta'') + n^l \int_0^{\zeta_s} d\zeta' \int_0^{\zeta'} d\zeta'' (h_{jk,l}(\zeta'') + h_{kl,j}(\zeta'') - h_{jl,k}(\zeta'')) \right] \right\}.$$

Here,  $\mathbf{n}$  is the direction to the source,  $\omega_0$  is the emitted frequency of the photon,  $\zeta$  parameterizes the path of the photon  $\tau(\zeta) = \tau_0 + \omega_0 \zeta$ ,  $x^i(\zeta) = -\zeta \omega_0 n^i$ ,  $h_{ij}(\tau, \mathbf{x})$  is treated as a function of  $\zeta$  through this parameterization of the photon path,  $\zeta_s$  is the value of  $\zeta$  at the emission event of the photon at the source, and the spacetime metric is

$$ds^2 = a(\tau)^2 \{ -d\tau^2 + [\delta_{ij} + h_{ij}(\tau, \mathbf{x})] dx^i dx^j \}.$$

We then specialize to the limit in which the sources are many gravitational wavelengths away and to plane waves propagating in the direction  $\mathbf{p}$  to obtain a simple formula, which generalizes a previous result of Pyne et al (198). We find that the deflection, as a function of time  $\tau$  and direction on the sky  $\mathbf{n}$ , is given by

$$\delta n^{\hat{i}}(\tau, \mathbf{n}) = \frac{n^i + p^i}{2(1 + \mathbf{p} \cdot \mathbf{n})} h_{jk}(0) n_j n_k - \frac{1}{2} h_{ij}(0) n_j,$$

where  $\mathbf{p}$  is the direction of propagation of the GW, and  $h_{ij}(0)$  is the GW field evaluated at the observer,  $h_{ij}(\tau, \mathbf{0})$ .

The main result of this chapter is a computation of the statistical properties of the angular deflection resulting from a SGWB, which is carried out in Sections 4.4 and 4.5. The apparent angular deflection caused by such a GW background is a stationary, zero-mean, Gaussian random process. We compute the fluctuations in  $\delta\mathbf{n}$  by making two different approximations: (i) The GW modes which contribute to the deflection have wavelengths  $\lambda$  which are short compared to the horizon size  $cH_0^{-1}$  today. (ii) The mode wavelengths  $\lambda$  are short compared to the distances to the sources;

this same approximation is made in pulsar timing searches for GWs (203). Since our calculations are only valid for GWs with wavelengths much smaller than the horizon, the contribution from waves with wavelengths comparable to the horizon scale will cause a small deviation from our results (on the order of a few percent for a white GW spectrum).

The total power in angular fluctuations is then

$$\langle \delta \mathbf{n}(\mathbf{n}, t)^2 \rangle = \theta_{\text{rms}}^2 = \frac{1}{4\pi^2} \int d \ln f \left( \frac{H_0}{f} \right)^2 \Omega_{\text{gw}}(f). \quad (4.3)$$

Taking a time derivative gives the spectrum of fluctuations of angular velocity or proper motion:

$$\langle \delta \dot{\mathbf{n}}(\mathbf{n}, t)^2 \rangle = \int d \ln f H_0^2 \Omega_{\text{gw}}(f),$$

which gives a rms angular velocity  $\omega_{\text{rms}}(f)$  of order

$$\omega_{\text{rms}}(f) \sim H_0 \sqrt{\Omega_{\text{gw}}} \sim 10^{-2} \mu\text{as yr}^{-1} \left( \frac{\Omega_{\text{gw}}}{10^{-6}} \right)^{1/2}.$$

This is the signal that we hope to detect.

We now discuss how the angular fluctuations are distributed on different angular scales, or equivalently how the power is distributed in the spherical harmonic index  $l$ . The total angular fluctuations can be written as

$$\langle \delta \mathbf{n}(\mathbf{n}, t)^2 \rangle = \int d \ln f \sum_{l=2}^{\infty} [\theta_{\text{rms},l}^E(f)^2 + \theta_{\text{rms},l}^B(f)^2]. \quad (4.4)$$

Here  $\theta_{\text{rms},l}^E(f)^2$  is the total electric-type power in angular fluctuations per unit logarithmic frequency in multipole sector  $l$ , and  $\theta_{\text{rms},l}^B(f)^2$  is the corresponding magnetic-type power. These quantities can be written as

$$\theta_{\text{rms},l}^Q(f)^2 = \theta_{\text{rms}}^2 g_Q \sigma(f) \alpha_l^{QQ}, \quad (4.5)$$

where  $Q = E$  or  $B$ . The various factors in this formula are as follows. The factors  $g_E$  and  $g_B$  are the fractions of the total power carried by E-modes and B-modes, respectively, and satisfy  $g_E + g_B = 1$ . Their values are  $g_E = g_B = 1/2$ , implying that electric and magnetic type fluctuations have equal power. The function  $\sigma(f)$  describes how the power is distributed in frequency, and is the same for all multipoles, both electric and magnetic. It is normalized so that  $\int d \ln f \sigma(f) = 1$ , and is given explicitly by [cf. Eq. (4.3) above]

$$\sigma(f) = \frac{f^{-2} \Omega_{\text{gw}}(f)}{\int d \ln f' f'^{-2} \Omega_{\text{gw}}(f')}. \quad (4.6)$$

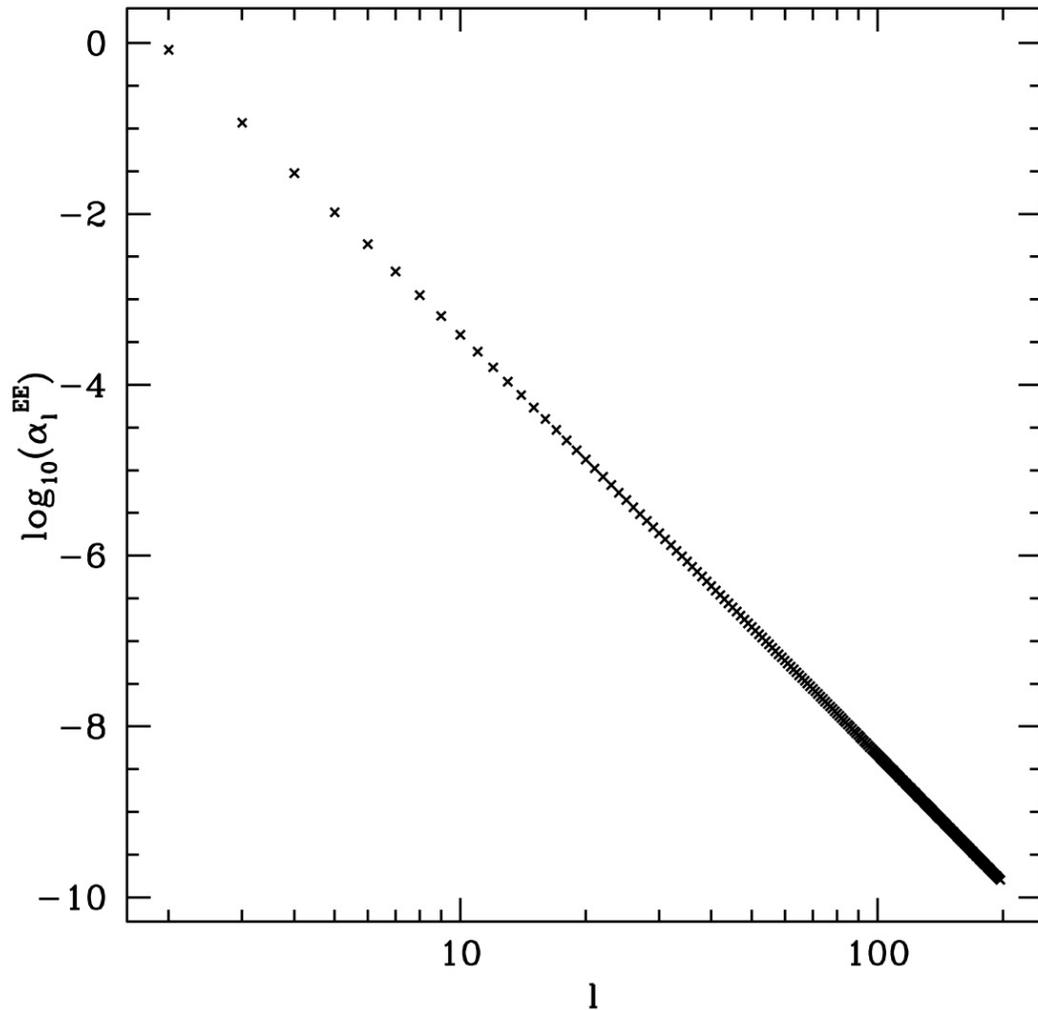


Figure 4.1: Here we plot the coefficients  $\alpha_l^{EE}$  as defined in Eq. (4.87) vs. multipole  $l$ .

Finally, the angular spectra  $\alpha_l^{EE}$  and  $\alpha_l^{BB}$  describe how the power is distributed in different multipoles, starting with the quadrupole at  $l = 2$ , and are normalized so that

$$\sum_{l=2}^{\infty} \alpha_l^{QQ} = 1 \quad (4.7)$$

for  $Q = E$  and  $Q = B$ . We show that  $\alpha_l^{EE} = \alpha_l^{BB}$ , and this spectrum is plotted in Figure 4.1 and tabulated in Table 4.1. These coefficients are well fit by the power law  $\alpha_l^{EE} = 32.34 l^{-4.921}$ . We note that the result for the quadrupole,  $\alpha_2^{EE} = 5/6$ , has previously been derived using a different method in Ref. (198).

Table 4.1: First 10 nonzero multipole coefficients  $\alpha_l^{EE}$  as defined in Eq. (4.87) and plotted in Figure 4.1

$l$	$\alpha_l^{EE}$
2	0.833333
3	0.116667
4	0.03
5	0.0104762
6	0.00442177
7	0.00212585
8	0.00112434
9	0.000639731
10	0.000385675
11	0.000243696

## 4.2 Calculation of Astrometric Deflection in a Minkowski Background Spacetime

### 4.2.1 Setting the Stage—Minkowski Calculation

We will first calculate the angular deflection due to a small GW perturbation on a flat background metric,

$$ds^2 \equiv g_{\mu\nu} dx^\mu dx^\nu = -dt^2 + (\delta_{ij} + h_{ij}) dx^i dx^j. \quad (4.8)$$

We are considering the effect of these GWs on a photon traveling from a source to an observer, with an unperturbed worldline  $x_0^\alpha(\lambda) = \omega_0(\lambda, -\lambda\mathbf{n}) + (t_0, 0, 0, 0)$ , where  $-\mathbf{n}$  is the direction of the photon's travel,  $\omega_0$  is its unperturbed frequency, and the photon is observed at the origin at time  $t_0$ . The photon's unperturbed 4-momentum is given by  $k_0^\alpha = \omega_0(1, -\mathbf{n})$ .

To calculate the geodesics that the photon, source and observer follow, we need the connection coefficients in this metric. There are three nonzero connection coefficients:

$$\Gamma_{0i}^k = \frac{1}{2} h_{ki,0}, \quad \Gamma_{ij}^0 = \frac{1}{2} h_{ij,0}, \quad \Gamma_{ij}^k = \frac{1}{2} [h_{ki,j} + h_{kj,i} - h_{ij,k}]. \quad (4.9)$$

First, using the geodesic equation

$$\frac{d^2 x^\alpha}{d\tau^2} = -\Gamma_{\beta\gamma}^\alpha u^\beta u^\gamma, \quad (4.10)$$

it is straightforward to verify that the paths of stationary observers in these coordinates are geodesics. Therefore we can assume that both the source and observer are stationary in these coordinates, with

$$\begin{aligned}
x_{obs}^i(t) &= 0 \\
x_s^i(t) &= x_s^i = \text{constant}.
\end{aligned}$$

The affine parameter of the source is therefore

$$\lambda_s = -\frac{|\mathbf{x}_s|}{\omega_0}.$$

### 4.2.2 Photon Geodesic

Next, we solve the geodesic equation (4.10) for the path of a photon traveling from the source to the observer in the perturbed metric. We write this path as the sum of contributions of zeroth and first order in  $h$ ,

$$x^\alpha(\lambda) = x_0^\alpha(\lambda) + x_1^\alpha(\lambda). \quad (4.11)$$

Similarly, the photon 4-momentum is

$$k^\alpha(\lambda) = k_0^\alpha(\lambda) + k_1^\alpha(\lambda). \quad (4.12)$$

We note that the connection coefficients are all first order in  $h$ , so keeping only first-order terms, we will use only the unperturbed photon 4-momentum in the geodesic equation, yielding

$$\frac{d^2 x_1^0}{d\lambda^2} = -\frac{\omega_0^2}{2} n^i n^j h_{ij,0}, \quad (4.13)$$

$$\frac{d^2 x_1^k}{d\lambda^2} = -\frac{\omega_0^2}{2} [-2n^i h_{ki,0} + n^i n^j (h_{ki,j} + h_{kj,i} - h_{ij,k})]. \quad (4.14)$$

We now integrate the geodesic equation to obtain the perturbed photon 4-momentum and trajectory. The right-hand sides are evaluated along the photon's unperturbed path from  $\lambda = 0$  at the present time back to  $\lambda$ , since they are already first order in  $h$ . We define

$$\begin{aligned}
\mathcal{I}_{ij}(\lambda) &= \int_0^\lambda d\lambda' h_{ij,0}(\lambda'), & \mathcal{J}_{ijk}(\lambda) &= \int_0^\lambda d\lambda' h_{ij,k}(\lambda'), \\
\mathcal{K}_{ij}(\lambda) &= \int_0^\lambda d\lambda' \int_0^{\lambda'} d\lambda'' h_{ij,0}(\lambda''), & \mathcal{L}_{ijk}(\lambda) &= \int_0^\lambda d\lambda' \int_0^{\lambda'} d\lambda'' h_{ij,k}(\lambda''),
\end{aligned} \quad (4.15)$$

where  $h_{ij}(\lambda)$  means  $h_{ij}(t_0 + \omega_0\lambda, -\omega_0\lambda\mathbf{n})$ . We find

$$\begin{aligned}
k_1^0(\lambda) &= -\frac{\omega_0^2}{2} n^i n^j \mathcal{I}_{ij}(\lambda) + I_0, & k_1^j(\lambda) &= -\frac{\omega_0^2}{2} n^i R_{ij} + J_0^j, \\
x_1^0(\lambda) &= -\frac{\omega_0^2}{2} n^i n^j \mathcal{K}_{ij}(\lambda) + I_0 \lambda + K_0, & x_1^j(\lambda) &= -\frac{\omega_0^2}{2} n^i S_{ij} + J_0^j \lambda + L_0^j,
\end{aligned} \tag{4.16}$$

where  $I_0$ ,  $J_0^j$ ,  $K_0$  and  $L_0^j$  are constants of integration, and we have defined the quantities

$$R_{ij}(\lambda) \equiv [-2\mathcal{I}_{ij}(\lambda) + n^k (\mathcal{J}_{ijk}(\lambda) + \mathcal{J}_{jki}(\lambda) - \mathcal{J}_{ikj}(\lambda))], \tag{4.17}$$

$$S_{ij}(\lambda) \equiv [-2\mathcal{K}_{ij}(\lambda) + n^k (\mathcal{L}_{ijk}(\lambda) + \mathcal{L}_{jki}(\lambda) - \mathcal{L}_{ikj}(\lambda))]. \tag{4.18}$$

### 4.2.3 Boundary Conditions

We determine the eight integration constants  $I_0$ ,  $J_0^j$ ,  $K_0$  and  $L_0^j$  using the boundary conditions of the problem, namely that the photon path passes through the detection event  $x_{obs}^\mu = (t_0, 0, 0, 0)$ , that it is null, that the photon is emitted with the unperturbed frequency  $\omega_0$ , and that the photon path intersects the path of the source at some earlier time.

#### 1. Photon path must pass through detection event

First, the perturbed photon trajectory must pass through the detection event  $t = t_0$ ,  $x^i = 0$ . Therefore,

$$x^\mu(0) = x_0^\mu(0) + x_1^\mu(0) = (t_0, 0, 0, 0),$$

giving

$$K_0 = \frac{\omega_0^2}{2} n^i n^j \mathcal{K}_{ij}(0) = 0, \quad L_0^j = \frac{\omega_0^2}{2} n^i S_{ij}(0) = 0, \tag{4.19}$$

where we have used the fact that by definition  $\mathcal{K}_{ij}(0) = S_{ij}(0) = 0$ .

#### 2. Photon geodesic is null

The geodesic of the photon must be null, which gives one more constraint:  $g_{\mu\nu} k^\mu k^\nu = 0$ . This is already true to zeroth order. To first-order we get:

$$0 = h_{\mu\nu} k_0^\mu k_0^\nu + 2\eta_{\mu\nu} k_1^\mu k_0^\nu,$$

where  $k_0^\alpha = \omega_0(1, -\mathbf{n})$ . Inserting the expression for the perturbed 4-momentum  $k_1^\alpha$  given by Eqs. (4.15), (4.16) and (4.17), and simplifying using

$$\frac{d}{d\lambda} h_{ij} = \omega_0 h_{ij,0} - \omega_0 n_k h_{ij,k} \quad (4.20)$$

shows that all of the terms involving  $\lambda$  cancel out, as they must, leaving the condition

$$I_0 + n_i J_0^i = \frac{1}{2} \omega_0 n^i n^j h_{ij}(0). \quad (4.21)$$

### 3. Photon is emitted with frequency $\omega_0$

The photon is emitted at the source with the unperturbed frequency  $\omega_0 = -g_{\mu\nu} k^\mu u_s^\nu$ . The 4-velocity of the source is  $u_s^\mu = (1, 0, 0, 0)$  as it has constant spatial coordinate position, so the constraint becomes  $-g_{\mu 0} k^\mu = \omega_0$ . The source emits the photon at  $\lambda = \lambda_s$ , so from Eq. (4.16) this yields

$$I_0 = \frac{\omega_0^2}{2} n^i n^j \mathcal{I}_{ij}(\lambda_s) \quad (4.22)$$

### 4. Perturbed photon path must hit source worldline somewhere

The constraint that the perturbed photon trajectory must hit the source worldline somewhere can be written as

$$x^j(\tilde{\lambda}_s) = x_s^j = x_0^j(\tilde{\lambda}_s) + x_1^j(\tilde{\lambda}_s) \quad (4.23)$$

for some  $\tilde{\lambda}_s$ . To zeroth order we have  $\tilde{\lambda}_s = \lambda_s$ , but there will be a first-order correction. Inserting the expression (4.16) for the perturbation of the geodesic gives

$$x_s^j = -\omega_0 \tilde{\lambda}_s n^j - \frac{\omega_0^2}{2} n^i S_{ij}(\tilde{\lambda}_s) + \tilde{\lambda}_s J_0^j. \quad (4.24)$$

Projecting this equation perpendicular to  $\mathbf{n}$  gives a formula for the perpendicular component of  $J_0^i$ :

$$J_0^i{}_{\perp} = \frac{\omega_0^2}{2\lambda_s} (\delta^{ik} - n^i n^k) n^j S_{jk}(\lambda_s). \quad (4.25)$$

Here on the right-hand side we have replaced  $\tilde{\lambda}_s$  with  $\lambda_s$ , which is valid to linear order. Adding to this our earlier result for the component of  $J_0^i$  parallel to  $\mathbf{n}$  in Eqs. (4.21) and (4.22) gives

$$J_0^i = \frac{\omega_0^2}{2\lambda_s} n^j S_{jk}(\lambda_s) (\delta^{ik} - n^i n^k) - \frac{\omega_0^2}{2} n^i n^j n^k \mathcal{I}_{jk}(\lambda_s) + \frac{1}{2} \omega_0 n^i n^j n^k h_{jk}(0). \quad (4.26)$$

#### 4.2.4 Perturbation to Observed Frequency

We calculate the observed photon frequency  $\omega_{obs} = -g_{\mu\nu}k^\mu u_{obs}^\nu$ , where  $u_{obs}^\nu = (1, 0, 0, 0)$ , and check our result against standard formulae for the frequency shift, used in pulsar timing searches for GWs (204). The observed frequency is, from Eqs. (4.16) and (4.22),

$$\omega_{obs} = k^0(0) = \omega_0 + I_0 = \omega_0 + \frac{\omega_0^2}{2} n^i n^j \mathcal{I}_{ij}(\lambda_s). \quad (4.27)$$

Using the definition (4.15), the perturbed redshift is therefore

$$z \equiv \frac{\omega_0 - \omega_{obs}}{\omega_0} = -\frac{\omega_0}{2} n^i n^j \int_0^{\lambda_s} d\lambda' h_{ij,0}(\lambda'). \quad (4.28)$$

For a plane wave traveling in the direction of the unit vector  $\mathbf{p}$ , we have

$$h_{ij} = h_{ij}(t - \mathbf{p} \cdot \mathbf{x}) = h_{ij}[\omega_0 \lambda (1 + \mathbf{p} \cdot \mathbf{n})],$$

giving

$$h_{ij,0} \equiv \frac{\partial}{\partial t} h_{ij} = \frac{1}{\omega_0(1 + \gamma)} \frac{\partial}{\partial \lambda} h_{ij},$$

where  $\gamma = \mathbf{p} \cdot \mathbf{n}$ . This gives for the redshift

$$z = -\frac{1}{2(1 + \gamma)} n^i n^j [h_{ij}(\lambda_s) - h_{ij}(0)], \quad (4.29)$$

which agrees with (204) up to a sign, which is an error in their calculation (205).

#### 4.2.5 Local Proper Reference Frame of Observer

We must also account for the changes induced in the basis vectors of the observer's local proper reference frame due to the presence of the GW. We introduce a set of orthonormal basis vectors  $\vec{e}_{\hat{\alpha}}$  which are parallel transported along the observer's worldline, with  $\vec{e}_{\hat{0}} = \vec{u}$ . The parallel transport equation for the spatial vectors gives

$$u^\alpha e_{j;\alpha}^\beta = u^\alpha \left[ \partial_\alpha e_j^\beta + \Gamma_{\alpha\gamma}^\beta e_j^\gamma \right] = 0. \quad (4.30)$$

We separate the basis vectors into two pieces,  $e_j^i = \delta_j^i + \delta e_j^i$ , where we assume that the unperturbed basis vectors are aligned with the coordinate basis directions.

Using  $\vec{u} = \partial_t$ , and the connection coefficients (4.9) of the metric, Eq. (4.30) gives us an expression for the perturbation to the basis tetrad:

$$\delta e_j^i(t) = -\frac{1}{2}h_{i\hat{j}}(t, \mathbf{0}) + \omega_{i\hat{j}},$$

where  $\omega_{\hat{j}}^i$  is a matrix of constants. Now, we observe that  $e_{\hat{j}}$  is an orthonormal set of three-vectors, which gives us six constraints on the constants  $\omega_{\hat{j}}^i$ :

$$(\eta_{mn} + h_{mn}) \left( \delta_j^m + \delta e_{\hat{j}}^m \right) \left( \delta_k^n + \delta e_{\hat{k}}^n \right) = \delta_{\hat{j}\hat{k}}.$$

This is identically correct to zeroth order; to first order we get  $\delta e_{j\hat{k}} + \delta e_{k\hat{j}} + h_{jk} = 0$ , or, inserting our equation for  $\delta e$ , and assuming that  $h_{ij} = h_{ji}$ , we find  $\omega_{ij} = -\omega_{ji}$ , i.e., that the constants  $\omega_{ij}$  are antisymmetric in their indices. These constants parameterize an arbitrary infinitesimal time-independent rotation. Evaluating now at the detection event gives

$$\delta e_j^i = -\frac{1}{2}h_{i\hat{j}}(0) + \omega_{i\hat{j}}. \quad (4.31)$$

For the remainder of this chapter we will set to zero the term  $\omega_{i\hat{j}}$ , since it corresponds to a time-independent, unobservable angular deflection. The deflections caused by GWs will be observable because of their time dependence.

## 4.2.6 Observed Angular Deflection

We can express the four-momentum of the incoming photon in the above reference frame as

$$k^\alpha(0) = \omega_{obs} u^\alpha - \omega_{obs} n^{\hat{j}} e_j^\alpha, \quad (4.32)$$

where  $\delta_{\hat{j}\hat{k}} n^{\hat{j}} n^{\hat{k}} = 1$ ,  $u^\alpha$  is the observer's 4-velocity, and  $\omega_{obs}$  is given by equation 4.27. Note that we evaluate all quantities at the detection event  $t = t_0$ ,  $\mathbf{x} = 0$ . Plugging in our results for the perturbed 4-momentum and the observed frequency, we obtain an equation for the observed direction to the source  $n^{\hat{j}}$

$$\begin{aligned} k^i(0) &= -\omega_0 n^i + \frac{\omega_0^2}{2\lambda_s} n^j S_{jk}(\lambda_s) (\delta^{ik} - n^i n^k) - \frac{\omega_0^2}{2} n^i n^j n^k \mathcal{I}_{jk}(\lambda_s) + \frac{1}{2} \omega_0 n^i n^j n^k h_{jk}(0) \\ &= -\left( \omega_0 + \frac{\omega_0^2}{2} n^k n^l \mathcal{I}_{kl}(\lambda_s) \right) n^{\hat{j}} \left( \delta_j^i - \frac{1}{2} h_j^i(0) \right). \end{aligned} \quad (4.33)$$

We decompose the direction to the source into zeroth- and first-order pieces as  $n^{\hat{j}} = n_0^{\hat{j}} + \delta n^{\hat{j}}$ . The zeroth-order terms in Eq. (4.33) give us  $n_0^{\hat{j}} = n^{\hat{j}}$ . Plugging this into the first-order terms and simplifying, we find the perturbation to the source direction

$$\delta n^{\hat{i}} = \frac{1}{2} \left\{ n^j h_{ij}(0) - \frac{\omega_0}{\lambda_s} n^j S_{jk}(\lambda_s) (\delta^{ik} - n^i n^k) - n^i n^j n^k h_{jk}(0) \right\}.$$

Inserting our definition of  $S_{jk}$ , we obtain the solution to the source direction perturbation in Minkowski space

$$\begin{aligned} \delta n^{\hat{i}} = \frac{1}{2} \left\{ n^j h_{ij}(0) - n^i n^j n^k h_{jk}(0) - \frac{\omega_0}{\lambda_s} (\delta^{ik} - n^i n^k) n^j \right. \\ \left. \times \left[ -2 \int_0^{\lambda_s} d\lambda' \int_0^{\lambda'} d\lambda'' h_{jk,0}(\lambda'') + n^l \int_0^{\lambda_s} d\lambda' \int_0^{\lambda'} d\lambda'' (h_{jk,l}(\lambda'') + h_{kl,j}(\lambda'') - h_{jl,k}(\lambda'')) \right] \right\}. \end{aligned} \quad (4.34)$$

As a check of the calculation, we see that  $\delta n^{\hat{i}}$  is orthogonal to  $n^i$ , so that  $n^i + \delta n^{\hat{i}}$  is a unit vector, as expected.

We now specialize to the case of a plane wave propagating in the direction of the unit vector  $\mathbf{p}$ ,

$$h_{ij}(t, \mathbf{x}) = h_{ij}(t - \mathbf{p} \cdot \mathbf{x}).$$

Using the identity (4.20) we can reduce the double integrals in Eq. (4.34) to single integrals, obtaining

$$\delta n^{\hat{i}} = (\delta^{ik} - n^i n^k) n^j \left\{ -\frac{1}{2} h_{jk}(0) + \frac{p_k n_l}{2(1 + \mathbf{p} \cdot \mathbf{n})} h_{jl}(0) + \frac{1}{\lambda_s} \int_0^{\lambda_s} d\lambda \left[ h_{jk}(\lambda) - \frac{p_k n_l}{2(1 + \mathbf{p} \cdot \mathbf{n})} h_{jl}(\lambda) \right] \right\}. \quad (4.35)$$

Evaluating this explicitly for the plane wave

$$h_{ij}(t, \mathbf{x}) = \text{Re} \left[ \mathcal{H}_{ij} e^{-i\Omega(t - \mathbf{p} \cdot \mathbf{x})} \right]$$

gives

$$\begin{aligned} \delta n^{\hat{i}} = \text{Re} \left[ \left( \left\{ 1 + \frac{i(2 + \mathbf{p} \cdot \mathbf{n})}{\omega_0 \lambda_s \Omega (1 + \mathbf{p} \cdot \mathbf{n})} \left[ 1 - e^{-i\Omega \omega_0 (1 + \mathbf{p} \cdot \mathbf{n}) \lambda_s} \right] \right\} n^i \right. \right. \\ \left. \left. + \left\{ 1 + \frac{i}{\omega_0 \lambda_s \Omega (1 + \mathbf{p} \cdot \mathbf{n})} \left[ 1 - e^{-i\Omega \omega_0 (1 + \mathbf{p} \cdot \mathbf{n}) \lambda_s} \right] \right\} p^i \right) \frac{n^j n^k \mathcal{H}_{jk} e^{-i\Omega t_0}}{2(1 + \mathbf{p} \cdot \mathbf{n})} \right. \\ \left. - \left\{ \frac{1}{2} + \frac{i}{\omega_0 \lambda_s \Omega (1 + \mathbf{p} \cdot \mathbf{n})} \left[ 1 - e^{-i\Omega \omega_0 (1 + \mathbf{p} \cdot \mathbf{n}) \lambda_s} \right] \right\} n^j \mathcal{H}_j^i e^{-i\Omega t_0} \right]. \end{aligned} \quad (4.36)$$

If we define the observed angles  $(\theta, \phi)$  by  $n^{\hat{i}} = (\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta)$ , then the observed angular deflections are

$$\delta\theta = e_{\hat{\theta}}^i \delta n^i, \quad \delta\phi = \frac{e_{\hat{\phi}}^i \delta n^i}{\sin\theta}, \quad (4.37)$$

where  $e_{\hat{\theta}}^i = (\cos\theta \cos\phi, \cos\theta \sin\phi, -\sin\theta)$  and  $e_{\hat{\phi}}^i = (-\sin\phi, \cos\phi, 0)$ .

As another check of our calculation, we now compare our result with the coordinate (gauge-dependent) angular deflection computed by Yoo et al. (206). Starting from our Eq. (4.34), we disregard the first term, which arises from the change from the coordinate basis to the parallel transported orthonormal basis. The remaining terms in Eq. (4.34) give the coordinate angular deflection  $\delta n^i$ . Simplifying using the identity (4.20) and the identity  $\int_0^x dx' \int_0^{x'} dx'' f(x'') = \int_0^x dx' (x-x')f(x')$  gives

$$\delta n^i = -\frac{1}{2}n^i n^j n^k h_{jk}(0) + (\delta^{ij} - n^i n^j) \int_0^{\lambda_s} d\lambda \left\{ \frac{h^{jk}(\lambda) - h^{jk}(0)}{\lambda_s} n_k + \frac{\omega_0}{2} \left( \frac{\lambda_s - \lambda}{\lambda_s} \right) \partial_j (n^k n^l h_{kl}) \right\}. \quad (4.38)$$

When combined with Eqs. (4.37), this agrees with Eqs. (13) and (14) of (206), specialized to only tensor perturbations, up to an overall sign. The sign flip is due to the fact that Ref. (206) uses a convention for the sign of angular deflection, explained after their Eq. (16), which is opposite to ours.

#### 4.2.7 The Distant Source Limit

We now specialize to the limit where the distance  $\omega_0 |\lambda_s|$  to the source is large compared to the wavelength  $\sim c\Omega^{-1}$  of the GWs. As discussed in the Introduction, astrometry is potentially sensitive to waves with a broad range of frequencies, extending from the inverse of the observation time (a few years) down to the Hubble frequency. Therefore this assumption is a nontrivial limitation on the domain of validity of our analysis. However, for sources at cosmological distances (the most interesting case), this assumption is not a significant limitation.

In this limit, we can neglect the second term in each of the three small square brackets in Eq. (4.36), giving

$$\delta n^i(t, \mathbf{n}) = \text{Re} \left[ (n^i + p^i) \frac{\mathcal{H}_{jk} n_j n_k e^{-i\Omega t}}{2(1 + \mathbf{p} \cdot \mathbf{n})} - \frac{1}{2} \mathcal{H}_{ij} n_j e^{-i\Omega t} \right], \quad (4.39)$$

where we have written  $t$  for  $t_0$ . This result agrees with and generalizes a calculation of Pyne et al. (198). We note that this same approximation is used in pulsar timing searches for GWs (204). In that context the approximation is essentially always valid, since pulsar distances are large compared to a few light years, and the properties of pulsar frequency noise imply that that pulsar timing is only sensitive to GWs with periods of order the observation time, and not much lower frequencies, unlike the case for astrometry.

### 4.3 Generalization to Cosmological Spacetimes

Of course, we do not live in Minkowski space. The apparent homogeneity and isotropy of the Universe imply that our Universe has an FRW geometry, with line element:

$$ds^2 = g_{\alpha\beta} dx^\alpha dx^\beta = a(\tau)^2 \{-d\tau^2 + [\delta_{ij} + h_{ij}(\tau, \mathbf{x})] dx^i dx^j\}, \quad (4.40)$$

where  $\tau$  is conformal time, and we specialize to the transverse traceless gauge in which  $\delta^{ij} h_{ij} = \delta^{ij} \partial_i h_{jk} = 0$ . To translate our calculation in Minkowski spacetime to this new metric, we define an unphysical, conformally related metric  $\bar{g}_{\alpha\beta} = a(\tau)^{-2} g_{\alpha\beta}$  given by

$$\bar{g}_{\alpha\beta} dx^\alpha dx^\beta = -d\tau^2 + [\delta_{ij} + h_{ij}(\tau, \mathbf{x})] dx^i dx^j, \quad (4.41)$$

which has an associated unphysical derivative operator  $\bar{\nabla}_\alpha$ .

#### 4.3.1 Stationary Observers are Freely Falling

As before, it is straightforward to check that observers who are stationary in the coordinates (4.40) are freely falling. Therefore we assume as before that the observer and source are stationary:

$$x_{obs}^i(t) = 0, \quad x_s^i(t) = x_s^i.$$

#### 4.3.2 Null Geodesic in the Conformal Metric

Let us consider a photon traveling from a distant source to us, which follows a null geodesic in the physical metric  $g_{\alpha\beta}$ . Its path is also a null geodesic of the conformally related metric  $\bar{g}_{\alpha\beta}$ , though it is not affinely parameterized in this metric (207). Specifically, the physical 4-momentum of the photon  $k^\mu$  must satisfy the geodesic equation  $k^\mu \nabla_\mu k_\nu = 0$ . If we define a conformally related, unphysical 4-momentum  $\bar{k}_\mu = k_\mu$ , whose contravariant components are then related to those of the physical 4-momentum by

$$\bar{k}^\mu = \bar{g}^{\mu\nu} \bar{k}_\nu = a(\tau)^2 g^{\mu\nu} \bar{k}_\nu = a(\tau)^2 g^{\mu\nu} k_\nu = a(\tau)^2 k^\mu, \quad (4.42)$$

then we find that

$$\bar{k}^\mu \bar{\nabla}_\mu \bar{k}_\nu = a(\tau)^2 k^\mu \bar{\nabla}_\mu k_\nu. \quad (4.43)$$

From (207) we know that for any vector  $v^\alpha$ , and conformally related derivatives  $\nabla_\alpha$  and  $\bar{\nabla}_\alpha$ , we have  $\nabla_\alpha v_\beta = \bar{\nabla}_\alpha v_\beta - C_{\alpha\gamma}^\gamma v_\gamma$ , where  $C_{\alpha\gamma}^\gamma = 2\delta_{(\alpha}^\gamma \nabla_{\beta)} \ln a - g_{\alpha\beta} g^{\gamma\delta} \nabla_\delta \ln a$ . Thus, we find

$$\begin{aligned}
\bar{k}^\mu \bar{\nabla}_\mu \bar{k}_\nu &= a(\tau)^2 k^\mu \nabla_\mu k_\nu + a(\tau)^2 k^\mu k_\rho \left( 2\delta_{(\mu}^\rho \nabla_{\nu)} \ln a - g_{\mu\nu} g^{\rho\sigma} \nabla_\sigma \ln a \right) \\
&= a(\tau)^2 k^\mu \nabla_\mu k_\nu + a(\tau)^2 (k^\rho k_\rho \nabla_\nu \ln a + k^\mu k_\nu \nabla_\mu \ln a - k_\nu k^\sigma \nabla_\sigma \ln a) \\
&= a(\tau)^2 k^\mu \nabla_\mu k_\nu,
\end{aligned} \tag{4.44}$$

where to get the last line we have used that the geodesic is null. Therefore, if  $k^\mu$  is a null geodesic of the physical metric  $g_{\mu\nu}$ , then  $\bar{k}^\mu$  is a null geodesic of the conformally related metric  $\bar{g}_{\mu\nu}$ . If  $\lambda$  is an affine parameter of the geodesic, it will not be an affine parameter of the geodesic in the unphysical metric. The affine parameter  $\bar{\lambda}$  in the unphysical metric is related to  $\lambda$  by

$$\frac{d\bar{\lambda}}{d\lambda} = \frac{1}{a(\tau(\lambda))^2}. \tag{4.45}$$

### 4.3.3 Parallel Transport of Basis Vectors in FRW Background Spacetime

We next investigate the parallel transport of the observer's basis tetrad in a FRW background spacetime. From the form (4.40) of the metric, we anticipate that the basis vectors must scale as  $a^{-1}$  to remain normalized. Thus, we will define the basis vectors and their perturbations as

$$e_j^i = \frac{1}{a} \left( \delta_j^i + \delta e_j^i \right). \tag{4.46}$$

The relevant connection coefficients are

$$\Gamma_{0k}^i = \frac{\dot{a}}{a} \delta_k^i + \frac{1}{2} \delta^{im} h_{mk,0}. \tag{4.47}$$

The parallel transport equation (4.30) for the spatial basis vectors gives us

$$\partial_0 e_j^i + \Gamma_{0k}^i e_j^k = 0. \tag{4.48}$$

Plugging in the connection coefficients (4.47) and the basis vector expansion (4.46), we get

$$\partial_0 \delta e_j^i + \frac{1}{2} \delta^{im} h_{mj} = 0, \tag{4.49}$$

the same equation as before. The solution, as before, will be

$$\delta e_j^i(t) = -\frac{1}{2} h_j^i(t). \tag{4.50}$$

### 4.3.4 Generalization of Angular Deflection Computation

We parametrize the photon path in the background spacetime by

$$\tau(\zeta) = \tau_0 + \omega_0 \zeta, \quad x^i(\zeta) = -\zeta \omega_0 n^i, \quad (4.51)$$

where  $\zeta$  is an affine parameter of the unphysical metric (4.41) (denoted  $\bar{\lambda}$  above). From the decomposition (4.32), the observed source direction is

$$n^{\hat{j}} = \frac{g_{\alpha\beta} k^\alpha e_j^\beta}{g_{\alpha\beta} k^\alpha u^\beta}. \quad (4.52)$$

We rewrite all the quantities in this expression in terms of their conformally transformed versions

$$\bar{g}_{\alpha\beta} = a^{-2} g_{\alpha\beta}, \quad \bar{k}^\alpha = a^2 k^\alpha, \quad \bar{u}^\alpha = a u^\alpha, \quad \bar{e}_j^\alpha = a e_j^\alpha, \quad (4.53)$$

which are the quantities that are used in the Minkowski spacetime calculation of Section 4.2. This gives

$$n^{\hat{j}} = \frac{\bar{g}_{\alpha\beta} \bar{k}^\alpha \bar{e}_j^\beta}{\bar{g}_{\alpha\beta} \bar{k}^\alpha \bar{u}^\beta}, \quad (4.54)$$

the same expression as in Minkowski spacetime. Therefore, the final result is the same expression (4.34) as before, except that it is written in terms of the non-affine parameter  $\zeta$ :

$$\begin{aligned} \delta n^{\hat{i}} = & \frac{1}{2} \left\{ n^j h_{ij}(0) - n^i n^j n^k h_{jk}(0) - \frac{\omega_0}{\zeta_s} (\delta^{ik} - n^i n^k) n^j \right. \\ & \left. \times \left[ -2 \int_0^{\zeta_s} d\zeta' \int_0^{\zeta'} d\zeta'' h_{jk,0}(\zeta'') + n^l \int_0^{\zeta_s} d\zeta' \int_0^{\zeta'} d\zeta'' (h_{jk,l}(\zeta'') + h_{kl,j}(\zeta'') - h_{jl,k}(\zeta'')) \right] \right\}. \end{aligned} \quad (4.55)$$

### 4.3.5 The Distant Source Limit

We now specialize again to the limit where the distance to the source is large compared to the wavelength  $\sim c \Omega^{-1}$  of the GWs. We also assume that the wavelength  $c \Omega^{-1}$  is small compared to the horizon scale, but we allow the sources to be at cosmological distances.

Starting from Eq. (4.55) and paralleling the derivation of Eq. (4.38) we obtain

$$\delta n^{\hat{i}}(\tau_0, \mathbf{n}) = \frac{1}{2} s_{ik} n_j h_{jk}(0) + \frac{s_{ik} n_j}{\zeta_s} \int_0^{\zeta_s} d\zeta [h_{jk}(\zeta) - h_{jk}(0)] + \frac{\omega_0 s_{ik}}{2} \int_0^{\zeta_s} d\zeta \left( \frac{\zeta_s - \zeta}{\zeta_s} n_j n_l h_{jl,k}(\zeta) \right), \quad (4.56)$$

where  $s_{ik} = \delta_{ik} - n_i n_k$ . Now the wave equation satisfied by the metric perturbation is

$$\left[ \partial_\tau^2 + 2\frac{a',\tau}{a} \partial_\tau - \nabla^2 \right] h_{ij}(\tau, \mathbf{x}) = 0,$$

and plane wave solutions are of the form

$$h_{ij}(\tau, \mathbf{x}) = \text{Re} \left\{ \mathcal{H}_{ij} e^{i\Omega \mathbf{p} \cdot \mathbf{x}} q_\Omega(\tau) \right\},$$

where the mode function  $q_\Omega$  satisfies

$$q_\Omega'' + 2\frac{a'}{a} q_\Omega' + \Omega^2 q_\Omega = 0. \quad (4.57)$$

We now evaluate the angular deflection (4.56) for such a plane wave, in the limit where  $\varepsilon \equiv a'/(\Omega a) \ll 1$ , i.e., the limit where the wavelength  $\sim a/\Omega$  of the GW is much smaller than the horizon scale  $\sim a^2/a'$ . In the second term in (4.56), the term  $h_{jk}(\zeta)$  is rapidly oscillating, and so its integral can be neglected compared to the integral of  $h_{jk}(0)$ ; corrections will be suppressed by powers of  $\varepsilon$ . In the third term in (4.56), the integrand is rapidly oscillating, and so the integral will be dominated by contributions near the endpoints, up to  $\mathcal{O}(\varepsilon)$  corrections. However the integrand vanishes at  $\zeta = \zeta_s$ , and thus the integral is dominated by the region near  $\zeta = 0$ . In that region we can use the leading order WKB approximation to the mode function solution of (4.57),

$$q_\Omega(\tau) = \frac{1}{a(\tau)} e^{-i\Omega\tau},$$

and to a good approximation we can replace  $a(\tau)$  by  $a(\tau_0)$ . Thus we see that the same answer is obtained for distant sources as in our Minkowski spacetime calculation, even for sources at cosmological distances. From Eq. (4.39) we obtain

$$\delta n^{\hat{i}}(\tau_0, \mathbf{n}) = \frac{n^i + p^i}{2(1 + \mathbf{p} \cdot \mathbf{n})} h_{jk}(0) n_j n_k - \frac{1}{2} h_{ij}(0) n_j \quad (4.58)$$

for plane waves in the direction  $\mathbf{p}$ .

## 4.4 Calculation of Angular Deflection Correlation Function

Now that we have calculated the deflection of the observed direction to a distant source due to an arbitrary metric perturbation  $h_{ij}$ , we would like to determine the properties of the deflection produced by a SGWB, such as that produced by inflation.

#### 4.4.1 Description of SGWB as a Random Process

In the distant source limit, the angular deflection (4.58) depends only on the GW field  $h_{ij}$  evaluated at the location of the observer for each direction of propagation  $\mathbf{p}$ . Moreover, we have restricted attention to modes with wavelengths short compared to the Hubble time. Therefore, it is sufficient to use a flat spacetime mode expansion to describe the stochastic background. This expansion is (see, e.g., Ref. (191))

$$h_{ij}(\mathbf{x}, t) = \sum_{A=+, \times} \int_0^\infty df \int d^2\Omega_{\mathbf{p}} h_{A\mathbf{p}}(f) e^{2\pi i f(\mathbf{p}\cdot\mathbf{x}-t)} e_{ij}^{A,\mathbf{p}} + c.c., \quad (4.59)$$

where  $f$  and  $\mathbf{p}$  are the frequency and direction of propagation of individual GW modes,  $h_{A\mathbf{p}}$  are the stochastic amplitudes of modes with polarization  $A$  and direction  $\mathbf{p}$ , and the polarization tensors  $e_{ij}^{A,\mathbf{p}}$  are normalized such that  $e_{ij}^{A,\mathbf{p}} e_{ij}^{B,\mathbf{p}*} = 2\delta^{AB}$ .

We will assume that  $h_{ij}(\mathbf{x}, t)$  is a Gaussian random process, as it is likely to be the sum of a large number of random processes. We also assume that it is zero-mean and stationary. It follows that the mode amplitudes  $h_{A\mathbf{p}}(f)$  satisfy

$$\begin{aligned} \langle h_{A\mathbf{p}}(f) h_{B\mathbf{p}'}(f') \rangle &= 0, \\ \langle h_{A\mathbf{p}}(f) h_{B\mathbf{p}'}(f')^* \rangle &= \frac{3H_0^2 \Omega_{\text{gw}}(f)}{32\pi^3 f^3} \delta(f - f') \delta_{AB} \delta^2(\mathbf{p}, \mathbf{p}') \end{aligned} \quad (4.60)$$

for  $f, f' \geq 0$ , where  $H_0$  is the Hubble parameter and  $\delta^2(\mathbf{p}, \mathbf{p}')$  is the delta function on the unit sphere (see, e.g., (191)).

Since the angular deflection  $\delta\mathbf{n}(\mathbf{n}, t)$  depends linearly on the metric perturbation, it will also be a stationary, zero-mean, Gaussian random process, whose statistical properties are determined by its two point correlation function  $\langle \delta n^i \delta n^j \rangle$ . Specializing our expression (4.58) for the angular deflection to the form (4.59) of the metric perturbation, we find

$$\delta n^i(\mathbf{n}, t) = \sum_{A=+, \times} \int_0^\infty df \int d^2\Omega_{\mathbf{p}} h_{A\mathbf{p}}(f) e^{-2\pi i f t} \mathcal{R}_{ikl}(\mathbf{n}, \mathbf{p}) e_{kl}^{A,\mathbf{p}} + c.c., \quad (4.61)$$

where

$$\mathcal{R}_{ikl}(\mathbf{n}, \mathbf{p}) = \frac{1}{2} \left( \frac{[n_i + p_i] n_k n_l}{1 + \mathbf{p} \cdot \mathbf{n}} - n_k \delta_{il} \right). \quad (4.62)$$

#### 4.4.2 Power Spectrum of the Astrometric Deflections of the SGWB

So, we need only evaluate the two-point correlation function to gain full knowledge of the statistical properties of the angular deflection due to the SGWB. Writing out this quantity explicitly using Eq.

(4.61),

$$\begin{aligned} \langle \delta n^i(\mathbf{n}, t) \delta n^j(\mathbf{n}', t') \rangle = & \sum_{A,B=+,\times} \int_0^\infty df df' \int d^2\Omega_{\mathbf{p}} d^2\Omega_{\mathbf{p}'} \left\langle \left[ h_{A\mathbf{p}}(f) e^{-2\pi i f t} \mathcal{R}_{ikl}(\mathbf{n}, \mathbf{p}) e_{kl}^{A,\mathbf{p}} + c.c. \right] \right. \\ & \left. \times \left[ h_{B\mathbf{p}'}(f')^* e^{2\pi i f' t'} \mathcal{R}_{jrs}(\mathbf{n}', \mathbf{p}') (e_{rs}^{B,\mathbf{p}'})^* + c.c. \right] \right\rangle. \end{aligned} \quad (4.63)$$

The average, which is an average over ensembles, acts only on the stochastic amplitudes  $h_{A\mathbf{p}}$ . Using the mode 2 point function (4.60) in Eq. (4.63), we get the simplified result

$$\langle \delta n^i(\mathbf{n}, t) \delta n^j(\mathbf{n}', t') \rangle = \int_0^\infty df \frac{3H_0^2}{32\pi^3} f^{-3} \Omega_{\text{gw}}(f) e^{-2\pi i f(t-t')} H_{ij}(\mathbf{n}, \mathbf{n}') + c.c., \quad (4.64)$$

where we have defined

$$H_{ij}(\mathbf{n}, \mathbf{n}') = \sum_{A=+,\times} \int d^2\Omega_{\mathbf{p}} \mathcal{R}_{ikl}(\mathbf{n}, \mathbf{p}) e_{kl}^{A,\mathbf{p}} \mathcal{R}_{jrs}(\mathbf{n}', \mathbf{p}) (e_{rs}^{A,\mathbf{p}})^*. \quad (4.65)$$

#### 4.4.3 Basis Tensors and Their Symmetries

We simplify the expression (4.65) for  $H_{ij}$  further using the identity

$$\sum_{A=+,\times} e_{ij}^{A,\mathbf{p}} (e_{kl}^{A,\mathbf{p}})^* = 2P_{ijkl}, \quad (4.66)$$

where  $P_{ijkl}$  is the projection tensor onto the space of traceless symmetric tensors orthogonal to  $\mathbf{p}$ , given by

$$2P_{ijkl} = \delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk} - \delta_{ij}\delta_{kl} + p_i p_j p_k p_l - \delta_{ik} p_j p_l - \delta_{jl} p_i p_k - \delta_{il} p_j p_k - \delta_{jk} p_i p_l + \delta_{ij} p_k p_l + \delta_{kl} p_i p_j. \quad (4.67)$$

This gives

$$H_{ij}(\mathbf{n}, \mathbf{n}') = 2 \int d^2\Omega_{\mathbf{p}} \mathcal{R}_{ikl}(\mathbf{n}, \mathbf{p}) P_{klrs} \mathcal{R}_{jrs}(\mathbf{n}', \mathbf{p}). \quad (4.68)$$

Noting that the correlation function (4.64) is perpendicular to  $\mathbf{n}$  on its first index and  $\mathbf{n}'$  on its second, we can decompose it onto a basis of tensors with this property:

$$H_{ij}(\mathbf{n}, \mathbf{n}') = \alpha(\mathbf{n}, \mathbf{n}') A_i A_j + \beta(\mathbf{n}, \mathbf{n}') A_i C_j + \gamma(\mathbf{n}, \mathbf{n}') B_i A_j + \sigma(\mathbf{n}, \mathbf{n}') B_i C_j, \quad (4.69)$$

for some scalar functions  $\alpha$ ,  $\beta$ ,  $\gamma$  and  $\sigma$ . Here we have defined

$$\mathbf{A} = \mathbf{n} \times \mathbf{n}', \quad \mathbf{B} = \mathbf{n} \times \mathbf{A}, \quad \mathbf{C} = -\mathbf{n}' \times \mathbf{A}. \quad (4.70)$$

We can deduce from Eq. (4.68) that  $H_{ij}(\mathbf{n}, \mathbf{n}')^* = H_{ji}(\mathbf{n}', \mathbf{n})$ . Noting that  $A_i(\mathbf{n}', \mathbf{n}) = -A_i(\mathbf{n}, \mathbf{n}')$ , and  $B_i(\mathbf{n}', \mathbf{n}) = -C_i(\mathbf{n}, \mathbf{n}')$ , this symmetry applied to the expansion (4.69) gives

$$\alpha(\mathbf{n}, \mathbf{n}')^* = \alpha(\mathbf{n}', \mathbf{n}), \quad \sigma(\mathbf{n}, \mathbf{n}')^* = \sigma(\mathbf{n}', \mathbf{n}), \quad \beta(\mathbf{n}, \mathbf{n}')^* = \gamma(\mathbf{n}', \mathbf{n}).$$

We see from Eq. (4.65) that  $H_{ij}$  transforms as tensor under rotations. This implies that the functions  $\alpha$ ,  $\beta$ ,  $\gamma$  and  $\sigma$  must be invariant under rotations, and can only depend on the angle  $\Theta$  between  $\mathbf{n}$  and  $\mathbf{n}'$ . Thus,  $\alpha(\mathbf{n}, \mathbf{n}') = \alpha(\mathbf{n}', \mathbf{n}) = \alpha(\Theta)$  and so forth, so  $\alpha$  and  $\sigma$  must be real.

Next, we note that the expression (4.68) for  $H_{ij}(\mathbf{n}, \mathbf{n}')$  is invariant under the parity transformation  $\mathbf{n} \rightarrow -\mathbf{n}$  and  $\mathbf{n}' \rightarrow -\mathbf{n}'$ . Looking then at the basis tensors, we see that  $\mathbf{A}$  is invariant under this transformation, while  $\mathbf{B}$  and  $\mathbf{C}$  change sign. Thus, in order to insure that  $H_{ij}$  is invariant, it can only have terms multiplying  $A_i A_j$  and  $B_i C_j$ , so  $\beta(\Theta) = 0 = \gamma(\Theta)$ .

Having taken the symmetries of the problem into consideration, we have found  $H_{ij}$  to be of the form

$$H_{ij}(\mathbf{n}, \mathbf{n}') = \alpha(\Theta) A_i A_j + \sigma(\Theta) B_i C_j. \quad (4.71)$$

#### 4.4.4 Solving the General Integral

We can evaluate the coefficients in the expansion (4.71) of  $H_{ij}$  by contracting it with the basis tensors:

$$A^i A^j H_{ij} = \sin^4(\Theta) \alpha(\Theta), \quad B^i C^j H_{ij} = \sin^4(\Theta) \sigma(\Theta).$$

Rewriting these using Eq. (4.68), we find

$$\alpha(\Theta) = \frac{2}{\sin^4(\Theta)} \int d^2 \Omega_{\mathbf{p}} A^i \mathcal{R}_{ikl}(\mathbf{n}, \mathbf{p}) P_{klrs} A^j \mathcal{R}_{jrs}(\mathbf{n}', \mathbf{p})^*, \quad (4.72)$$

$$\sigma(\Theta) = \frac{2}{\sin^4(\Theta)} \int d^2 \Omega_{\mathbf{p}} B^i \mathcal{R}_{ikl}(\mathbf{n}, \mathbf{p}) P_{klrs} C^j \mathcal{R}_{jrs}(\mathbf{n}', \mathbf{p})^*. \quad (4.73)$$

To simplify the calculation, we define the quantities  $\kappa = \mathbf{n} \cdot \mathbf{p}$ ,  $\kappa' = \mathbf{n}' \cdot \mathbf{p}$ ,  $\lambda = \mathbf{n} \cdot \mathbf{n}'$ ,  $\mu = \mathbf{A} \cdot \mathbf{p}$ , which satisfy  $\mu^2 + \lambda^2 + \kappa^2 + \kappa'^2 = 1 + 2\lambda\kappa\kappa'$ . Using these definitions and the definition (4.62) of  $\mathcal{R}_{ikl}$ , we can write

$$\begin{aligned}
A^i \mathcal{R}_{ikl}(\mathbf{n}, \mathbf{p}) &= \frac{1}{2} n_k \left( \frac{\mu n_l}{1 + \kappa} - A_l \right), & A^j \mathcal{R}_{jrs}(\mathbf{n}', \mathbf{p}) &= \frac{1}{2} n'_r \left( \frac{\mu n'_s}{1 + \kappa'} - A_s \right), \\
B^i \mathcal{R}_{ikl}(\mathbf{n}, \mathbf{p}) &= \frac{1}{2} n_k \left( -\frac{\kappa' + \lambda}{1 + \kappa} n_l + n'_l \right), & C^j \mathcal{R}_{jrs}(\mathbf{n}', \mathbf{p}) &= \frac{1}{2} n'_r \left( -\frac{\kappa + \lambda}{1 + \kappa'} n'_s + n_s \right).
\end{aligned}$$

We can then rewrite our expressions for  $\alpha$  and  $\sigma$

$$\alpha(\Theta) = \frac{1}{4 \sin^4(\Theta)} \int d^2 \Omega_{\mathbf{p}} 2P_{klrs} n_k \left( \frac{\mu n_l}{1 + \kappa} - A_l \right) n'_r \left( \frac{\mu n'_s}{1 + \kappa'} - A_s \right), \quad (4.74)$$

$$\sigma(\Theta) = \frac{1}{4 \sin^4(\Theta)} \int d^2 \Omega_{\mathbf{p}} 2P_{klrs} n_k \left( -\frac{\kappa' + \lambda}{1 + \kappa} n_l + n'_l \right) n'_r \left( -\frac{\kappa + \lambda}{1 + \kappa'} n'_s + n_s \right). \quad (4.75)$$

Let's define two new variables  $\nu^2 = (1 - \kappa^2)$ ,  $\nu'^2 = (1 - \kappa'^2)$ . Applying the definition (4.67) of the projection tensor  $P_{klrs}$ , we can calculate the necessary contractions of  $P_{klrs}$  for  $\alpha$ :

$$\begin{aligned}
2P_{klrs} n_k A_l n'_r A_s &= (\lambda - \kappa \kappa') (1 - \lambda^2 - \mu^2), & 2P_{klrs} n_k n_l n'_r A_s &= \mu (\kappa' \kappa^2 - 2\lambda \kappa + \kappa'), \\
2P_{klrs} n_k A_l n'_r n'_s &= \mu (\kappa \kappa'^2 - 2\lambda \kappa' + \kappa), & 2P_{klrs} n_k n_l n'_r n'_s &= \nu^2 \nu'^2 - 2\mu^2,
\end{aligned} \quad (4.76)$$

and for  $\sigma$ :

$$\begin{aligned}
2P_{klrs} n_k n'_l n'_r n_s &= \nu^2 \nu'^2, & 2P_{klrs} n_k n_l n'_r n_s &= \nu^2 (\lambda - \kappa \kappa'), \\
2P_{klrs} n_k n_l n'_r n'_s &= \nu^2 \nu'^2 - 2\mu^2, & 2P_{klrs} n_k n'_l n'_r n'_s &= \nu'^2 (\lambda - \kappa \kappa').
\end{aligned} \quad (4.77)$$

Plugging these back into Eqs. (4.74) and (4.75) and simplifying, we find

$$\alpha(\Theta) = \frac{1}{4 \sin^4(\Theta)} \int d^2 \Omega_{\mathbf{p}} \left[ (\lambda - \kappa \kappa') (1 - \lambda^2) - \mu^2 (1 + \lambda) + \frac{2\mu^2 (\lambda + \kappa) (\lambda + \kappa')}{(1 + \kappa) (1 + \kappa')} \right] = -\sigma(\Theta).$$

Noticing that we can do the integrals  $\int d^2 \Omega_{\mathbf{p}} \mu^2 = \frac{4\pi}{3} \sin^2 \Theta$  and  $\int d^2 \Omega_{\mathbf{p}} \kappa \kappa' = \frac{4\pi}{3} \cos \Theta$ , but that the last term is more complicated, we find

$$\alpha(\Theta) = -\sigma(\Theta) = \frac{\pi (\cos \Theta - 1)}{3 \sin^2 \Theta} + \frac{1}{2 \sin^4 \Theta} \int d^2 \Omega_{\mathbf{p}} \frac{\mu^2 (\lambda + \kappa) (\lambda + \kappa')}{(1 + \kappa) (1 + \kappa')}. \quad (4.78)$$

We can reduce the two dimensional integral (4.78) to a one dimensional integral by parameterizing  $\mathbf{p}$  in spherical polar coordinates  $\theta_p$  and  $\phi_p$ , choosing  $\mathbf{n} = (0, \sin(\Theta/2), \cos(\Theta/2))$  and  $\mathbf{n}' = (0, -\sin(\Theta/2), \cos(\Theta/2))$  and integrating over  $\phi_p$ . This gives

$$\alpha(\Theta) = -\sigma(\Theta) = \frac{\pi}{3} \frac{(\cos(\Theta) - 1)}{\sin^2(\Theta)} + \frac{\pi}{2 \sin^2 \Theta} \int_0^\pi d\theta_p \sin \theta_p \{ \sin^2 \theta_p + 8 \cos(\Theta/2) [\cos \theta_p + \cos(\Theta/2)] [g(\theta_p, \Theta) - 1] \}, \quad (4.79)$$

where

$$g(\theta_p, \Theta) = \frac{|\cos \theta_p + \cos(\Theta/2)|}{[1 + \cos \theta_p \cos(\Theta/2)]}. \quad (4.80)$$

We perform the integral over  $\theta_p$ , and find the final form of the function  $\alpha(\Theta)$

$$\alpha(\Theta) = -\sigma(\Theta) = \frac{\pi}{3 \sin^2 \Theta} (7 \cos \Theta - 5) - \frac{32\pi}{\sin^4 \Theta} \ln (\sin(\Theta/2)) \sin^6(\Theta/2). \quad (4.81)$$

A plot of the function  $\alpha(\Theta)$  is shown in Figure 4.2.

To summarize, we have now completed the calculation of the angular deflection correlation function. The final answer is given by Eq. (4.64), with  $H_{ij}(\mathbf{n}, \mathbf{n}')$  given from Eqs. (4.71) and (4.81) as

$$H_{ij}(\mathbf{n}, \mathbf{n}') = \alpha(\Theta) (A_i A_j - B_i C_j). \quad (4.82)$$

Here the vectors  $\mathbf{A}$ ,  $\mathbf{B}$ , and  $\mathbf{C}$  are defined by Eqs. (4.70), and  $\alpha(\Theta)$  is given by Eq. (4.81).

#### 4.4.5 Special Case: Coincidence

As a check of our calculation, we can solve for the two-point correlation function exactly in the case that  $\mathbf{n} = \mathbf{n}'$ . Using Eqs. (4.62), (4.67) and (4.68), the integral simplifies to

$$H_{ij}(\mathbf{n}, \mathbf{n}) = \frac{1}{4} \int d^2 \Omega_{\mathbf{p}} [1 - (\mathbf{p} \cdot \mathbf{n})^2] (\delta_{ij} - n_i n_j).$$

We can solve this integral analytically, getting

$$H_{ij}(\mathbf{n}, \mathbf{n}) = \frac{2\pi}{3} (\delta_{ij} - n_i n_j). \quad (4.83)$$

This corresponds to the limit of  $\alpha(\Theta)(A^i A^j - B^i C^j)$  as  $\mathbf{n} \rightarrow \mathbf{n}'$ , with  $\alpha(\Theta) = 2\pi/(3\Theta^2) + \mathcal{O}(\Theta^{-1})$  from Eq. (4.81). Inserting the coincidence limit (4.83) into the correlation function (4.64) yields the formula (4.3) for the total rms angular fluctuations discussed in the introduction.

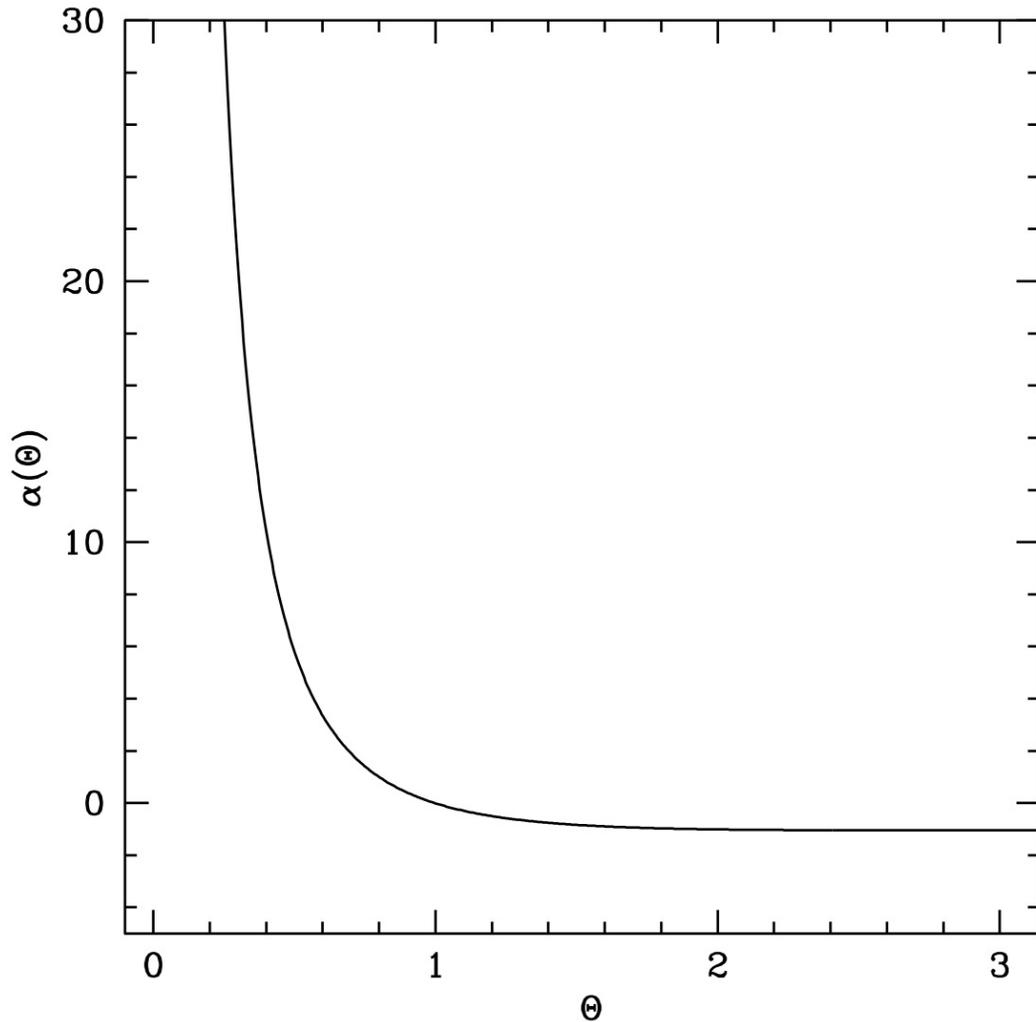


Figure 4.2: Here we plot the function  $\alpha(\Theta)$ , the coefficient of  $H_{ij}(\mathbf{n}, \mathbf{n}')$  as shown in Eq. (4.82), as a function of the angle  $\Theta$  between  $\mathbf{n}$  and  $\mathbf{n}'$ .

## 4.5 Spectrum of Angular Deflection Fluctuations

### 4.5.1 Overview

In the previous section we computed the correlation function  $\langle \delta n^i(\mathbf{n}, t) \delta n^j(\mathbf{n}', t') \rangle$  as a function of the unit vectors  $\mathbf{n}$  and  $\mathbf{n}'$ . However for many purposes it is more useful to perform a multipole decomposition of the angular deflection, and to compute the spectrum of fluctuations on different angular scales  $l$ , as is done with cosmic microwave background anisotropies. We decompose  $\delta \mathbf{n}(\mathbf{n}, t)$  as

$$\delta \mathbf{n}(\mathbf{n}, t) = \sum_{lm} \delta n_{Elm}(t) \mathbf{Y}_{lm}^E(\mathbf{n}) + \delta n_{Blm}(t) \mathbf{Y}_{lm}^B(\mathbf{n}), \quad (4.84)$$

where  $\mathbf{Y}_{lm}^E$  and  $\mathbf{Y}_{lm}^B$  are the electric- and magnetic-type transverse vector spherical harmonics defined by

$$\mathbf{Y}_{lm}^E(\mathbf{n}) = (l(l+1))^{-1/2} \nabla Y_{lm}(\mathbf{n}), \quad \mathbf{Y}_{lm}^B(\mathbf{n}) = (l(l+1))^{-1/2} (\mathbf{n} \times \nabla) Y_{lm}(\mathbf{n}). \quad (4.85)$$

We will show in this section that the statistical properties of the coefficients are given by

$$\langle \delta n_{Qlm}(t) \delta n_{Q'l'm'}(t')^* \rangle = \delta_{QQ'} \delta_{ll'} \delta_{mm'} \int_0^\infty df \cos[2\pi f(t-t')] S_{Ql}(f) \quad (4.86)$$

for  $Q, Q' = E$  or  $B$ , for some spectrum  $S_{Ql}(f)$ , a function of frequency  $f$  and of angular scale  $l$ . The formula (4.86) shows that different multipoles of the angular deflection are statistically independent, as required by spherical symmetry of the stochastic background. Also the electric-type and magnetic-type fluctuations are uncorrelated, as required by parity invariance of the stochastic background (see below).

The spectrum  $S_{Ql}(f)$  is given by

$$S_{Ql}(f) = \frac{4\pi}{2l+1} \theta_{rms}^2 \frac{\sigma(f)}{f} g_Q \alpha_l^{QQ}. \quad (4.87)$$

Here  $\theta_{rms}^2$  is the total rms angular fluctuation squared, given by Eq. (4.3) in the introduction. The function  $\sigma(f)$  describes how the power is distributed in frequency. It is the same for all multipoles, is normalized according to  $\int d(\ln f) \sigma(f) = 1$ , and is given explicitly by Eq. (4.6) in the introduction. The quantities  $g_E$  and  $g_B$  are the fraction of the total power in electric-type and magnetic-type fluctuations, and are  $g_E = g_B = 1/2$ . Finally the angular spectra  $\alpha_l^{EE}$  and  $\alpha_l^{BB}$  describe the dependence on angular scale, which is the same for all frequencies. They are normalized according to

$$\sum_{l=2}^{\infty} \alpha_l^{QQ} = 1, \quad (4.88)$$

and are the same for  $E$  and  $B$  modes,  $\alpha_l^{EE} = \alpha_l^{BB}$ . This spectrum is plotted in Fig 4.1 and the first 10 values are listed in Table 4.1. We note that these coefficients are well fit by the power law  $\alpha_l^{EE} = 32.34 l^{-4.921}$ .

Before proceeding with the derivation of the spectrum (4.86), we first derive from (4.86) the expression (4.4) discussed in the introduction for the total fluctuation power. Squaring the expansion (4.84), taking an expected value, and then using (4.86) gives

$$\begin{aligned}
\langle \delta \mathbf{n}(\mathbf{n}, t)^2 \rangle &= \sum_{Qlm} \sum_{Q'l'm'} \mathbf{Y}_{lm}^Q(\mathbf{n}) \mathbf{Y}_{l'm'}^{Q'}(\mathbf{n})^* \langle \delta n_{Qlm}(t) \delta n_{Q'l'm'}(t')^* \rangle \\
&= \sum_{Ql} \int_0^\infty \frac{\sigma(f)}{f} \sum_{m=-l}^l \left| \mathbf{Y}_{lm}^Q(\mathbf{n}) \right|^2 \theta_{rms}^2 \frac{4\pi}{2l+1} g_Q \alpha_l^{QQ}.
\end{aligned} \tag{4.89}$$

Using Unsöld's theorem for vector spherical harmonics,

$$\sum_{m=-l}^l \left| \mathbf{Y}_{lm}^Q(\mathbf{n}) \right|^2 = \frac{2l+1}{4\pi},$$

gives

$$\langle \delta \mathbf{n}(\mathbf{n}, t)^2 \rangle = \sum_{Ql} \int_0^\infty \theta_{rms}^2 \frac{\sigma(f)}{f} g_Q \alpha_l^{QQ}, \tag{4.90}$$

which reduces to Eq. (4.4). Note that using the normalization conventions for  $\alpha_l^{QQ}$  and  $\sigma(f)$  now gives  $\langle \delta \mathbf{n}(\mathbf{n}, t)^2 \rangle = \theta_{rms}^2 (g_E + g_B) = \theta_{rms}^2$ , showing consistency of the definitions.

## 4.5.2 Derivation

We now turn to a derivation of the spectrum (4.87). First we note that the vector spherical harmonics are transverse in the sense that  $\mathbf{Y}_{lm}^Q(\mathbf{n}) \cdot \mathbf{n} = 0$  for  $Q = E, B$ , and are orthogonal in the sense that

$$\int d^2\Omega_{\mathbf{n}} Y_{lmi}^Q(\mathbf{n}) Y_{l'm'}^{Q'i*}(\mathbf{n}) = \delta_{QQ'} \delta_{ll'} \delta_{mm'}.$$

Using this orthogonality property, we can extract the coefficients of the expansion (4.84)

$$\delta n_{Qlm}(t) = \int d^2\Omega_{\mathbf{n}} \delta n_i(\mathbf{n}, t) Y_{lm}^{Qi*}(\mathbf{n}).$$

Thus we can write for the correlation function between two of these coefficients

$$\langle \delta n_{Qlm}(t) \delta n_{Q'l'm'}(t')^* \rangle = \int d^2\Omega_{\mathbf{n}} d^2\Omega_{\mathbf{n}'} Y_{lmi}^{Q*}(\mathbf{n}) Y_{l'm'j}^{Q'}(\mathbf{n}') \langle \delta n^i(\mathbf{n}, t) \delta n^j(\mathbf{n}', t') \rangle, \tag{4.91}$$

or more explicitly, using Eq. (4.64)

$$\langle \delta n_{Qlm}(t) \delta n_{Q'l'm'}(t')^* \rangle = \frac{3H_0^2}{16\pi^3} \int_0^\infty df \cos[2\pi f(t-t')] \frac{\Omega_{\text{gw}}(f)}{f^3} C_{QlmQ'l'm'}, \tag{4.92}$$

where

$$C_{QlmQ'l'm'} = \int d^2\Omega_{\mathbf{n}} d^2\Omega_{\mathbf{n}'} Y_{lm}^{Q*}(\mathbf{n}) Y_{l'm'}^{Q'}(\mathbf{n}') H_{ij}(\mathbf{n}, \mathbf{n}'). \quad (4.93)$$

We now argue that the  $EB$  cross-correlation vanishes. From Eq. (4.85), we see that  $\mathbf{Y}_{lm}^E(\mathbf{n})$  has the same parity under  $\mathbf{n} \rightarrow -\mathbf{n}$  as  $Y_{lm}(\mathbf{n})$ , while the parity of  $\mathbf{Y}_{lm}^B(\mathbf{n})$  is opposite. From Section 4.4.3 above  $H_{ij}(\mathbf{n}, \mathbf{n}')$  is invariant under both  $\mathbf{n} \rightarrow -\mathbf{n}$  and  $\mathbf{n}' \rightarrow -\mathbf{n}'$ . Thus, if  $Q = E$ ,  $Q' = B$  in Eq. (4.92), the integral will be symmetric under  $\mathbf{n} \rightarrow -\mathbf{n}$  but antisymmetric under  $\mathbf{n}' \rightarrow -\mathbf{n}'$ , causing the integral over  $d^2\Omega_{\mathbf{n}'}$  to vanish. Therefore,  $EB$  cross correlations vanish, and we need only calculate the  $EE$  and  $BB$  correlation functions.

#### 4.5.2.1 EE correlation

Inserting the definition (4.85) of the electric vector spherical harmonics and the formula (4.82) for  $H_{ij}$  into Eq. (4.93) and integrating by parts, we obtain

$$C_{ElmE'l'm'} = \frac{1}{l(l+1)} \int d^2\Omega_{\mathbf{n}} d^2\Omega_{\mathbf{n}'} Y_{lm}^*(\mathbf{n}) Y_{l'm'}(\mathbf{n}') \beta^{EE}(\Theta), \quad (4.94)$$

where the function  $\beta^{EE}$  is given by

$$\beta^{EE}(\Theta) = \nabla_i \nabla'_j [H_{ij}(\mathbf{n}, \mathbf{n}')] = \nabla_i \nabla'_j \{ \alpha(\Theta) [A_i A_j - B_i C_j] \}. \quad (4.95)$$

Here  $\nabla_i$  and  $\nabla'_j$  denote normal three dimensional derivatives with respect to  $\mathbf{x}$  and  $\mathbf{x}'$ , where  $\mathbf{n} = \mathbf{x}/|\mathbf{x}|$  and  $\mathbf{n}' = \mathbf{x}'/|\mathbf{x}'|$ . Integration by parts on the unit sphere of this derivative operator is valid as long as the radial component of the integrand vanishes, from the identity  $\nabla_i v^i = \partial_r v^r + 2v_r/r + \nabla_A v^A$ , where  $\nabla_A$  denotes a covariant derivative on the unit sphere. It can be checked that the radial components do vanish in the above computation.

Next, we expand the function  $\beta^{EE}$  in terms of Legendre polynomials, and use the spherical harmonic addition theorem, which gives

$$\begin{aligned} \beta^{EE}(\Theta) &= \sum_l \beta_l^{EE} P_l(\cos \Theta) \\ &= \sum_{lm} \frac{4\pi}{2l+1} \beta_l^{EE} Y_{lm}(\mathbf{n}) Y_{lm}(\mathbf{n}')^* \end{aligned} \quad (4.96)$$

Inserting this into Eq. (4.94) and using the orthogonality of spherical harmonics gives

$$C_{ElmE'l'm'} = \delta_{ll'} \delta_{mm'} \frac{1}{l(l+1)} \frac{4\pi}{2l+1} \beta_l^{EE}. \quad (4.97)$$

Inserting this into Eq. (4.92) now yields the correlation function given by Eqs. (4.86) and (4.87),

and using the definitions (4.3) and (4.6) of  $\theta_{rms}^2$  and  $\sigma(f)$  allows us to read off the electric multipole spectrum

$$g_E \alpha_l^{EE} = \frac{3}{4\pi l(l+1)} \beta_l^{EE}. \quad (4.98)$$

We will show below that  $g_E = 1/2$ .

It remains to explicitly evaluate the function  $\beta^{EE}(\Theta)$  defined in Eq. (4.95) and evaluate its expansion coefficients. We have

$$\beta^{EE}(\Theta) \equiv \nabla_i \nabla'_j [\alpha(\Theta) T^{ij}] = [\nabla_i \nabla'_j \alpha(\Theta)] T^{ij} + [\nabla_i \alpha(\Theta)] (\nabla'_j T^{ij}) + [\nabla'_j \alpha(\Theta)] (\nabla_i T^{ij}) + \alpha(\Theta) (\nabla_i \nabla'_j T^{ij}), \quad (4.99)$$

where we have defined  $T^{ij} = (A^i A^j(\mathbf{n}, \mathbf{n}') - B^i C^j(\mathbf{n}, \mathbf{n}'))$ . Using  $A^i = \epsilon^{ijk} n_j n'_k$ ,  $B^i = (\mathbf{n} \cdot \mathbf{n}') n^i - n^i$ ,  $C^i = (\mathbf{n} \cdot \mathbf{n}') n^i - n^i$ , we can write the tensor  $T^{ij}$  in Cartesian coordinates as

$$T^{ij} = \epsilon^{ikl} \epsilon^{jrs} n_k n'_l n_r n'_s - ((\mathbf{n} \cdot \mathbf{n}') n^i - n^i) ((\mathbf{n} \cdot \mathbf{n}') n^j - n^j).$$

Using  $\nabla_i n_j = \delta_{ij} - n_i n_j$ ,  $\nabla'_i n'_j = \delta_{ij} - n'_i n'_j$ ,  $\nabla'_i n^j = \nabla_i n'^j = 0$ , and  $\nabla_i \epsilon^{ijk} = \nabla'_i \epsilon^{ijk} = 0$ , we calculate the derivatives

$$\begin{aligned} \nabla_i T^{ij} &= (1 - 3(\mathbf{n} \cdot \mathbf{n}')) ((\mathbf{n} \cdot \mathbf{n}') n^j - n^j), & \nabla'_j T^{ij} &= (1 - 3(\mathbf{n} \cdot \mathbf{n}')) ((\mathbf{n} \cdot \mathbf{n}') n^i - n^i), \\ \nabla_i \nabla'_j T^{ij} &= -9(\mathbf{n} \cdot \mathbf{n}')^2 + 2(\mathbf{n} \cdot \mathbf{n}') + 3. \end{aligned} \quad (4.100)$$

For the gradients of  $\alpha$ , we use the fact that  $\cos(\Theta) = \mathbf{n} \cdot \mathbf{n}'$ , so that  $-\sin(\Theta) \nabla_i \Theta = n'_i - (\mathbf{n} \cdot \mathbf{n}') n_i$ , and similarly for  $\nabla'_j$ . Thus, we find

$$\begin{aligned} \nabla_i \alpha(\Theta) &= -\alpha'(\Theta) \frac{n'_i - (\mathbf{n} \cdot \mathbf{n}') n_i}{\sin(\Theta)}, & \nabla'_j \alpha(\Theta) &= -\alpha'(\Theta) \frac{n_j - (\mathbf{n} \cdot \mathbf{n}') n'_j}{\sin(\Theta)} \\ \nabla_i \nabla'_j \alpha(\Theta) &= \alpha'(\Theta) \left\{ \frac{\delta_{ij} - n_i n_j - n'_i n'_j + (\mathbf{n} \cdot \mathbf{n}') n_i n'_j}{-\sin(\Theta)} + \frac{\cos(\Theta) [n'_i - (\mathbf{n} \cdot \mathbf{n}') n_i] [n_j - (\mathbf{n} \cdot \mathbf{n}') n'_j]}{-\sin^3(\Theta)} \right\} \\ &+ \alpha''(\Theta) \frac{[n'_i - (\mathbf{n} \cdot \mathbf{n}') n_i] [n_j - (\mathbf{n} \cdot \mathbf{n}') n'_j]}{\sin^2(\Theta)}. \end{aligned} \quad (4.101)$$

Plugging Eqs. (4.100) and (4.101) into Eq. (4.99), we get

$$\begin{aligned}\beta^{EE}(\Theta) &= [-9\cos^2(\Theta) + 2\cos(\Theta) + 3]\alpha(\Theta) - \sin^2(\Theta)\alpha''(\Theta) \\ &\quad + [1 - 6\cos(\Theta)]\sin(\Theta)\alpha'(\Theta).\end{aligned}\tag{4.102}$$

Next, we insert the expression (4.81) for  $\alpha(\Theta)$  to obtain

$$\beta^{EE}(\Theta) = \frac{4\pi}{3}\left(4 + (1 - \cos\Theta)\{12\ln[\sin(\Theta/2)] - 1\}\right).\tag{4.103}$$

We numerically compute the coefficients  $\beta_l^{EE}$  of the Legendre polynomial expansion (4.96) of  $\beta^{EE}(\Theta)$ , and from them compute  $\alpha_l^{EE}$  using Eq. (4.98). The result is plotted in Figure 4.1 and tabulated in Table 4.1.

#### 4.5.2.2 BB correlation

We now calculate the  $BB$  correlation in a similar manner to the  $EE$  case above. Inserting into Eq. (4.93) the definition (4.85) of magnetic vector spherical harmonics and integrating by parts, we find

$$C_{BlmBl'm'} = \frac{1}{l(l+1)} \int d^2\Omega_{\mathbf{n}} d^2\Omega_{\mathbf{n}'} Y_{lm}^*(\mathbf{n}) Y_{l'm'}(\mathbf{n}') \beta^{BB}(\Theta),$$

where

$$\beta^{BB}(\Theta) = \nabla_l \nabla_p' [\epsilon_{ikl} \epsilon_{jmp} n_k n_m' \alpha(\Theta) T_{ij}].\tag{4.104}$$

As before, we can derive from here the form (4.86) and (4.87) of the spectrum, with  $\alpha_l^{BB}$  given by

$$g_B \alpha_l^{BB} = \frac{3}{4\pi l(l+1)} \beta_l^{BB}.$$

We now show that  $\beta^{BB}(\Theta) = \beta^{EE}(\Theta)$ , from which it follows that  $g_E = g_B = 1/2$  and that  $\alpha_l^{EE} = \alpha_l^{BB}$ . To see this we evaluate the cross products in (4.104) using  $\mathbf{n} \times \mathbf{A} = \mathbf{B}$ ,  $\mathbf{n} \times \mathbf{B} = -\mathbf{A}$ ,  $\mathbf{n}' \times \mathbf{C} = \mathbf{A}$ . This gives

$$\epsilon_{ikl} \epsilon_{jmp} n_k n_m' H_{ij} = H_{lp},$$

and using the definitions (4.95) and (4.104) of  $\beta^{EE}$  and  $\beta^{BB}$ , it follows that  $\beta^{BB} = \beta^{EE}$ .

## Chapter 5

# Bipolar Spherical Harmonic Decomposition<sup>1</sup>

### 5.1 Introduction

In this chapter we present the method of bipolar spherical harmonic (BiPoSH) decomposition as a way to characterize departures from Gaussianity and isotropy of a field on the sky. We then discuss two example applications, the temperature of the cosmic microwave background (CMB) radiation and the intensity of 21 cm radiation from the dark ages. The remainder of this chapter is structured as follows: background on the CMB, the 21 cm radiation, weak gravitational lensing, and BiPoSH decomposition is given in the rest of Section 5.1. In Section 5.2 we review the mathematics of BiPoSHs, and in Section 5.3 we present a detailed calculation of the expected BiPoSH signals from the lensing of the CMB (or another field) by intervening scalar and tensor perturbations. In Section 5.4 we calculate the possible constraint on chiral GWs, as an example of a parity-violating process, using the correlation between opposite parity BiPoSH coefficients and that between opposite-parity BiPoSH and CMB polarization modes. Finally, in Section 5.5 we discuss the possibility of also applying the BiPoSH formalism to future 21 cm intensity maps, and the resulting constraint on primordial GWs that could be obtained.

#### 5.1.1 Cosmic Microwave Background

The CMB is a nearly instantaneous snapshot of the surface of last scatter at redshift  $z \sim 1100$ , when electrons and protons combined to form neutral atoms, allowing photons to stream freely. The spectrum of cosmic microwave background (CMB) radiation is an almost perfect black-body

---

<sup>1</sup>The material in this chapter was adapted from “Odd-Parity Bipolar Spherical Harmonics,” L. Book, M. Kamionkowski, and T. Souradeep, PRD **85**, 023010 (2012) and “Lensing of 21-cm Fluctuations by Primordial Gravitational Waves,” L. Book, M. Kamionkowski, and F. Schmidt, submitted to PRL. Reproduced here with permission, copyright (2012) by the American Physical Society. Introductory material has been significantly supplemented, the discussion of non-cosmological sources of odd-parity BiPoSH in Section 5.3.5 was increased, and Figure 5.6 and associated discussion was added to Section 5.5.

spectrum with a temperature of 2.73 K (208). The prediction of this spectral shape, from the idea that the Universe had expanded and cooled from a hot and dense initial state, was one of the great successes of the big bang model. This radiation is almost completely isotropic, with fluctuations around five orders of magnitude smaller than its overall amplitude. The detection of anisotropies in the CMB (209) has revolutionized the precision with which cosmological measurements can be made. This precision comes from the fact that, unlike large-scale structure, the small photon perturbations from early times remain linear all the way to the present time, and therefore provide a far simpler system from which to extract cosmological information.

Most of the information that has been obtained from the CMB so far has come from its power spectrum, the two-point correlation function, under the assumptions of isotropy and homogeneity. Indeed, the amount and precision of information available from this function is extraordinary. The power spectrum has peaks and troughs resulting from acoustic oscillations of the photon-baryon fluid prior to recombination, and the precise position and height of these peaks imparts a great deal of information about the content of the Universe at those times. For example, the relative height of odd- and even-numbered peaks gives direct information about the density of baryons, due to the retarding effect of baryons on the sound speed. The amplitude of the peaks is affected not only by the primordial amplitude and spectral tilt of perturbations, but also by the optical depth to reionization, which has the effect to wipe out perturbations, the presence of tensor perturbations, the matter and baryon densities, the cosmological constant, and the curvature of the Universe, as well as many other parameters at a lower level. While it may seem that this large number of parameters must necessarily be degenerate, the varying effects of these parameters on the heights of each peak, their positions, and the amplitude of large angle fluctuations, together with information from other sources such as supernovae and galaxy surveys, allows for extremely precise determination of all of these parameters together.

Motivated by the extreme utility of the CMB temperature power spectrum, in recent years, attention has been paid to effects that go beyond the power spectrum, such as weak lensing (210), cosmic birefringence (211; 212), and departures from statistical isotropy (SI) (213; 214; 215) and from Gaussianity (216; 217; 218). In this chapter we present the method of bipolar spherical harmonic (BiPoSH) decomposition, which are a general way to parametrize departures from statistical isotropy and Gaussianity, and thereby to efficiently search for information in the CMB beyond the power spectrum.

### 5.1.2 21 cm Radiation

Between the time of recombination, described above as the era probed by the CMB, and reionization, when radiation from the first stars and galaxies dissociated neutral atoms, the Universe was filled with neutral atoms, mostly hydrogen with some helium and heavier isotopes and elements. Thus,

the strongest hope of obtaining information about the Universe during this long period of time, known as the cosmic dark ages (from  $1100 < z < 10$ ) comes from detecting the neutral hydrogen.

Although the optical depth to reionization can be obtained with the CMB, the most promising method of obtaining detailed, three-dimensional maps of the matter during the dark ages comes from the 21 cm line of neutral hydrogen. This is the radiation emitted when the proton of a hydrogen atom flips its spin relative to that of the electron, and is emitted, as its name implies, at a wavelength of 21 cm. This transition is forbidden, with a lifetime of  $\sim 3 \times 10^7$  years, which implies that it is far from saturation (219). This then means that it can be used to probe the density (and not just the presence) of neutral hydrogen. Combined with the fact that this is a spectral line, and therefore also gives a redshift, this allows us to construct three-dimensional maps of the neutral hydrogen density.

While the potential of the 21 cm line to probe cosmology has been known for quite some time, only recently has technology advanced to the point of possible observation of the cosmological 21 cm signal. Several experiments are poised to soon detect the 21 cm signal from the epoch of reionization (220), and there are longer-term prospects to delve into the dark ages (221). Due to these exciting observational prospects, a great deal of effort has gone into the theory of these pre-structure-formation-era density fluctuations (222; 219; 223). In Section 5.5 we contribute to this effort by considering the use of BiPoSH decomposition on 21 cm intensity maps, and probe their utility to detect a background of stochastic GWs.

Atomic hydrogen in the redshift range  $30 \lesssim z \lesssim 200$  can absorb radiation deep in the Rayleigh-Jeans region of the CMB (222). The angular power spectrum of these 21 cm fluctuations extends to multipole moments  $l \sim 10^7$  (limited only by the baryonic Jeans mass) (222), far larger than those,  $l \sim 3000$ , to which the CMB power spectrum extends (beyond which fluctuations are suppressed by Silk damping).

### 5.1.3 Weak Gravitational Lensing

The prediction of the lensing of photons by the sun, and its subsequent observation in 1919, was one of the great successes of general relativity. The subsequent detection of strongly lensed systems such that multiple images of the same source are formed, though predicted much earlier, occurred only in 1979 (224). Here, we will consider weak gravitational lensing of large-scale perturbations, known as cosmic shear. Other important applications of gravitational lensing for cosmology include the study of the inner regions of galaxy clusters due to lensed background galaxies (225; 226; 227; 228) and gravitational microlensing of stars in the galaxy to study the presence of dark matter substructure (229; 230; 231; 232).

Weak gravitational lensing of galaxies by large-scale density perturbations (233) was detected in 2000 (234) and is now a chief aim of a number of ongoing and future galaxy surveys (235; 236; 237). These efforts seek the lensing-induced distortions of galaxy shapes. Weak lensing of the CMB by

density perturbations was detected in correlation with large scale structure (238) and more recently using the four-point correlation function of CMB temperature maps (239). The observational signatures here are lensing-induced position-dependent departures from statistical isotropy in the two-point CMB correlation functions, or equivalently, the four-point correlation functions induced by lensing (240).

Primordial gravitational waves can likewise lens both galaxies and the CMB (241; 242; 243). The most general lensing pattern can, like the CMB polarization, be decomposed into curl and curl-free parts (244). Since density perturbations produce (to linear order in the deflection angle) no curl in the lensing pattern, the curl component provides an IGW probe.

#### 5.1.4 Previous Work with Bipolar Spherical Harmonics

Bipolar spherical harmonics (BiPoSHs) are an orthonormal basis for functions of two angles on the sphere. These functions, also used in nuclear physics to characterize the wavefunctions of two-particle systems, provide an elegant and general formalism for quantifying departures from statistical isotropy and Gaussianity (245; 246; 247). If the CMB map is Gaussian and statistically isotropic, then its statistics are specified entirely in terms of the power spectrum  $C_l$ , the expectation value of the squared magnitude of the spherical-harmonic coefficients  $a_{lm}$  for the map, and there are no correlations between different  $a_{lm}$ s. A wide variety of departures from SI and Gaussianity induce correlations between different  $a_{lm}$ s. The point of the BiPoSH formalism is to parametrize correlations between two different coefficients,  $a_{lm}$  and  $a_{l'm'}$ , that represent two different “angular-momentum” states, in terms of total angular momenta  $L$  and  $M$ .

Since their introduction as tools for cosmology in 2003 (245), bipolar spherical harmonics have been applied to the CMB in several ways. The pattern of BiPoSH coefficients due to non-standard cosmic topologies has been calculated (248), those due to anisotropy in primordial power have been tested for in WMAP data (249; 215), and model-independent departures from SI have been tested for (250; 251; 247; 252; 253; 254; 255). They have also been used to test for asymmetric beams (256) and/or other systematic artifacts in WMAP (257). BiPoSHs for polarization have been proposed to search for position-dependent rotation of the CMB polarization (258; 259; 260; 261).

In this chapter, we discuss another application of BiPoSHs. The majority of physical effects considered so far induce only even-parity BiPoSH coefficients, and thus previous work neglected odd-parity BiPoSH modes (257) or disregarded BiPoSH parity altogether. In this chapter we show that odd-parity BiPoSH modes can provide probes of both cosmological effects and systematic artifacts that would remain elusive with the even-parity BiPoSHs. We give a detailed treatment of the BiPoSH coefficients induced by lensing, and show that lensing by gravitational waves (GWs) can excite odd-parity BiPoSHs. We also discuss observational errors that could excite these modes, and apply BiPoSH to a new system—future 21 cm intensity maps.

## 5.2 Review of Bipolar Spherical Harmonics

### 5.2.1 Statistically Isotropic and Gaussian Maps

A CMB temperature map  $T(\hat{n})$ , as a function of position  $\hat{n}$  on the sky, can be decomposed into spherical-harmonic coefficients

$$a_{lm} = \int d^2\hat{n} T(\hat{n}) Y_{lm}^*(\hat{n}).$$

If the map is statistically isotropic and Gaussian, then the statistics can be determined entirely in terms of the power spectrum  $C_l$ , defined by

$$\langle a_{lm} a_{l'm'}^* \rangle = C_l \delta_{ll'} \delta_{mm'}, \quad (5.1)$$

where the angle brackets denote an average over all realizations, and  $\delta_{ll'}$  and  $\delta_{mm'}$  are Kronecker deltas. Eq. (5.1) states that all of the  $a_{lm}$  are uncorrelated, and Gaussianity further dictates that the probability distribution function for any  $a_{lm}$  to take on a particular value is a Gaussian distribution with variance  $C_l$ .

The spatial temperature autocorrelation function is defined to be  $C(\hat{n}, \hat{n}') \equiv \langle T(\hat{n}) T(\hat{n}') \rangle$ . Most generally it is a function of the two directions  $\hat{n}$  and  $\hat{n}'$ . However, if the map is statistically isotropic and Gaussian, then the spatial correlation function depends only on the angle  $\theta$ , given by  $\cos \theta = \hat{n} \cdot \hat{n}'$ , between the two directions. In this case,

$$C(\hat{n}, \hat{n}') = \sum_l \frac{(2l+1)}{4\pi} C_l P_l(\hat{n} \cdot \hat{n}'),$$

where  $P_l(x)$  are the Legendre polynomials.

### 5.2.2 Departures from Gaussianity/SI

Departures from Gaussianity and/or SI will induce correlations between different  $a_{lms}$ . The most general correlation between any two  $a_{lms}$  can be written,

$$\langle a_{lm} a_{l'm'}^* \rangle = C_l \delta_{ll'} \delta_{mm'} + \sum_{LM; L>0} (-1)^{m'} \langle l m l', -m' | LM \rangle A_{ll'}^{LM}, \quad (5.2)$$

where  $C_l$  is the (isotropic) power spectrum,  $\langle l m l' m' | LM \rangle$  are Clebsch-Gordan coefficients, and the  $A_{ll'}^{LM}$  are BiPoSH coefficients. The spatial two-point correlation function is then

$$C(\hat{n}, \hat{n}') = \sum_l \frac{(2l+1)}{4\pi} C_l P_l(\hat{n} \cdot \hat{n}') + \sum_{l'l''LM} A_{l'l''}^{LM} \{Y_l(\hat{n}) \otimes Y_{l'}(\hat{n}')\}_{LM}, \quad (5.3)$$

where

$$\{Y_l(\hat{n}) \otimes Y_{l'}(\hat{n}')\}_{LM} = \sum_{mm'} \langle l m l' m' | LM \rangle Y_{lm}(\hat{n}) Y_{l'm'}(\hat{n}'), \quad (5.4)$$

are the bipolar spherical harmonics (BiPoSHs). These BiPoSHs constitute a complete orthonormal basis for functions of  $\hat{n}$  and  $\hat{n}'$  in terms of total-angular-momentum states labeled by quantum numbers  $L$  and  $M$  composed of angular-momentum states with  $lm$  and  $l'm'$ ; they are an alternative to the outer product of the  $\{l, m\}$  and  $\{l', m'\}$  bases.

### 5.2.3 Odd-Parity Bipolar Spherical Harmonics

It is instructive to decompose  $A_{ll'}^{LM}$  into its odd and even parity parts,

$$A_{ll'}^{LM} = A_{ll'}^{\oplus LM} \frac{[1 + (-1)^{l+l'+L}]}{2} + A_{ll'}^{\ominus LM} \frac{[1 - (-1)^{l+l'+L}]}{2}, \quad (5.5)$$

where  $A_{ll'}^{\oplus LM}$  ( $A_{ll'}^{\ominus LM}$ ) are zero for the sum  $l+l'+L$  being odd (even). It follows from the symmetry  $C(\hat{n}, \hat{n}') = C(\hat{n}', \hat{n})$  that  $A_{ll'}^{\oplus LM}$  ( $A_{ll'}^{\ominus LM}$ ) are (anti) symmetric in  $l$  and  $l'$ . We also infer that  $[A_{ll'}^{\oplus LM}]^* = (-1)^M A_{ll'}^{\oplus L-M}$  and  $[A_{ll'}^{\ominus LM}]^* = (-1)^{M+1} A_{ll'}^{\ominus L-M}$ . Thus, odd-parity BiPoSHs vanish for  $l = l'$ . Prior literature has considered physical effects (e.g., nontrivial topologies (262), SI violation (247; 263)) that produce only  $A_{ll'}^{\oplus LM}$ , the even-parity BiPoSHs, and measurements have been carried out with WMAP data only for the  $A_{ll'}^{\oplus LM}$  (250; 264). In this chapter, we consider also the odd-parity BiPoSHs  $A_{ll'}^{\ominus LM}$ .

Estimators for the BiPoSH coefficients (both the  $\oplus$  and  $\ominus$  modes) can be constructed from a map of the CMB temperature field  $T(\hat{n})$ , as follows:

$$\widehat{A_{ll'}^{LM}} = \sum_{mm'} W_l^{-1} W_{l'}^{-1} a_{lm}^{\text{map}} a_{l'm'}^{*\text{map}} (-1)^{m'} \langle l m l' -m' | LM \rangle, \quad (5.6)$$

and this estimator has a variance, under the null hypothesis (a SI Gaussian map),

$$\left\langle \widehat{A_{ll'}^{LM}} \widehat{A_{l'l'}^{L'M'}}^* \right\rangle = \delta_{LL'} \delta_{MM'} \left[ \delta_{ll'} \delta_{l'l'} + (-1)^{l+l'+L} \delta_{ll'} \delta_{l'l'} \right] C_l^{\text{map}} C_{l'}^{\text{map}} W_l^{-2} W_{l'}^{-2}, \quad (5.7)$$

where  $a_{lm}^{\text{map}} = W_l a_{lm} + a_{lm}^{\text{noise}}$  and  $C_l^{\text{map}} = W_l^2 C_l + N_l$  are the temperature spherical-harmonic coefficients and power spectrum corrected for detector noise and finite resolution. The Gaussian detector window function, which encapsulates the effects of finite detector resolution, is given by  $W_l \equiv \exp[-l^2 \theta_{\text{FWHM}}^2 / (16 \ln 2)]$ , where  $\theta_{\text{FWHM}}$  is the full width at half maximum of the detector. The instrumental noise contribution to the temperature power spectrum is given by

$$N_l = \frac{4\pi(\text{NET})^2}{t_{\text{obs}} \sqrt{f_{\text{sky}}}},$$

where  $f_{\text{sky}}$  is the fraction of the sky observed, NET is the noise equivalent temperature of the detector, and  $t_{\text{obs}}$  is the length of time over which the CMB was observed by a particular survey. We notice that the variance in Eq. (5.7) vanishes for odd parity and  $l = l' = \bar{l} = \bar{l}'$ , which is expected given that odd-parity BiPoSHs with  $l = l'$  vanish.

The noise in any individual  $A_{ll'}^{LM}$  is large, and so a search for a statistically significant departure from zero in one or a handful of  $A_{ll'}^{LM}$ s will probably not be too effective. It is better to consider specific models and/or parameterizations for departures from SI/Gaussianity and then combine the  $A_{ll'}^{LM}$ s into a minimum-variance estimator for the SI/Gaussianity-violating parameters of those models. For example, Ref. (245; 246) considered the bipolar power spectrum  $\kappa_L \equiv \sum_{ll'M} |A_{ll'}^{LM}|^2$  as a parameterization for departures from SI. As another example, Ref. (249) combined  $A_{ll'}^{LM}$ s with  $L = 2$  and  $l' = l, l \pm 2$  to derive minimum-variance estimators for the amplitude of an inflation-induced primordial-power quadrupole of the type considered in Ref. (265).

## 5.3 Gravitational Lensing

### 5.3.1 Gradient and Curl-Type Deflections

Consider a statistically isotropic and homogeneous Gaussian temperature map  $T_g(\hat{n})$  on the sphere, where  $\hat{n}$  is a position on the sky. Now suppose that each point on the sky  $\hat{n}$  has been deflected from an original direction  $\hat{n} + \vec{\Delta}(\hat{n})$  so that the observed temperature is  $T(\hat{n}) = T_g(\hat{n} + \vec{\Delta}) \simeq T_g(\hat{n}) + \vec{\Delta} \cdot \vec{\nabla}_{\hat{\theta}} T_g(\hat{n})$ . This deflection might come about cosmologically through weak gravitational lensing or may arise as an instrumental/measurement artifact (for example, if there are pointing errors).

The most general deflection field  $\vec{\Delta}$  can be written in vector notation as

$$\vec{\Delta} = \vec{\nabla}_{\hat{\theta}} \phi(\hat{n}) + \vec{\nabla}_{\hat{\theta}} \times \Omega(\hat{n}), \quad (5.8)$$

or in component notation,  $\Delta_i = (\nabla_{\hat{\theta}})_i \phi(\hat{n}) + \epsilon_{ij} (\nabla_{\hat{\theta}})_j \Omega(\hat{n})$ ,<sup>2</sup> in terms of two scalar functions  $\phi(\hat{n})$  and  $\Omega(\hat{n})$  on the sphere, where  $\vec{\nabla}_{\hat{\theta}}$  is the angular covariant derivative on the unit sphere. In other words, the most general vector field on a two-sphere can be written as the gradient of some scalar field  $\phi(\hat{n})$  plus the curl of some other field  $\Omega(\hat{n})$ . Weak gravitational lensing by density perturbations gives rise, at linear order in the lensing potential, only to the gradient component. A curl component can arise cosmologically from second-order terms in the deflection field or from lensing by GWs. Systematic measurement effects may conceivably give rise to both types of deflections.

We now show that the  $A_{ll'}^{\oplus LM}$  and  $A_{ll'}^{\ominus LM}$  BiPoSH coefficients are induced, respectively, by the

<sup>2</sup>Here, the Levi-Civita symbol on the unit sphere can be defined in terms of its three-dimensional equivalent as  $\epsilon_{ij} = -\epsilon_{ijk} r_k$ . The choice of sign here can be understood as the choice to have the spherical polar coordinates  $(\theta, \phi)$  form a right-handed coordinate system on the sky, since it will ensure that the basis vectors satisfy  $\hat{e}_{\theta} \times \hat{e}_{\phi} = 1$ .

gradient and curl components of the deflection field. The change in the temperature moments induced by lensing is (at first order in  $\phi$  and  $\Omega$ ),

$$\begin{aligned}\delta a_{lm} &= \int d^2\hat{n} Y_{lm}^*(\hat{n}) \left\{ \left[ \vec{\nabla}_{\hat{\theta}} \phi \right] \cdot \left[ \vec{\nabla}_{\hat{\theta}} T(\hat{n}) \right] + \left[ \vec{\nabla}_{\hat{\theta}} \Omega(\hat{n}) \right] \times \left[ \vec{\nabla}_{\hat{\theta}} T(\hat{n}) \right] \right\} \\ &= \sum_{LM; L>0} \sum_{l'm'} a_{l'm'} \int d^2\hat{n} Y_{lm}^*(\hat{n}) \left\{ \phi_{LM} \left[ \vec{\nabla}_{\hat{\theta}} Y_{LM}(\hat{n}) \right] \cdot \left[ \vec{\nabla}_{\hat{\theta}} Y_{l'm'}(\hat{n}) \right] + \Omega_{LM} \left[ \vec{\nabla}_{\hat{\theta}} Y_{LM}(\hat{n}) \right] \times \left[ \vec{\nabla}_{\hat{\theta}} Y_{l'm'}(\hat{n}) \right] \right\},\end{aligned}$$

where in the second line we have decomposed

$$\phi(\hat{n}) = \sum_{L=1}^{\infty} \sum_{M=-L}^L Y_{LM}(\hat{n}) \phi_{LM}, \quad (5.9)$$

and similarly for  $\Omega(\hat{n})$ . We do not consider  $L = 0$  modes of  $\phi$  and  $\Omega$  since they would not cause a deflection. In the notation of Ref. (266),

$$\vec{\nabla}_{\hat{\theta}} Y_{lm} = \sqrt{\frac{l(l+1)}{2}} [{}_1Y_{lm} \hat{m}_+ - {}_{-1}Y_{lm} \hat{m}_-],$$

where  ${}_1Y_{lm}$  and  ${}_{-1}Y_{lm}$  are spin-weighted spherical harmonics, the null coordinates  $\hat{m}_{\pm} = (\hat{e}_{\theta} \mp i\hat{e}_{\phi})/\sqrt{2}$ , and the only non-trivial products of the null coordinates are  $\hat{m}_+ \cdot \hat{m}_- = 1$ , and  $\hat{m}_+ \times \hat{m}_- = i$ .

Thus, it is obtained that

$$\begin{aligned}\left( \vec{\nabla}_{\hat{\theta}} Y_{LM} \right) \cdot \left( \vec{\nabla}_{\hat{\theta}} Y_{l'm'} \right) &= -\frac{\sqrt{L(L+1)l'(l'+1)}}{2} [({}_1Y_{LM}) ({}_{-1}Y_{l'm'}) + ({}_{-1}Y_{LM}) ({}_1Y_{l'm'})], \\ \left( \vec{\nabla}_{\hat{\theta}} Y_{LM} \right) \times \left( \vec{\nabla}_{\hat{\theta}} Y_{l'm'} \right) &= -\frac{i\sqrt{L(L+1)l'(l'+1)}}{2} [({}_1Y_{LM}) ({}_{-1}Y_{l'm'}) - ({}_{-1}Y_{LM}) ({}_1Y_{l'm'})].\end{aligned}$$

Using the triple integral (266) of spin-weighted spherical harmonics, the  $\delta a_{lm}$  for the gradient and curl terms are obtained as

$$\delta a_{lm} = \sum_{LM; L>0} \sum_{l'm'} \frac{(-1)^{M+m} a_{l'm'} G_{ll'}^L}{\sqrt{(2L+1)l(l+1)}} \left[ \phi_{LM} \frac{[1 + (-1)^{l+l'+L}]}{2} - i\Omega_{LM} \frac{[1 - (-1)^{l+l'+L}]}{2} \right] \langle l m l', -m' | LM \rangle,$$

where

$$G_{ll'}^L \equiv \sqrt{\frac{L(L+1)l(l+1)l'(l'+1)(2l+1)(2l'+1)}{4\pi}} \langle l 0 l' 1 | L 1 \rangle.$$

Up to linear order in the deflection coefficients  $\phi_{LM}$  and  $\Omega_{LM}$ , the even- and odd-parity BiPoSH coefficients are then,

$$A_{ll'}^{\oplus LM} = \frac{\phi_{LM}}{\sqrt{2L+1}} \left[ \frac{C_l G_{ll'}^L}{\sqrt{l'(l'+1)}} + \frac{C_{l'} G_{ll'}^L}{\sqrt{l(l+1)}} \right] = Q_{ll'}^{\oplus L} \phi_{LM}, \quad (5.10)$$

$$A_{ll'}^{\ominus LM} = \frac{i\Omega_{LM}}{\sqrt{2L+1}} \left[ \frac{C_l G_{ll'}^L}{\sqrt{l'(l'+1)}} - \frac{C_{l'} G_{ll'}^L}{\sqrt{l(l+1)}} \right] = Q_{ll'}^{\ominus L} \Omega_{LM}, \quad (5.11)$$

where we have defined the quantities

$$\begin{aligned} Q_{l'l'}^{\oplus L} &= \frac{1}{\sqrt{2L+1}} \left[ \frac{C_l G_{l'l}^L}{\sqrt{l'(l'+1)}} + \frac{C_{l'} G_{ll'}^L}{\sqrt{l(l+1)}} \right], \\ Q_{l'l'}^{\ominus L} &= \frac{i}{\sqrt{2L+1}} \left[ \frac{C_l G_{l'l}^L}{\sqrt{l'(l'+1)}} - \frac{C_{l'} G_{ll'}^L}{\sqrt{l(l+1)}} \right]. \end{aligned} \quad (5.12)$$

Clearly, the gradient part contributes only to  $A_{l'l'}^{\oplus LM}$  and the curl part only to  $A_{l'l'}^{\ominus LM}$ . Further, it is explicit that the gradient and curl parts of the deflection correspond, respectively, to the symmetric and antisymmetric (in  $\{l'l'\}$ ) parts of the total  $A_{l'l'}^{LM}$ .

Suppose the  $A_{l'l'}^{LM}$ s have been measured using the estimators in Eq. (5.6). If we then assume that lensing is the dominant source of BiPoSHs we can use Eqs. (5.7), (5.10), and (5.11) to construct maximum-likelihood estimators for the gradient and curl components of the deflection field,

$$\widehat{\phi}_{LM} = \frac{\sum_{l'l'} Q_{l'l'}^{\oplus L*} \widehat{A}_{l'l'}^{\oplus LM} / (W_l^{-2} W_{l'}^{-2} C_l^{\text{map}} C_{l'}^{\text{map}})}{\sum_{l'l'} |Q_{l'l'}^{\oplus L}|^2 / (W_l^{-2} W_{l'}^{-2} C_l^{\text{map}} C_{l'}^{\text{map}})}, \quad (5.13)$$

$$\widehat{\Omega}_{LM} = \frac{\sum_{l'l'} Q_{l'l'}^{\ominus L*} \widehat{A}_{l'l'}^{\ominus LM} / (W_l^{-2} W_{l'}^{-2} C_l^{\text{map}} C_{l'}^{\text{map}})}{\sum_{l'l'} |Q_{l'l'}^{\ominus L}|^2 / (W_l^{-2} W_{l'}^{-2} C_l^{\text{map}} C_{l'}^{\text{map}})}. \quad (5.14)$$

The variance of these estimators, under the null hypothesis of no lensing, is given by

$$\langle \widehat{\phi}_{LM} \widehat{\phi}_{L'M'}^* \rangle \equiv \delta_{LL'} \delta_{MM'} (\sigma_L^\phi)^2 \equiv 2 \delta_{LL'} \delta_{MM'} \left[ \sum_{l'l'} |Q_{l'l'}^{\oplus L}|^2 / (W_l^{-2} W_{l'}^{-2} C_l^{\text{map}} C_{l'}^{\text{map}}) \right]^{-1}, \quad (5.15)$$

$$\langle \widehat{\Omega}_{LM} \widehat{\Omega}_{L'M'}^* \rangle \equiv \delta_{LL'} \delta_{MM'} (\sigma_L^\Omega)^2 \equiv 2 \delta_{LL'} \delta_{MM'} \left[ \sum_{l'l'} |Q_{l'l'}^{\ominus L}|^2 / (W_l^{-2} W_{l'}^{-2} C_l^{\text{map}} C_{l'}^{\text{map}}) \right]^{-1}, \quad (5.16)$$

where the sums in Eqs. (5.13) and (5.15) only include pairs of  $l, l'$  for which  $l+l'+L$  is even, while those in Eqs. (5.14) and (5.16) only include pairs for which this quantity is odd.

### 5.3.2 Deflection Field from Metric Perturbations

Cosmic shear, weak gravitational lensing due to density perturbations or GWs along the line of sight to the CMB, will produce displacements like those in Eq. (5.8). Our goal here will be to calculate the displacement spherical-harmonic coefficients  $\phi_{LM}$  and  $\Omega_{LM}$  that arise from gravitational lensing due to density perturbations and GWs. There is a vast literature on lensing by density perturbations and also specifically on lensing of the CMB by density perturbations (210). Our density-perturbation

results follow most closely those of Refs. (267; 266). Lensing by GWs has been considered in Ref. (268). We follow primarily the approach of Refs. (269; 270), who calculated  $\Omega_{LM}$  due to GWs, but extend their results to include  $\phi_{LM}$  from GWs, reproducing the results of Ref. (271). We make use in this Section of relevant work on lensing and/or differential analysis on the celestial sphere in Refs. (266; 267; 272; 273).

We write the metric for the perturbed spacetime as

$$ds^2 = a^2(\eta) [-d\eta^2 + (\delta_{ij} + h_{ij}) dx^i dx^j],$$

where  $h_{ij}$  is the metric perturbation in the synchronous gauge, and  $\eta$  is the conformal time. Now consider a photon that we observe to come from the direction  $\hat{n}$  on the sky. In the absence of perturbations, this photon travels along a path  $\vec{x}(\eta) = (\eta_0 - \eta) \hat{n}$  as a function of conformal time  $\eta$ , where  $\eta_0$  is the conformal time today. Metric perturbations will induce perturbations in this trajectory, which we can calculate by integrating the geodesic equation back over the photon path to find the direction of propagation of the photon when it was emitted at a conformal time  $\eta$ . To first order in the metric perturbation  $h$ , we find the original direction of propagation of the photon on the sky to be  $\hat{n} + \vec{\Delta}$ , where (274)

$$\Delta^i(\hat{n}) = \frac{P_{im}}{\eta_0 - \eta} \int_{\eta_0}^{\eta} d\eta' \left[ h_{mj} \hat{n}_j - \frac{1}{2} (\eta' - \eta) \hat{n}_k \hat{n}_l \partial_m h_{kl} \right]_{[\eta', (\eta_0 - \eta') \hat{n}]}. \quad (5.17)$$

Here, we have ignored the observer terms  $h_{ij}(\eta_0)$ , and we have defined the projection tensor  $P_{im} = \delta_{im} - n_i n_m$  onto the space perpendicular to the unit vector  $\hat{n}$ . The subscript indicates that the quantities in the integral are evaluated at time and space coordinates  $(\eta, \vec{x}) = (\eta', (\eta_0 - \eta') \hat{n})$ ; i.e., they are evaluated along the unperturbed path of the photon. In our case, the source is the CMB, and  $\eta = \eta_{\text{ls}}$  is the conformal time at the surface of last scatter. However, the calculation could also be applied to the lensing of galaxies in which case the relevant conformal time would be that corresponding to redshifts  $z \sim 1$ .

The functions  $\phi(\hat{n})$  and  $\Omega(\hat{n})$  in the decomposition in Eq. (5.8) can be obtained from

$$\nabla_{\hat{\theta}}^2 \phi(\hat{n}) = \vec{\nabla}_{\hat{\theta}} \cdot \vec{\Delta}(\hat{n}), \quad \nabla_{\hat{\theta}}^2 \Omega(\hat{n}) = -\vec{\nabla}_{\hat{\theta}} \times \vec{\Delta}(\hat{n}), \quad (5.18)$$

where as before  $\vec{\nabla}_{\hat{\theta}}$  is the angular covariant derivative on the unit sphere. As Ref. [25] notes, the standard lensing convergence is  $\kappa = -(1/2) \nabla_{\hat{\theta}}^2 \phi$  and the lensing rotation is  $\omega = (1/2) \nabla_{\hat{\theta}}^2 \Omega$ .

The gradient component is obtained from

$$\begin{aligned} \nabla_{\vec{\theta}}^2 \phi(\hat{n}) = \vec{\nabla}_{\vec{\theta}} \cdot \vec{\Delta} = & -\frac{1}{\eta_0 - \eta} \left\{ \int_{\eta}^{\eta_0} d\eta' (\eta_0 - \eta') (\delta_{ik} - \hat{n}_i \hat{n}_k) \left[ -\partial_k (h_{ij} n^j) + \frac{1}{2} (\eta' - \eta) \partial_i \partial_k (h_{lm} \hat{n}_l \hat{n}_m) \right] \right. \\ & \left. + \int_{\eta}^{\eta_0} d\eta' [3\hat{n}_i \hat{n}_j h_{ij} - h_{ii} + (\eta' - \eta) (\hat{n}_j \partial_i h_{ij} - 2\hat{n}_i \hat{n}_j \hat{n}_k \partial_k h_{ij})]_{[\eta', (\eta_0 - \eta') \hat{n}]} \right\}, \end{aligned} \quad (5.19)$$

where we have used the fact that  $\vec{\nabla}_{\vec{\theta}}$ , which acts on the unit vector  $\hat{n}$ , behaves as  $\nabla_{\vec{\theta}}^i = (\eta_0 - \eta') (\delta_{ik} - \hat{n}_i \hat{n}_k) \partial_k$  inside the integral due to the dependence of  $\vec{x}$  on  $\hat{n}$  as defined in the integrand subscript.

Let us now consider the curl component. For this calculation we must use  $\nabla_{\vec{\theta}}^2 \Omega = -\vec{\nabla}_{\vec{\theta}} \times \vec{\Delta}$  and then note that, as before,  $\nabla_{\vec{\theta}}^i = (\eta_0 - \eta') (\delta_{ik} - \hat{n}_i \hat{n}_k) \partial_k$  inside the integrand. Applying this to Eq. (5.17), we have (269)

$$\nabla_{\vec{\theta}}^2 \Omega(\hat{n}) = - \int_{\eta}^{\eta_0} d\eta' (n_i n_l \epsilon_{ijk} \partial_j h_{kl})_{[\eta', \hat{n}(\eta_0 - \eta')]} \quad (5.20)$$

### 5.3.3 Lensing by Density (Scalar Metric) Perturbations

Let us first consider scalar perturbations. In the conformal-Newtonian gauge in the absence of anisotropic stresses, the metric is given by

$$ds^2 = a^2(\eta) [-(1 - 2\Phi)d\eta^2 + (1 + 2\Phi)\delta_{ij}dx^i dx^j].$$

Noting that a conformal transformation preserves null geodesics, our calculations of the photon path will be unaffected if we work in a synchronous metric obtained from the conformal-Newtonian form through multiplication by  $(1 + 2\Phi)$ . Assuming that  $\Phi$  is small and keeping terms only to linear order, we find the conformally related metric,

$$ds^2 = a^2(\eta) [-d\eta^2 + (1 + 4\Phi)\delta_{ij}dx^i dx^j].$$

Using this metric perturbation  $h_{ij} = 4\Phi\delta_{ij}$  in Eq. (5.19) above, we find that the first, third, and fourth terms vanish, giving for the gradient-type lensing caused by scalar perturbations,

$$\nabla_{\vec{\theta}}^2 \phi^{\text{sca}}(\hat{n}) = -\frac{2}{\eta_0 - \eta} \int_{\eta}^{\eta_0} d\eta' (\eta' - \eta) [(\delta_{ij} - \hat{n}_i \hat{n}_j) (\eta_0 - \eta') \partial_i \partial_j \Phi - 2\hat{n}_i \partial_i \Phi].$$

For small-scale fluctuations, the second term will be negligible compared with the first, so it can be dropped. We can rewrite the spatial derivatives in terms of  $\vec{\nabla}_{\vec{\theta}}$  to find

$$\nabla_{\vec{\theta}}^2 \phi^{\text{sca}}(\hat{n}) = -\frac{2}{\eta_0 - \eta} \int_{\eta}^{\eta_0} d\eta' \frac{\eta' - \eta}{\eta_0 - \eta'} \nabla_{\vec{\theta}}^2 \Phi(\eta', (\eta_0 - \eta') \hat{n}),$$

and we can remove the angular derivatives to obtain the usual expression for the projected potential

$$\phi^{\text{sca}}(\hat{n}) = -2 \int_{\eta}^{\eta_0} d\eta' \frac{\eta' - \eta}{(\eta_0 - \eta)(\eta_0 - \eta')} \Phi(\eta', (\eta_0 - \eta')\hat{n}).$$

We can once again decompose  $\phi(\hat{n})$  in terms of its spherical-harmonic coefficients as in Eq. (5.9).

We then find

$$\begin{aligned} \phi_{LM}^{\text{sca}} &\equiv \int d^2\hat{n} Y_{LM}^*(\hat{n}) \phi^{\text{sca}}(\hat{n}) \\ &= -2 \int_{\eta}^{\eta_0} d\eta' \frac{\eta' - \eta}{(\eta_0 - \eta)(\eta_0 - \eta')} \int d^2\hat{n} Y_{LM}^*(\hat{n}) \Phi(\eta', (\eta_0 - \eta')\hat{n}). \end{aligned} \quad (5.21)$$

Thus, lensing by density perturbations with a given projected potential is characterized by nonzero even bipolar spherical harmonics  $A_{l'l}^{\oplus LM}$  given by Eq. (5.10) with  $\phi_{LM}$  given by  $\phi_{LM}^{\text{sca}}$  above. Scalar perturbations cause no curl-type lensing, which we can see in several ways. For scalar perturbations,  $h_{ij} \propto \Phi \delta_{ij}$ , and so the left-hand side of Eq. (5.20) vanishes. Then, by taking a Laplacian of the mode expansion  $\Omega_{LM} = \int d^2\hat{n} \Omega(\hat{n}) Y_{LM}^*(\hat{n})$ , and noting that the spherical harmonics are eigenfunctions of the Laplacian with eigenvalue  $L(L+1)$ , we can write

$$\Omega_{LM} = \frac{1}{L(L+1)} \int d^2\hat{n} Y_{LM}^*(\hat{n}) \nabla_{\hat{\theta}}^2 \Omega(\hat{n}). \quad (5.22)$$

Thus, we find that all of the  $\Omega_{LM}^{\text{sca}}$ , except possibly for the unphysical  $L=0$  mode, vanish. Equivalently, an argument can be made that scalar perturbations have no preferred direction, and so could not generate curl-modes, which do have a preferred direction. Thus, scalar modes produce no odd bipolar spherical harmonics  $A_{l'l}^{\ominus LM}$ .

We can go on to find the autocorrelation power spectrum of the  $\phi_{LM}^{\text{sca}}$ . Starting from Eq. (5.21), we use the fact that the potential perturbations  $\Phi(\eta, \vec{k})$  today are related to their primordial values  $\Phi_P(\vec{k})$  by

$$\Phi(\eta, \vec{k}) = \frac{9}{10} \Phi_P(\vec{k}) T^{\text{sca}}(k) \frac{D_1(\eta)}{a(\eta)},$$

where  $a(\eta)$  is the scale factor,  $T^{\text{sca}}(k)$  is the scalar transfer function that describes the evolution of scalar modes through the epochs of horizon crossing and matter-radiation equality, and  $D_1(\eta)$  is the growth function that captures the scale-independent evolution of scalar modes at later times (76). The transfer function can be approximated using the fitting form of Ref. (275),

$$T^{\text{sca}}(x \equiv k/k_{\text{eq}}) = \frac{\ln(1 + 0.17x)}{0.171x} [1 + 0.284x + (1.18x)^2 + (0.399x)^3 + (0.490x)^4]^{-0.25},$$

where  $k_{\text{eq}}$  is the wavenumber of the mode that crossed the horizon at matter-radiation equality,

defined as  $k_{\text{eq}} \equiv a_{\text{eq}} H(a_{\text{eq}}) = \sqrt{2} H_0 a_{\text{eq}}^{-1/2}$ . We can write the growth function, under the assumption of cosmological-constant dark energy, as

$$D_1(\eta) = \frac{5\Omega_m}{2} \frac{H(\eta)}{H_0} \int_o^{a(\eta)} \frac{da'}{(a' H(a')/H_0)^3}.$$

We also write the autocorrelation of the primordial scalar fluctuations  $\langle \Phi_P(\vec{k}) \Phi_P^*(\vec{k}') \rangle = (2\pi)^3 \delta^3(\vec{k} - \vec{k}') P_\Phi(k)$ , where the primordial power spectrum is given by

$$P_\Phi(k) = \frac{50\pi^2}{9k^3} \left(\frac{k}{H_0}\right)^{n_s-1} \Delta_R^2 \left(\frac{\Omega_m}{D_1(a=1)}\right)^2.$$

With these ingredients, and after using the partial-wave decomposition,

$$e^{ik(\eta_0-\eta')\cos\theta} = \sum_{L=0}^{\infty} i^L (2L+1) j_L(k(\eta_0-\eta')) P_L(\cos\theta), \quad (5.23)$$

we find the autocorrelation power spectrum to be

$$C_L^{\phi\phi\text{sca}} = \frac{2}{\pi} \left[ \frac{9}{5(\eta-\eta_0)} \right]^2 \int dk k^2 P_\Phi(k) T^{\text{sca}}(k)^2 \left\{ \int_\eta^{\eta_0} d\eta' \frac{(\eta'-\eta)}{(\eta_0-\eta')} \frac{D_1(\eta')}{a(\eta')} j_L[(\eta_0-\eta')k] \right\}^2. \quad (5.24)$$

To calculate the magnitude and shape of this autocorrelation function, we employ the WMAP 7-year cosmological parameters of Ref. (276). We plot the result of our calculation in green squares in Figure 5.1.

### 5.3.4 Lensing by GWs (Tensor Metric Perturbations)

If the metric perturbation  $h_{ij}$  is caused by GWs, we can decompose it into plane waves,

$$h_{ij}(\vec{x}, \eta) = \int \frac{d^3k}{(2\pi)^3} e^{i\vec{k}\cdot\vec{x}} T(k, \eta) \sum_{\alpha=+, \times} h^\alpha(\vec{k}) \epsilon_{ij}^\alpha(\vec{k}), \quad (5.25)$$

where we sum over the two GW polarizations  $+$  and  $\times$ , the plane-wave amplitudes are  $h^\alpha(\vec{k})$ , and  $\epsilon_{ij}^\alpha$  are the polarization tensors, which are transverse, traceless matrices. Here,  $T(k, \eta)$  is the GW transfer function, which gives the conformal-time evolution of the mode; Ref. (269) notes that it is well approximated by  $T(k, \eta) = 3j_1(k\eta)/(k\eta)$ .

Now consider a single GW mode propagating in the  $\hat{z}$  direction with wavenumber  $k$  and  $+$  polarization. In this case the polarization tensor is

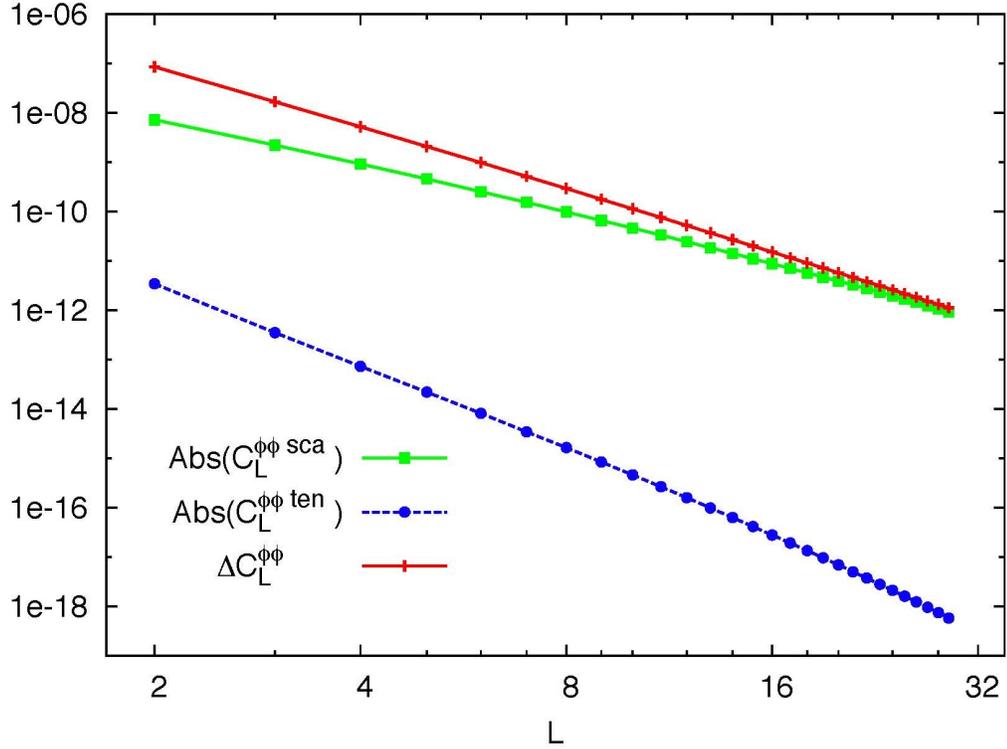


Figure 5.1: Here we plot the autocorrelation power spectrum  $C_L^{\phi\phi}$  of the gradient-type  $\phi$  modes of cosmic shear. In green squares we show the autocorrelation of the  $\phi$  modes from lensing by scalar perturbations, and in blue circles that of the  $\phi$  modes induced by tensor perturbations. We use the WMAP-7 cosmological parameters, and assume the maximum allowable tensor-to-scalar ratio  $r = 0.24$  from the WMAP-7 data combined with BAO and the  $H_0$  measurement (276), to calculate the tensor contribution. The error with which these power spectra could be measured using the parameters of the Planck satellite is shown as red +s.

$$\epsilon_{ij}^+(k\hat{z}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

The only nonzero metric-perturbation components are then  $h_{xx} = -h_{yy} = h^+(\vec{k}) e^{ikz} T(k, \eta)$ . The unit vector  $\hat{n} = (\sin\theta \cos\varphi, \sin\theta \sin\varphi, \cos\theta)$ . The curl component of lensing of the CMB by tensor perturbations is then

$$\nabla_{\hat{\theta}}^2 \Omega^{\text{ten}}(\hat{n}) = ik h^+(\vec{k}) \sin^2\theta \sin 2\varphi \int_{\eta}^{\eta_0} d\eta' T(k, \eta') e^{ik(\eta_0 - \eta') \cos\theta}. \quad (5.26)$$

A GW with the  $\times$  polarization is the same as that with the  $+$  polarization, but rotated by  $45^\circ$  to the right. The  $\Omega^{\text{ten}}(\hat{n})$  pattern is therefore the same, but with  $\sin 2\varphi$  replaced by  $-\cos 2\varphi$ . We thus see that *lensing by GWs will give rise to nonvanishing  $A_W^{\ominus LM}$* .

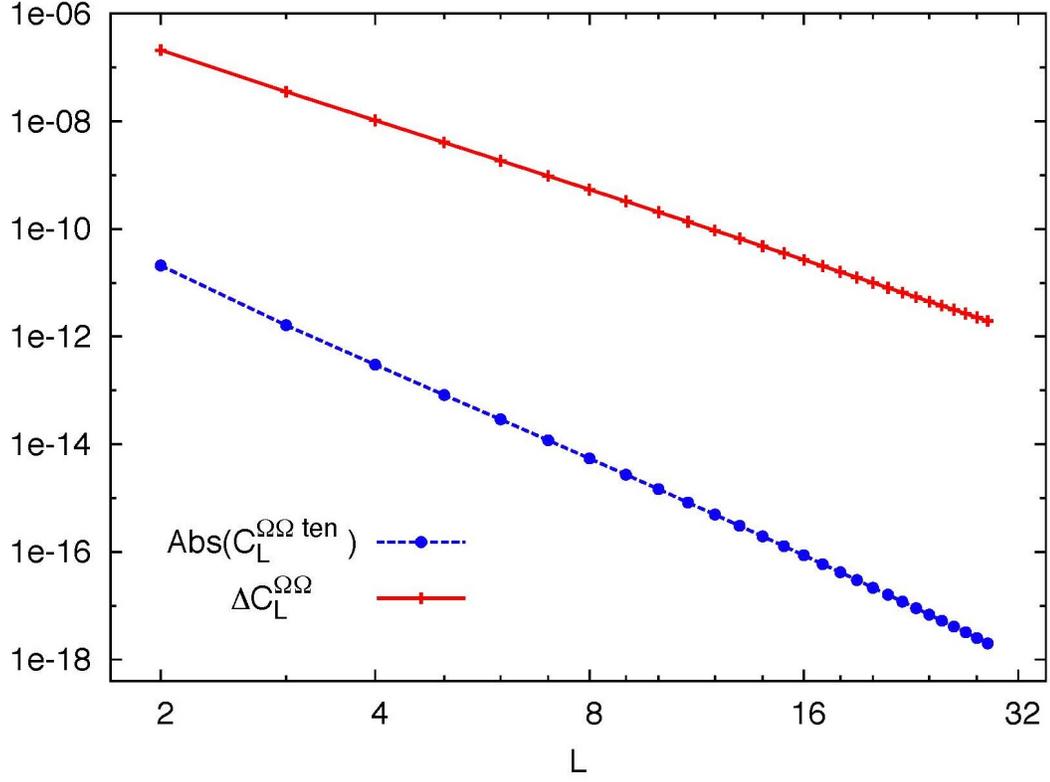


Figure 5.2: Here we plot the autocorrelation power spectrum  $C_L^{\Omega\Omega}$  of the curl-type  $\Omega$  modes of the weak lensing of the CMB temperature field. These modes can only be induced by tensor perturbations. We show the signal in blue circles and the error with which they could be measured using the parameters of the Planck satellite as red +s.

The gradient component of cosmic shear due to tensor perturbations is a bit more complicated; it is

$$\begin{aligned} \nabla_{\hat{\theta}}^2 \phi^{\text{ten}}(\hat{n}) = & -\frac{h^+(\vec{k})}{\eta_0 - \eta} \sin^2 \theta \cos 2\varphi \int_{\eta}^{\eta_0} d\eta' T(k, \eta') \left\{ 3 - 2ik(\eta' - \eta) \cos \theta \right. \\ & \left. + (\eta_0 - \eta') \left[ ik \cos \theta - \frac{k^2}{2} (\eta' - \eta) \sin^2 \theta \right] \right\} e^{ik(\eta_0 - \eta') \cos \theta}. \end{aligned} \quad (5.27)$$

This can be further simplified by noting that

$$-ik \cos \theta e^{ik(\eta_0 - \eta') \cos \theta} = \frac{\partial}{\partial \eta'} e^{ik(\eta_0 - \eta') \cos \theta},$$

which then leads to

$$\begin{aligned} \nabla_{\theta}^2 \phi^{\text{ten}}(\hat{n}) = & -\frac{h^+(\vec{k})}{\eta_0 - \eta} \sin^2 \theta \cos 2\varphi \int_{\eta}^{\eta_0} d\eta' T(k, \eta') \left\{ 3 + 2(\eta' - \eta) \frac{\partial}{\partial \eta'} \right. \\ & \left. - (\eta_0 - \eta') \left[ \frac{\partial}{\partial \eta'} + \frac{(\eta' - \eta)}{2} \left( k^2 + \frac{\partial^2}{\partial \eta'^2} \right) \right] \right\} e^{ik(\eta_0 - \eta') \cos \theta}. \end{aligned} \quad (5.28)$$

For the  $\times$  polarization, we replace  $\cos 2\varphi$  by  $\sin 2\varphi$ .

Note that the expressions for  $\nabla_{\theta}^2 \phi^{\text{ten}}$  and  $\nabla_{\theta}^2 \Omega^{\text{ten}}$  differ only in two ways: (1) The curl mode has a  $\sin 2\varphi$  dependence on the azimuthal angle  $\varphi$ , while the scalar mode has a  $\cos 2\varphi$  dependence (for the  $+$  polarization). (2) The  $\eta'$  dependences of the two integrands differ.

We now find the spherical-harmonic coefficients  $\phi_{LM}^{\text{ten}} = \int d^2 \hat{n} \phi^{\text{ten}}(\hat{n}) Y_{LM}^*(\hat{n})$  and  $\Omega_{LM}^{\text{ten}} = \int d^2 \hat{n} \Omega^{\text{ten}}(\hat{n}) Y_{LM}^*(\hat{n})$ . Taking the angular derivatives of this decomposition of the curl component, we find the result Eq. (5.22). We also expand these coefficients in terms of their polarization and  $\vec{k}$  modes,

$$\Omega_{LM}^{\text{ten}} = \int \frac{d^3 k}{(2\pi)^3} \sum_{\alpha=+, \times} \Omega_{LM}^{\text{ten} \alpha}(\vec{k}). \quad (5.29)$$

If we consider just one mode, with  $\alpha = +$  and  $\vec{k} = k\hat{z}$ , and use Eq. (5.26), its amplitude simplifies into an angular and a conformal time integral:

$$\Omega_{LM}^{\text{ten}+}(k\hat{z}) = -\frac{ik h^+(\vec{k})}{L(L+1)} \int_{\eta}^{\eta_0} d\eta' T(k, \eta') \int d^2 \hat{n} Y_{LM}^*(\hat{n}) \sin^2 \theta \sin(2\phi) e^{ik(\eta_0 - \eta') \cos \theta}.$$

The azimuthal integral is easily taken once the spherical harmonic is decomposed as

$$Y_{LM}^*(\hat{n}) = \sqrt{\frac{2L+1}{4\pi} \frac{(L-M)!}{(L+M)!}} e^{-iM\phi} P_{LM}(\cos \theta),$$

and yields the result that only  $M = \pm 2$  modes remain. The polar integral can then be taken by using the partial-wave decomposition Eq. (5.23) and by converting associated Legendre polynomials into regular Legendre polynomials and using their orthogonality. The final result that we obtain for the spherical-harmonic coefficients of the curl mode is

$$\Omega_{LM}^{\text{ten}+}(k\hat{z}) = i^L h^+(\vec{k}) (\delta_{M,2} - \delta_{M,-2}) \sqrt{\frac{2L+1}{2}} F_L^{\Omega}(k), \quad (5.30)$$

where

$$F_L^{\Omega}(k) = \sqrt{\frac{2\pi(L+2)!}{(L-2)!}} \frac{1}{L(L+1)} \int_{k\eta}^{k\eta_0} dw T(w) \frac{j_L(k\eta_0 - w)}{(k\eta_0 - w)^2} \quad (5.31)$$

is a transfer function for  $\Omega$ . Note that in writing Eq. (5.31) we have assumed that  $T(k, \eta) = T(k\eta)$ , and that for the  $\times$  polarization the  $\sin 2\varphi$  dependence of  $\Omega(\hat{n})$  is replaced by  $-\cos 2\varphi$ , so that the factor  $(\delta_{M,2} - \delta_{M,-2})$  is replaced by  $-i(\delta_{M,2} + \delta_{M,-2})$ .

Likewise, noting the similarities between Eqs. (5.26) and (5.28), and decomposing  $\phi_{LM}^{\text{ten}}$  into modes as in Eq. (5.29)

$$\phi_{LM}^{\text{ten}} = \int \frac{d^3k}{(2\pi)^3} \sum_{\alpha=+,\times} \phi_{LM}^{\text{ten}\alpha}(\vec{k}), \quad (5.32)$$

the result for the amplitude of the gradient mode with  $\alpha = +$  and  $\vec{k} = k\hat{z}$  is

$$\phi_{LM}^{\text{ten}+}(k\hat{z}) = i^L h^+(\vec{k})(\delta_{M,2} + \delta_{M,-2}) \sqrt{\frac{2L+1}{2}} F_L^\phi(k), \quad (5.33)$$

where

$$F_L^\phi(k) = -\sqrt{\frac{2\pi(L+2)!}{(L-2)!}} \frac{1}{L(L+1)} \int_{k\eta}^{k\eta_0} dw \frac{k\eta_0 - w}{k(\eta_0 - \eta)} T(w) \left[ \frac{\partial}{\partial w} + \frac{1}{2}(w - k\eta) \left( 1 + \frac{\partial^2}{\partial w^2} \right) \right] \frac{j_L(k\eta_0 - w)}{(k\eta_0 - w)^2}. \quad (5.34)$$

Again, the factor  $(\delta_{M,2} + \delta_{M,-2})$  is replaced by  $-i(\delta_{M,2} - \delta_{M,-2})$  for the  $\times$  polarization.

The contributions from this Fourier mode to the  $\phi^{\text{ten}}$  and  $\Omega^{\text{ten}}$  power spectra are  $C_L^{\phi\phi\text{ten}}(k\hat{z})_+ = \sum_M \langle |\phi_{LM}^{\text{ten}}|^2 \rangle / (2L+1)$  and  $C_L^{\Omega\Omega\text{ten}}(k\hat{z})_+ = \sum_M \langle |\Omega_{LM}^{\text{ten}}|^2 \rangle / (2L+1)$ . Note that it is only the  $M = \pm 2$  modes that contribute. By rotational invariance, the contribution from the  $\times$  polarization is the same, as is the contribution from any other mode with the same wavenumber  $k$  but pointing in a different direction. If the gravitational waves have power spectrum  $P_T(k)$ , defined by

$$\langle h_{\vec{k}}^i (h_{\vec{k}'}^j)^* \rangle = (2\pi)^3 \delta_D(\vec{k} - \vec{k}') \delta_{ij} P_T(k), \quad (5.35)$$

(with  $\{i, j\} = \{\times, +\}$ ), then the  $\phi$  and  $\Omega$  power spectra are

$$C_L^{XX\text{ten}} = 2 \int \frac{d^3k}{(2\pi)^3} P_T(k) [F_L^X(k)]^2 \quad (5.36)$$

for  $X = \{\phi, \Omega\}$ . In this chapter, we will assume a scale-invariant power spectrum

$$P_T(k) = \frac{\pi^2}{2k^3} \Delta_R^2 r, \quad (5.37)$$

where we have neglected the spectral tilt and adopt the parameters of WMAP7 (276).

We calculate the variance in the measurement of these autocorrelation functions from an observed CMB temperature map, under the null hypothesis of no GWs, and obtain an expression in terms of the variance of the  $\phi$  and  $\Omega$  estimators, Eqs. (5.15) and (5.16)

$$\Delta C_L^{\phi\phi} = \sqrt{\frac{2}{2L+1}} \left( \sigma_L^{\phi^2} + C_L^{\phi\phi \text{ sca}} \right), \quad (5.38)$$

$$\Delta C_L^{\Omega\Omega} = \sqrt{\frac{2}{2L+1}} \sigma_L^{\Omega^2}. \quad (5.39)$$

Here,  $\sigma_L^{\phi^2}$  and  $\sigma_L^{\Omega^2}$  are the variances of our estimators for  $\phi_{LM}$  and  $\Omega_{LM}$  as found in Eqs. (5.15) and (5.16). Note that under the null hypothesis of no GWs, there is no expected cosmological curl-type lensing signal, so this term is absent in Eq. (5.39).

To calculate these autocorrelation functions and their variances, we use the WMAP 7-year cosmological parameters of Ref. (276). We plot the results of our calculations in Figs. 5.1 and 5.2, where lensing from scalar perturbations is plotted in green squares (absent for  $C_L^{\Omega\Omega}$  since there are no scalar contributions to the curl modes), lensing from tensor perturbations is plotted in blue circles, and the variance of these measurements is shown in red +s. We use the parameters of the Planck satellite,  $NET = 62 \mu K s^{1/2}$ ,  $t_{\text{obs}} = 2 \text{ yr}$ ,  $\theta_{\text{FWHM}} = 2.0635 \times 10^{-3} \text{ rad}$ , and  $f_{\text{sky}} \approx 1$ . We can see that the scalar  $\phi$  signal is several orders of magnitude greater than the tensor signal, and that the variance with which the  $\phi$ - $\phi$  power spectrum could be measured with Planck is higher than the scalar signal for low multipoles. The corresponding variance with which the  $\Omega$ - $\Omega$  power spectrum could be measured is also significantly larger than the signal. In both of these cases, therefore, the signal to noise of measuring the lensing from tensors using all multipoles with Planck is negligible, and remains negligible even in the case of the ideal CMB experiment with zero noise.

Thus, a stochastic background of GWs with power spectrum  $P_T(k)$  predicts a spectrum of nonzero even and odd BiPoSHs given by Eqs. (5.10) and (5.11), with values of  $\phi_{LM}$  and  $\Omega_{LM}$  selected from Gaussian distributions with the variances  $C_L^{\phi\phi \text{ ten}}$  and  $C_L^{\Omega\Omega \text{ ten}}$  given by Eq. (5.36).

### 5.3.5 Non-Cosmological Odd-Parity BiPoSH

As well as the intrinsic sources of odd-parity BiPoSH considered above, there are many possible non-cosmological sources. For example, artifacts from incomplete sky coverage, such as removing foreground contaminants like the galaxy or bright point sources, introduce both even- and odd-parity BiPoSH modes.

Similarly, errors due to the pixellation of the sky introduce BiPoSH modes of both parities. The effect of pixellation is to average the temperature over the entire pixel, and then assign that value to the location of the pixel center. Clearly, this process will introduce both curl- and gradient-type deflections of the CMB temperature field, and as such induces both even- and odd-parity BiPoSH modes.

Finally, we consider odd-parity BiPoSH modes induced by pointing errors. A telescope pointing error can be described as a process that causes the positions of points on the sky to be mislabeled.

This then causes an effective deflection of the points on the sky  $\hat{n}_{\text{obs}} = \hat{n} + \vec{\Delta}(\hat{n})$ , where  $\hat{n}_{\text{obs}}$  is the direction that the telescope believes it is pointed in and  $\hat{n}$  is its actual pointing direction. As we saw in Section 5.3.1, we can decompose this deflection field  $\vec{\Delta}(\hat{n})$  into gradient and curl components, which source even- and odd-parity BiPoSHs, respectively. Thus, from Eq. (5.18) we can see that any pointing error that has a nonzero curl component  $\vec{\nabla}_{\hat{\sigma}} \times \vec{\Delta}(\hat{n})$  will excite odd-parity BiPoSHs.

Imagine, for example, that a satellite such as Planck misestimates the rate with which it is precessing. Since it is this precession that builds up observations of subsequent rings of the sky, such a misestimation would cause a shearing of each ring relative to its neighbors. This type of a deflection has a nonzero curl component, and thus would excite odd-parity BiPoSHs. Measurement of these BiPoSHs, and in particular the odd-parity BiPoSHs, can therefore provide a useful check for such pointing errors.

## 5.4 BiPoSHs as Probes of Parity Violation

### 5.4.1 Correlation of Opposite-Parity Lensing Components

Since the  $A_{l'l'}^{\oplus LM}$  and  $A_{l'l'}^{\ominus LM}$  have opposite parity for the same  $L$  and  $M$ , a cross-correlation between the two can arise only if there is some parity-breaking in the physics responsible for producing the departures from SI/Gaussianity. Here we mention, by way of example, chiral GWs as a mechanism to produce such a parity-violating correlation (277; 278; 279).

The contribution to the cross-correlation power spectrum from a single Fourier mode in the  $\hat{z}$  direction with + polarization is  $C_L^{\phi\Omega}(k\hat{z}) = \sum_m \langle \phi_{LM} \Omega_{LM}^* \rangle / (2L + 1) = 0$ ; it vanishes as the contribution from  $M = 2$  is canceled by that from  $M = -2$ . And if this is true, then by rotational invariance it is true for any other linearly polarized GW. We thus conclude that a stochastic GW background predicts  $C_L^{\phi\Omega} = 0$ . In other words, there is no cross-correlation between  $\phi$  and  $\Omega$ , and thus no cross-correlation between the even and odd BiPoSHs,  $A_{l'l'}^{\oplus LM}$  and  $A_{l'l'}^{\ominus LM}$ .

Following Ref. (277), however, consider a right-circularly polarized GW:  $h_R = h_+ + ih_\times$  (i.e., we sum a + polarization wave with a  $\times$  polarization wave out of phase by  $90^\circ$ ). The azimuthal-angle dependence for the wave is then  $e^{2i\varphi}$ , and  $\Omega_{LM}$  and  $\phi_{LM}$  have contributions only from  $M = 2$ . There is thus a nonzero cross-correlation between  $\phi$  and  $\Omega$ . Similarly for a left-circularly polarized GW  $h_L = h_+ - ih_\times$ , the  $\varphi$  dependence is  $e^{-2i\varphi}$ , and only  $M = -2$  modes are excited. There is again a cross-correlation between  $\phi$  and  $\Omega$ , but this time with the opposite sign.

In the standard inflationary scenario, there are equal numbers of right- and left-circularly polarized GWs, and the cross-correlation between  $\phi$  and  $\Omega$  therefore vanishes. But if for some reason there is an asymmetry between the number of right- and left-circularly polarized GWs (277; 278; 279; 280), a manifestation of parity breaking, then there may be a parity-violating cross-correlation between  $\phi$  and  $\Omega$ , and thus between  $A_{l'l'}^{\oplus LM}$  and  $A_{l'l'}^{\ominus LM}$ .

The chirality of the GW background can be parametrized by an amplitude  $A$  which can take values between  $-1$  and  $1$ , where  $A = +1$  denotes that all of the GWs are right-circularly polarized, and  $A = -1$  denotes that they are all left-circularly polarized. But we have seen that a right-handed GW contributes only to  $M = 2$  modes, while a left-handed one contributes only to  $M = -2$ . We can denote this by weighting  $M = 2$  components by  $(A + 1)/2$  and  $M = -2$  components by  $(A - 1)/2$ , so that our version of Eq. (5.30), for example, that is appropriate to the case of a chiral GW background will be

$$\Omega_{LM}^{\text{ten}+}(k\hat{z}) = i^L h^+(\vec{k}) \left[ \frac{(1+A)}{2} \delta_{M,2} - \frac{(1-A)}{2} \delta_{M,-2} \right] \sqrt{\frac{2L+1}{2}} F_L^\Omega(k), \quad (5.40)$$

and similarly for Eq. (5.33). In this way a fully right circularly-polarized GW background will have only contributions from  $M = 2$ , a fully left-circularly polarized background will have only contributions from  $M = -2$ , and if the amount of left and right-circularly polarized waves is equal, that is if the GW background is non-chiral, the contributions from  $M = 2$  and  $M = -2$  cancel. The  $\phi$ - $\Omega$  cross-correlation power spectrum is given by

$$C_L^{\phi\Omega} = A \int \frac{d^3k}{(2\pi)^3} P_T(k) F_L^\phi(k) F_L^\Omega(k). \quad (5.41)$$

Refs. (269; 271; 242) have shown that the amplitude of the stochastic gravitational-wave background is probably too small, even with the most optimistic assumptions, to produce a detectable gravitational-lensing signal in the CMB. The example of a chiral gravitational-wave background as a possible source of a detectable parity-breaking BiPoSH correlation is principally of academic interest. Still, Ref. (270) has recently argued that weak lensing of the CMB by GWs may be detectable in its cross-correlation with the CMB-polarization pattern induced by these GWs (272; 273; 281; 282; 283). We thus surmise that a chiral gravitational-wave background may still be able produce a detectable parity-breaking signal in BiPoSHs in cross-correlation with the CMB polarization, an idea we explore in the next section.

### 5.4.2 Large-Angle CMB Polarization Spectra

We follow the work of Ref. (270), finding the multipole moments of the CMB E- and B-type polarization spectra for large angular scales by considering only those modes that are produced after reionization. The spherical-harmonic coefficients of B-type polarization modes can be decomposed as

$$B_{lm} = \int \frac{d^3k}{(2\pi)^3} \sum_{\alpha=+,\times} B_{lm}^\alpha(\vec{k}), \quad (5.42)$$

where  $B_{lm}^\alpha(\vec{k})$  is the amplitude of polarization B modes multipole moment  $lm$  in the direction  $\vec{k}$ . The

general form of this amplitude is quite complicated, but we can simplify it if, as in Section 5.3.4, we consider only a single, +-polarized GW traveling in the  $\hat{z}$  direction with wavenumber  $k$ . In this case, the B-mode amplitude can be written

$$B_{lm}^+(k\hat{z}) = i^l h^\alpha(\vec{k}) (\delta_{m,2} - \delta_{m,-2}) \sqrt{\frac{2l+1}{2}} F_l^B(k), \quad (5.43)$$

$$F_l^B(k) = \frac{1}{2l+1} \sqrt{\frac{9\pi}{2}} \int_{\eta_{re}}^{\eta_0} d\eta \dot{\tau}(\eta) \{ (l+2)j_{l-1}[k(\eta_0 - \eta)] - (l-1)j_{l+1}[k(\eta_0 - \eta)] \} \int_{k\eta_{lss}}^{k\eta} dx \frac{-3j_2(x)}{x} \frac{j_2(k\eta - x)}{(k\eta - x)^2}, \quad (5.44)$$

where the  $h^\alpha(\vec{k})$  are the amplitudes of GW modes as defined in Eq. (5.25),  $\dot{\tau}(\eta)$  is the scattering rate  $\dot{\tau}(\eta) = n_e(\eta) \sigma_T a(\eta)$ , with  $n_e$  the electron density,  $\sigma_T$  the Thompson scattering cross section, and  $a$  the scale factor, and  $\eta_{re}$  and  $\eta_0$  are the conformal times at reionization and today, respectively. Since we are only interested in small scales, we find the approximation  $\eta_{lss} = 0$  is sufficient for our purposes, making the last integral significantly faster to evaluate. The result above agrees with the results of Ref. (270), whose method we followed in its derivation, up to a factor of  $i$ .

We find that the corresponding E-type polarization multipoles from tensor perturbations take the same form as  $B_{lm}$  above, except for the opposite sign in front of  $\delta_{m,-2}$  and a different factor in the curly brackets in Eq. (5.44). From Ref. (283) we find this alternative form to be  $(2l+1)/2 \{ -j_l(x) + j_l''(x) + 2j_l(x)/x^2 + 4j_l'(x)/x \}$ , where here  $x = [k(\eta_0 - \eta)]$ , and derivatives are with respect to  $x$ . Employing spherical Bessel function identities, we can then write the E-type polarization multipoles as

$$E_{lm} = \int \frac{d^3k}{(2\pi)^3} \sum_{\alpha=+,\times} E_{lm}^\alpha(\vec{k}), \quad (5.45)$$

$$E_{lm}^+(k\hat{z}) = i^l h^\alpha(\vec{k}) (\delta_{m,2} + \delta_{m,-2}) \sqrt{\frac{2l+1}{2}} F_l^E(k), \quad (5.46)$$

$$F_l^E(k) = \frac{1}{2l+1} \sqrt{\frac{9\pi}{2}} \int_{\eta_{re}}^{\eta_0} d\eta \dot{\tau}(\eta) \left\{ \frac{(2l+1)}{[k(\eta_0 - \eta)]^2} j_l[k(\eta_0 - \eta)] - \frac{(2l+1)(3l^2 + 3l - 4)}{(2l-1)(2l+3)} j_l[k(\eta_0 - \eta)] \right. \\ \left. + \frac{l(l+3)}{2(2l-1)} j_{l-2}[k(\eta_0 - \eta)] + \frac{(l+1)(l-2)}{2(2l+3)} j_{l+1}[k(\eta_0 - \eta)] \right\} \int_{k\eta_{lss}}^{k\eta} dx \frac{-3j_2(x)}{x} \frac{j_2(k\eta - x)}{(k\eta - x)^2}, \quad (5.47)$$

where the terms are defined as they were above for the  $B_{lm}$  amplitudes.

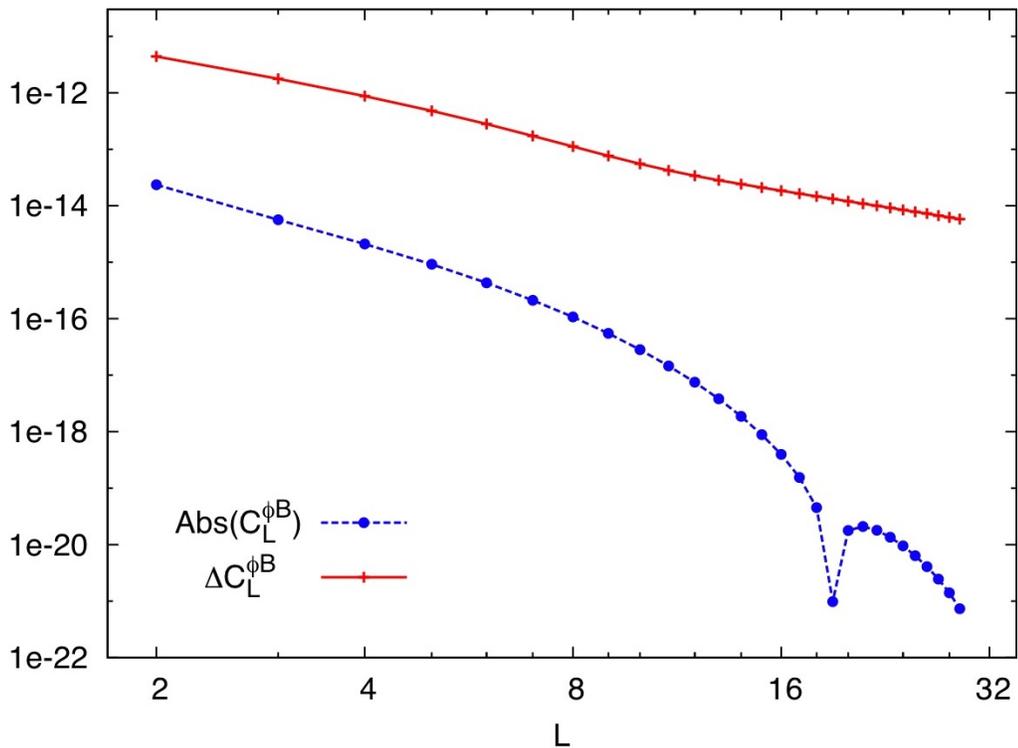


Figure 5.3: Here we plot the cross-correlation  $C_L^{\phi^B}$  between the gradient  $\phi$  modes of the weak lensing of cosmic shear with the curl-type  $B$  modes of the CMB polarization in blue circles, and the noise on this measurement due to cosmic variance and Planck satellite instrumental noise in red +s. Since these quantities are of opposite parity, in the absence of parity-breaking physics we expect this cross-correlation to vanish. However, if we assume, for example, that the entire allowable GW background is right-circularly polarized, such a cross-correlation could occur. The cross-correlation is linearly proportional to the chirality parameter  $A$ , defined such that  $A = 1$  denotes a completely right-circularly polarized GW background,  $A = -1$  denotes completely left-circularly polarized, and  $A = 0$  denotes an unpolarized background. Here we assume the maximum allowable tensor-to-scalar ratio  $r = 0.24$ , the limit from WMAP-7 data combined with BAO and the  $H_0$  measurement (276). Cusps in the absolute value of the correlation function correspond to sign changes of the correlation function.

### 5.4.3 Parity-Violating Correlations from Chiral GWs

We now want to calculate the expected cross-correlation between CMB-polarization multipole coefficients and weak-lensing-induced BiPoSHs of opposite parity. Note that these cross-correlations are directly related to the parity-odd three-point correlations discussed in Ref. (284). As we mentioned above, if there is no parity-violating physics, then in the cross-correlation of a parity-even and a parity-odd observable,  $M = 2$  terms and  $M = -2$  terms will cancel each other, giving a net zero cross-correlation. However, if for example the GW background is chiral, then parity is broken and we can get a nonzero cross-correlation between opposite parity observables. As we saw in Section 5.4.1, a right-handed GW contributes only to  $M = 2$  modes, while a left-handed one contributes only to

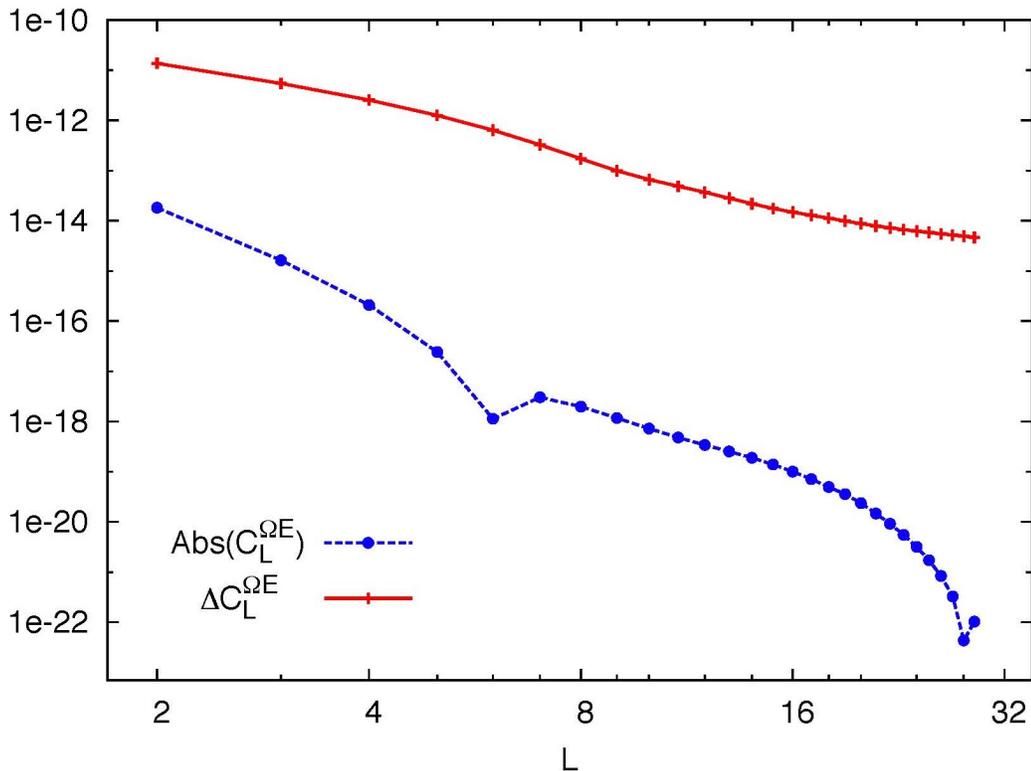


Figure 5.4: Here we plot the cross-correlation  $C_L^{\Omega E}$  between the curl-type  $\Omega$  modes of cosmic shear with the gradient-type  $E$ -modes of the CMB polarization in blue circles, and the noise on this measurement due to cosmic variance and Planck satellite instrumental noise in red +s. As with the  $\phi - B$  correlation, we assume a completely right-circularly polarized GW background, with the maximum currently permitted tensor-to-scalar ratio.

$M = -2$ . If we carry out a similar procedure for Eqs. (5.43), and (5.46) as we did in Eq. (5.40), weighting  $M = 2$  components by  $(A + 1)/2$  and  $M = -2$  components by  $(A - 1)/2$ , we can calculate parity-violating correlations between polarization and lensing components while accounting for the amplitude and handedness of a chiral GW background.

First considering the cross-correlation between  $B$ -modes of the CMB polarization and gradient-type modes of cosmic shear, we write

$$C_L^{\phi B} = \frac{1}{2L + 1} \sum_M \langle \phi_{LM} B_{LM}^* \rangle.$$

As before, by rotational invariance we know that both  $+$  and  $\times$  polarizations will contribute equally to  $C_L^{\phi B}$ , as will modes with any wavenumber  $\vec{k}$  whose magnitude  $k$  is the same. We can see that only  $\phi_{LM}^{\text{ten}}$  will contribute to this correlation, and not  $\phi_{LM}^{\text{sca}}$ , as the scalar perturbation field is not correlated, on average, with the tensor perturbation field. Then using Eqs. (5.32), (5.33), (5.35), (5.42) and (5.43), we can write this cross-correlation as

$$C_L^{\phi B} = A \int \frac{d^3 k}{(2\pi)^3} P_T(k) F_L^\phi(k) F_L^B(k). \quad (5.48)$$

Similarly, we can write the cross-correlation between  $E$ -type polarization modes and curl-type modes of cosmic shear, using Eqs. (5.29), (5.30), (5.35), (5.45), and (5.46), as

$$C_L^{\Omega E} = A \int \frac{d^3 k}{(2\pi)^3} P_T(k) F_L^\Omega(k) F_L^E(k), \quad (5.49)$$

where the GW power spectrum is given by

$$P_T(k) = \frac{\pi^2 r \Delta_R^2(k_0)}{2 k^3}.$$

We want to calculate the magnitude and shape of such correlations, to determine whether such a signal is observable. We use the WMAP 7-year cosmological parameters and assume the maximum allowable level of GWs from early universe physics, with a tensor-to-scalar ratio  $r = 0.24$ , the limit from the WMAP-7 data combined with BAO and the  $H_0$  measurement (276). We also assume that the GW background is entirely right-circularly polarized. As a first estimate, we calculate the level of such correlations while making several assumptions. We use the approximate form of the GW transfer function  $T(k, \eta) \simeq 3j_1(k\eta)/(k\eta)$ , assume that reionization happened instantaneously so that the electron density  $n_e$  is equal to a step function, and neglect contributions to the polarization modes that came from last scattering. The two last assumptions affect mostly the higher- $L$  multipoles, which in this cross-correlation are suppressed since we see that  $\phi_{LM}^{\text{ten}}$  and  $\Omega_{LM}^{\text{ten}}$  fall off very fast with  $L$ .

With these assumptions, we have calculated the correlation functions  $C_L^{\phi B}$  and  $C_L^{\Omega E}$ , and show them as the blue circles in Figs. 5.3 and 5.4. Note that the absolute value of the correlation functions are plotted, and that the cusps in the profiles result from sign changes. Note also that both correlation functions are linearly proportional to the chirality parameter  $A$ , so that they would flip in sign if the GW background were left instead of right-circularly polarized. We are only interested in low multipoles, since our assumptions break down for larger  $L$ , and such multipoles are strongly suppressed in correlation with the weak-lensing modes.

#### 5.4.4 Variance of $\phi$ -B and $\Omega$ -E Correlations

It is useful to know the variance with which we could measure such parity-violating cross-correlations. From Ref. (272) we see that the variance with which we could measure the cross-correlation  $C_L^{XY}$  of two distinct Gaussian random variables  $X$  and  $Y$  is given by

$$(\Delta C_L^{XY})^2 \equiv \left\langle \left( \widehat{C_L^{XY}} - C_L^{XY} \right)^2 \right\rangle,$$

where  $\widehat{C_L^{XY}} = 1/(2L+1) \sum_M X_{LM} Y_{LM}^*$  is the estimator for the cross-correlation, and  $C_L^{XY}$  is its theoretical value under the null hypothesis. Ref. (272) then evaluates this variance, assuming distinct  $X$  and  $Y$ , to be

$$(\Delta C_L^{XY})^2 = \frac{1}{2L+1} \left[ (C_L^{XY})^2 + C_L^{XX \text{ map}} C_L^{YY \text{ map}} \right], \quad (5.50)$$

where, as before,  $C_L^{XX \text{ map}} = W_L^2 C_L + N_L^{XX}$ , with  $W_L$  the window function defined in Section 5.2.3, and  $N_L^{XX}$  the noise in the measurement of  $C_L^{XX}$ .

In our case, the null hypothesis is that there is a GW background with the maximal tensor-to-scalar ratio, but it contains equal numbers of right- and left-circularly polarized GWs, i.e., it is not chiral. In this case, the theoretical value of parity-violating cross-correlations is zero, so that the first term in Eq. (5.50) vanishes. Then, assuming that  $\widehat{\phi_{LM}}$  and  $\widehat{\Omega_{LM}}$  are Gaussian random variables, a reasonable assumption since many uncorrelated noise processes are likely to contribute to this measured value, we find for the variances,

$$\left( \Delta C_L^{\phi B} \right)^2 = \frac{1}{2L+1} C_L^{\phi \phi \text{ map}} C_L^{BB \text{ map}} \quad (5.51)$$

$$\left( \Delta C_L^{\Omega E} \right)^2 = \frac{1}{2L+1} C_L^{\Omega \Omega \text{ map}} C_L^{EE \text{ map}}. \quad (5.52)$$

To calculate these errors, we know that the instrumental errors on the polarization power spectra are given by

$$N_L^{EE} = N_L^{BB} = \frac{8\pi(\text{NET})^2}{t_{\text{obs}} \sqrt{f_{\text{sky}}}}.$$

We use the Planck-satellite parameters, as in Section 5.3.4. We also use the CMB anisotropy calculator CAMB to calculate the temperature and polarization power spectra including effects at the surface of last scatter (285). The resulting errors are shown as red +s in Figs. 5.3 and 5.4. This noise, which combines instrumental and cosmic-variance sources, is at least an order of magnitude above the corresponding maximum signal level at low multipoles, and drops less rapidly with  $l$  so that the low multipoles yield the highest signal-to-noise.

#### 5.4.5 Signal-to-Noise Ratio of Chiral GW Background Detection

We finally wish to calculate the achievable signal-to-noise of a measurement of the magnitude of such cross-correlations given our calculations of their shapes and variances. Such a measurement would tell us about the presence or absence of a chiral GW background, or of parity violation in the processes that caused departures from Gaussianity/SI in general. We can phrase the aim of this

calculation as finding the error with which we could measure the chirality parameter  $A$ , which sets the amplitude of the cross-correlations relative to their maximum values in the case of a completely circularly polarized GW background, as in Eqs. (5.48) and (5.49). Let us calculate this for the case of the  $\phi$ - $B$  cross-correlation; the  $\Omega$ - $E$  case will be similar.

We define a new quantity  $C_{L \max}^{\phi B}$ , defined such that

$$C_L^{\phi B} = A C_{L \max}^{\phi B}.$$

If we assume that the instrumental noise on  $C_L^{\phi B}$  is Gaussian, so that  $\widehat{C}_L^{\phi B} \equiv W_L^{-2} C_L^{\phi B}$  is a random variable drawn from a Gaussian probability distribution with variance  $(\Delta C_L^{\phi B})^2$  and mean  $A C_{L \max}^{\phi B}$ , we can find the maximum-likelihood estimator for  $A$  to be

$$\widehat{A} = \frac{\sum_L \widehat{C}_L^{\phi B} C_{L \max}^{\phi B} (\Delta C_L^{\phi B})^{-2}}{\sum_L (C_{L \max}^{\phi B})^2 (\Delta C_L^{\phi B})^{-2}}.$$

Then, assuming that the instrumental noise is uncorrelated between different multipoles, the variance of this estimator is given by

$$\langle \widehat{A}^2 \rangle = \left[ \sum_L (C_{L \max}^{\phi B})^2 (\Delta C_L^{\phi B})^{-2} \right]^{-1}.$$

The maximum signal-to-noise with which we can measure this amplitude is given by

$$\left( \frac{S}{N} \right)_{\max}^{\phi B} = \frac{\widehat{A}_{\max}}{\sqrt{\langle \widehat{A}^2 \rangle}} = \left[ \sum_L (C_{L \max}^{\phi B})^2 (\Delta C_L^{\phi B})^{-2} \right]^{1/2}. \quad (5.53)$$

The same method can be used to calculate the obtainable signal-to-noise from the  $\Omega$ - $E$  cross correlation, giving

$$\left( \frac{S}{N} \right)_{\max}^{\Omega E} = \left[ \sum_L (C_{L \max}^{\Omega E})^2 (\Delta C_L^{\Omega E})^{-2} \right]^{1/2}. \quad (5.54)$$

Using the values of the cross-correlations and their errors calculated above, we find that the obtainable signal-to-noise from measurement of these cross-correlations is 0.002 for  $C_L^{\Omega E}$  and 0.01 for  $C_L^{\phi B}$ . These numbers are too small for us to have any reasonable expectation of detection using the Planck satellite. Recalculating the above errors assuming an ideal CMB experiment, with no instrumental noise and infinite resolution, the values of the signal-to-noise only change by a factor of two, indicating that this method is not likely to be a promising way to detect a chiral GW background.

## 5.5 Lensing of 21 cm Fluctuations by Primordial Gravitational Waves

In this section, we apply the mechanics of BiPoSH decomposition described above to a new system: intensity maps of the 21 cm radiation. Our work resembles in spirit that in Ref. (286) which argued that the huge number of Fourier modes available in 21 cm maps of the dark-age hydrogen distribution would provide considerable statistical significance in detecting the IGW distortion to matter fluctuations. However, they consider the *intrinsic* distortion to matter fluctuations by IGWs. On the other hand, we consider the distortion to the *images* of the matter distribution by lensing by IGWs. Our work is related to that of Ref. (287), who considered reconstruction of the lensing field due to density perturbations with 21 cm fluctuations.

As before, we consider the curl component of lensing induced by GWs, as given in Eqs. (5.30), (5.31) and (5.36). The angular power spectra for the lensing of sources at several redshifts are shown in Figure 5.5; for  $L \lesssim 6$ , the source-redshift dependence is weak for a scale-invariant gravitational-wave background.

Once again, we can write the minimum-variance estimator for the spherical-harmonic coefficients for the curl component of lensing is

$$\widehat{\Omega}_{LM} = \frac{\sum_{ll'} Q_{ll'}^{\ominus L*} \widehat{A}_{ll'}^{\ominus LM}}{\sum_{ll'} |Q_{ll'}^{\ominus L}|^2 / (C_l^{\text{map}} C_{l'}^{\text{map}})}, \quad (5.55)$$

where  $C_l^{\text{map}} = C_l + C_l^{\text{n}}$  is the angular power spectrum of the map with  $C_l$  now the power spectrum of the 21 cm intensity and  $C_l^{\text{n}}$  the noise power spectrum, and the sums are only over  $l + l' + L = \text{odd}$ . We use lower-case  $l$  for 21 cm fluctuations and upper-case  $L$  for the lensing-deflection field. Here,  $Q_{ll'}^{\ominus L}$  is given by Eq. (5.12), and the estimator for the BiPoSH coefficients is given by Eq. (5.6). As before, the estimator for the power spectrum of the curl component of the deflection field is then  $\widehat{C}_L^{\Omega} = \sum_m |\widehat{\Omega}_{LM}|^2 / (2L + 1)$ . The variance of  $\widehat{\Omega}_{LM}$  under the null hypothesis is given by

$$(\sigma_L^{\Omega})^2 \equiv \langle |\widehat{\Omega}_{LM}|^2 \rangle = 2 \left[ \sum_{ll'} |Q_{ll'}^{\ominus L}|^2 / (C_l^{\text{map}} C_{l'}^{\text{map}}) \right]^{-1}. \quad (5.56)$$

This noise power spectrum is plotted in Figure 5.5 using the 21 cm power spectra from Ref. (222) and taking the noise power spectrum  $C_l^{\text{n}} = 0$  for  $l < l_{\text{max}}$  and  $C_l^{\text{n}} = \infty$  for  $l > l_{\text{max}}$ . We show results for several  $l_{\text{max}}$  which are, roughly speaking, the maximum value of  $l$  with which the 21 cm power spectrum can be measured with high signal-to-noise. The signal-to-noise (squared) with

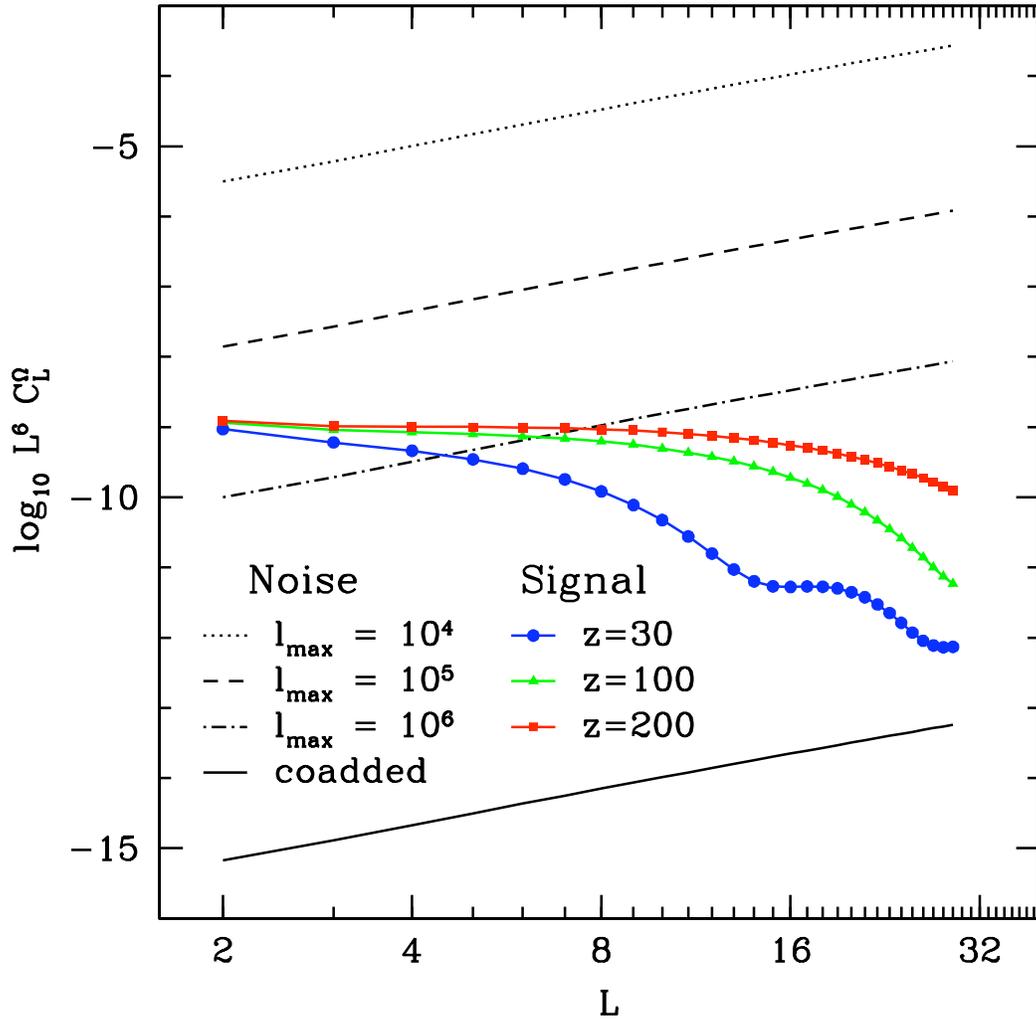


Figure 5.5: The power spectrum for the deflection-field curl component for lensing of sources at various redshifts by a scale-invariant spectrum of IGWs of the largest amplitude ( $r = 0.2$ ) consistent with current measurements. We also superimpose noise power spectra for lensing reconstruction carried out to various values of  $l_{\max}$ . Also shown is the noise power spectrum we estimate from co-adding the signals from all possible redshifts, assuming an  $l_{\max} = 10^6$ .

which IGWs can be detected is then

$$(S/N)^2 = \sum_L (L + 1/2) (C_L^\Omega)^2 / (\sigma_L^\Omega)^4. \quad (5.57)$$

Before reviewing the numerical results, it is instructive to consider an analytic estimate of the noise power spectrum  $(\sigma_L^\Omega)^2$ . To do so, we use the flat-sky approximation (242),

$$(\sigma_L^\Omega)^{-2} = \int \frac{d^2l}{(2\pi)^2} \frac{(\vec{L} \times \vec{l})^2 (C_l - C_{|\vec{L}-\vec{l}|})^2}{2C_l^{\text{map}} C_{|\vec{L}-\vec{l}|}^{\text{map}}}. \quad (5.58)$$

For  $L \ll l$  we approximate  $|\vec{L}-\vec{l}| \simeq l - L \cos \alpha$ , where  $\cos \alpha \equiv \hat{L} \cdot \hat{l}$ , and  $C_{|\vec{L}-\vec{l}|} \simeq C_l - L(\cos \alpha)(\partial C_l / \partial l)$ . If  $C_l \propto l^n$ , then

$$\begin{aligned} (\sigma_L^\Omega)^{-2} &= \int \frac{l dl}{4\pi^2} \int_0^{2\pi} d\alpha \frac{1}{2} L^4 \sin^2 \alpha \cos^2 \alpha \left( \frac{\partial \ln C_l}{\partial \ln l} \right)^2 \\ &\simeq L^4 n^2 l_{\text{max}}^2 / (64\pi). \end{aligned} \quad (5.59)$$

The flat-sky calculation is accurate for  $L \gtrsim 20$  and *overestimates* the noise by up to 30% at smaller  $L$ . As shown in Figure 2 in Ref. (222), the 21 cm power spectrum extends without suppression out to  $l \gtrsim 10^6$ , and values of  $l_{\text{max}} \sim 10^7$  are perhaps achievable with a bit more effort. However, given the rapid suppression of the 21 cm power spectrum at higher  $l$ , the return on the investment of noise reduction in terms of higher  $l_{\text{max}}$  will probably be small above  $l_{\text{max}} \simeq 10^7$ .

We now approximate the  $\Omega$  power spectrum (for  $r = 0.2$ ) as  $C_L^\Omega \simeq 10^{-11} (L/2)^{-6}$ . Although this approximation differs from the numerical results for different redshifts  $z$  at  $L \simeq 30$ , it is quite accurate for all  $30 \lesssim z \lesssim 200$  for the smallest  $L$  where most of the signal arises. From Eq. (5.57), the signal-to-noise with which the gravitational-wave background can be detected is

$$(S/N) \simeq 4.5 (l_{\text{max}}/10^6)^2 (n/2)^2 (L_{\text{min}}/2)^{-1}, \quad (5.60)$$

where  $L_{\text{min}}$  is the minimum  $L$  that can be measured.

There are several things to note about this result: (1) The signal-to-noise obtained with the adopted fiducial values for  $l_{\text{max}}$ ,  $L$ , and  $n$  is significant. (2) The scaling of the signal-to-noise with  $l_{\text{max}}$  is very rapid, and greater than what might have been expected ( $\propto l_{\text{max}}$ ) naively. The origin of this rapid scaling is similar to that for detection of the local-model trispectrum (288) (as the signal we are measuring here is, strictly speaking, an intensity trispectrum). Thus, the sensitivity to a gravitational-wave background increases very rapidly as the angular resolution of the map is improved. (3) The sensitivity decreases as  $L_{\text{min}}$  is increased, so good sky coverage is important for gravitational-wave detection.

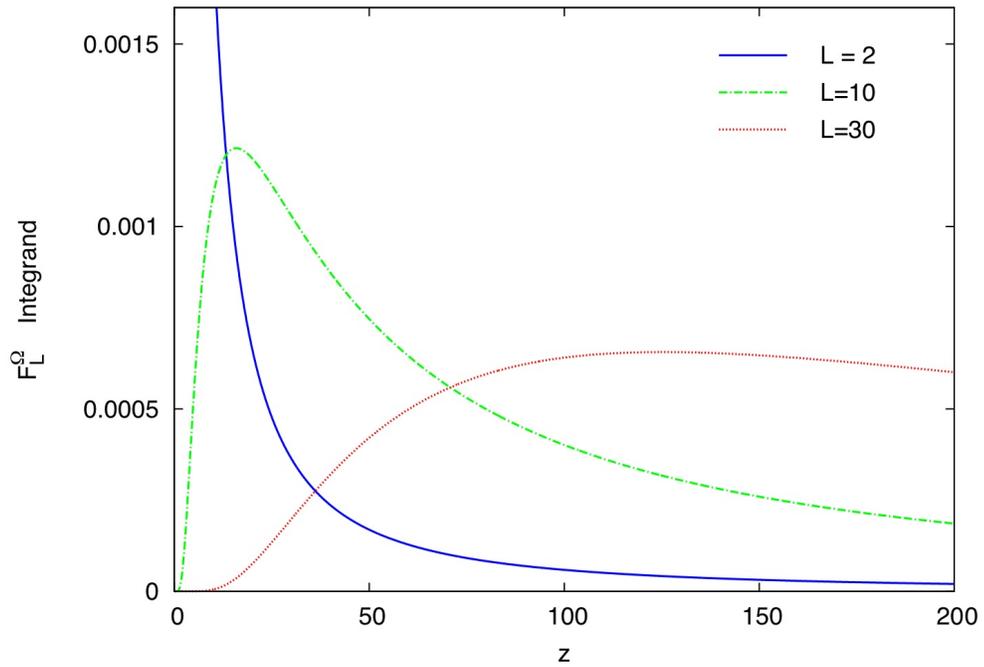


Figure 5.6: Shown are the integrands of the curl-mode transfer function Eq. (5.31) for three values of the angular scale of the lensing field. These curves qualitatively show the redshift distribution of the GWs which lens the background 21 cm field on the observed scale of the lensing distribution. The curves are normalized such that the area under each is unity. It is clear that, for the lowest multipoles of the lensing field, all of the lensing is occurring very near to the observer, so the assumption that the lensing field is the same for all 21 cm redshifts is justified.

While a signal-to-noise of 4.5 is respectable, and could be improved with even larger  $l_{\max}$ , we can go much further: By changing the frequency at which the 21 cm map is made, we look at spherical shells of atomic hydrogen at different redshifts. Suppose, then, that we have 21 cm maps at two different frequencies that correspond to spherical shells separated along the line of sight by a comoving distance  $\delta R$ . Those two maps are statistically independent at the highest  $l$  (where the vast majority of the signal-to-noise for IGW detection arises) if  $(\delta R/R) \gtrsim l^{-1}$ . If  $\Delta R$  is the separation in comoving radius corresponding to the entire frequency range covered by the observations (say, redshifts  $z \simeq 30-200$ ), then the total number of statistically independent maps that can be obtained is  $N_z \simeq (\Delta R/\delta R) \simeq l(\Delta R/R) \simeq 0.15l$ . If so, then each map contributes an independent upper limit to the GW background amplitude, and the signal-to-noise from these redshift ranges can be added in quadrature, increasing the total signal-to-noise by a factor  $N_z^{1/2}$ . But there may be room for even more improvement: If most of the lensing occurs at redshifts  $z \lesssim 30$ , then the lensing pattern is the same for all redshift shells in which case every redshift shell contributes coherently to an estimator for  $\Omega_{LM}$ . In this case,  $(\sigma_L^\Omega)^2$  is decreased by factor  $N_z^{-1}$ , and the signal-to-noise increased by a factor  $N_z$  relative to the single- $z$  estimate. In Figure 5.6, we show the redshift dependence of the integrand of the curl-mode transfer function Eq. (5.31) for three values of the lensing field multipole moment. This quantity directly shows the redshifts of the GWs that are contributing to this multipole of the lensing field, and the figure clearly shows that for low multipole moment, all of the lensing is occurring very close to the observer at  $z = 0$ . While these curves have been normalized for comparison, we know that almost all of the signal comes from the lowest multipole moments. In this case, then, the assumption that the lensing field for each redshift slice of 21 cm radiation is the same is justified.

Using that most of the signal comes from the lowest  $L$ , we estimate that the signal-to-noise for IGW detection obtained by coadding redshift shells will be

$$(S/N)_{\text{tot}} \simeq 6.8 \times 10^5 (l_{\max}/10^6)^3 (n/2)^2 (L_{\min}/2)^{-1}, \quad (5.61)$$

assuming (as above) the largest currently allowed IGW amplitude  $r \simeq 0.2$ . Put another way, the smallest tensor-to-scalar ratio that can be detected at the  $3\sigma$  level is

$$r \simeq 10^{-6} (L_{\min}/2) (l_{\max}/10^6)^{-3} (n/2)^{-2}. \quad (5.62)$$

Note that the dependence on  $l_{\max}$  is very steep, and including all the information to  $l_{\max} = 10^7$  could yield a detection threshold of  $r \simeq 10^{-9}$ . Note also that a more-sophisticated analysis, including the full structure of cross-correlations between redshift maps, may be able to improve upon the  $N_z^{1/2}$  scaling even in the case where the lensing signal is incoherent at different redshifts. The full-sky calculation, including a more-realistic shape of  $C_l$ , yields a result consistent with this estimate

(Figure 5.5).

To put this result in perspective, we note that the current upper bound  $r \lesssim 0.22$  comes from WMAP measurements of temperature-polarization correlations, although not from B-mode null searches. The forthcoming generation of sub-orbital B-mode experiments are targeting  $r \lesssim 0.1$ , and a dedicated CMB-polarization satellite might then get to  $r \sim 10^{-2}$  (289).

Measurement of gravitational-wave amplitudes  $r \lesssim 0.01$  with CMB polarization will have to contend with the additional contribution to B-mode polarization from gravitational lensing (by density perturbations) of primordial E modes (290). The two contributions (IGW and lensing) to B modes can be distinguished if the lensing deflection angle can be reconstructed with small-scale CMB fluctuations (291; 292). This may allow values  $r \sim 10^{-3}$  to be probed, although it requires a far more sophisticated CMB experiment (with far better angular resolution) than simple detection of B modes would require.

Further progress in separation of lensing and IGW contributions to B modes can be obtained with 21 cm measurements (287) of precisely the type we discuss here but of the curl-free lensing component (due to density perturbations) rather than the curl component from IGWs. Such measurements, when combined with a precise CMB polarization experiment, can in principle get to IGW amplitudes comparable to those we have discussed here. Measurement of the 21 cm curl component may therefore ultimately be competitive for the most sensitive probe of IGWs, even if a sensitive CMB-polarization experiment is done. Furthermore, if both 21 cm observations and a CMB-polarization map are available, then measurement of the 21 cm curl component can be used as a cross-check and to complement a measurement from the combination of B-mode polarization with 21 cm lensing subtraction.

While we have focussed here on the dark ages, similar measurements can be performed with 21 cm fluctuations from the epoch of reionization and with galaxy surveys; the critical issue will be how high  $l_{\text{max}}$  can get. While the 21 cm curl component induced by lensing by density perturbations at second order is too small to be an issue (242), a curl component may conceivably arise since the atomic-hydrogen distribution is not perfectly Gaussian due to non-linear gravitational collapse and baryonic effects. We speculate that this curl component will be small for the small- $L$  modes at which the IGW signal peaks. We also imagine that the information from multiple redshifts may be combined to separate the IGW and any bias-induced signal.

To close, we note that the measurements we describe will be challenging and are very futuristic compared to what current and next-generation experiments will accomplish. Still, 21 cm cosmology is an exciting and rapidly developing experimental arena, for a good number of scientific reasons (219), and we hope that the idea presented here provides one additional motivation to carry such work forward.

## 5.6 Conclusions

BiPoSHs are a formalism to describe correlations between two different spherical-harmonic coefficients of the CMB temperature field, which can occur if the CMB temperature field is not exactly Gaussian or statistically isotropic. This chapter introduces odd-parity BiPoSHs, a set of BiPoSHs that has not yet been studied, and details how they can be estimated from knowledge of the CMB temperature fluctuations.

We calculate the even- and odd-parity BiPoSHs that are sourced by gradient- and curl-type deflections of the CMB, respectively, and from this we obtain estimators for these deflections in terms of the BiPoSH coefficients. We show that lensing by scalar metric perturbations causes only gradient-type deflections, and thus only sources even-parity BiPoSHs. However, lensing by GWs produces both gradient- and curl-type deflections and thus sources both even- and odd-parity BiPoSHs. We calculate the expected power spectra of deflections due to scalar and tensor perturbations and their errors, and conclude that a reasonable signal-to-noise measurement of the amplitude of the GW background cannot be obtained from these autocorrelations even with the ideal CMB experiment, and thus from autocorrelations of the BiPoSH coefficients.

Although lensing by GWs produces both even- and odd-parity BiPoSHs, their opposite parity implies that they could not be correlated. However, in the presence of parity-violating physics, such as a chiral GW background, this parity argument breaks down and we might expect a correlation. We consider such a cross-correlation, and encourage its measurement even though the likelihood of observing a cosmological signal is low.

A GW background also produces signals in the  $E$ - and  $B$ -type CMB polarization spectra, which are of even and odd parity, respectively. We consider the possibility that a chiral GW background would produce cross-correlations between opposite-parity components of lensing and polarization, and calculate the expected magnitude and errors of such cross-correlations. Although we find that the likelihood of observing a cosmological signal is low, we encourage the measurement of these cross-correlations since such a detection would provide evidence of important systematic errors or even new parity-breaking physics.

We also apply the BiPoSH decomposition formalism to the 21 cm radiation, and in particular consider the detection of a GW background using the induced curl modes of lensing. We find that, with a futuristic 21 cm map at several redshifts, we can obtain very strong constraints on the amplitude of the GW background. We note that, to obtain comparable constraints, missions probing the B-modes of the CMB polarization will need to have a very precise knowledge of the scalar lensing field to remove contamination from lensed E-modes. This type of knowledge can be obtained from exactly the type of 21 cm experiment that we describe here, implying that these two methods may eventually be comparable, complementary tests of the GW background.

We note that the BiPoSH formalism can also be generalized to include the polarization as well. It may be that inclusion of the polarization improves the sensitivity to these parity-breaking, and other, signals. Finally, we note that weak-lensing distortions of distant galaxies can also be decomposed into curl and gradient components (267; 293). Similar tests for parity violation can thus also be carried out with weak lensing of galaxies.

## Appendix A

# Accretion Shock Calculation

The accretion shock model of (88) upon which we base our implementation makes several assumptions. The cluster is assumed to be spherically symmetric and in hydrodynamic equilibrium, with an effective equation of state  $P(r) \propto [\rho(r)]^{\gamma_{\text{eff}}}$ , where  $\gamma_{\text{eff}} = 1.2$ , and with the actual equation of state of a free monatomic gas,  $P = K(s)\rho^{5/3}$ , where  $K$  is a function of the specific entropy. The cluster potential is assumed to be of the NFW (294) type, and the post-shock velocity of the gas is assumed to be negligible.

With these assumptions, they obtain the following hydrostatic model of the gas within the accretion shock:

$$T(x) = T_{\Delta}g(x) \tag{A.1}$$

$$\rho(x) = \rho_g [g(x)]^{1/(\gamma_{\text{eff}}-1)} \tag{A.2}$$

$$P(x) = \frac{T_{\Delta}\rho_g}{\mu m_p} [g(x)]^{\gamma_{\text{eff}}/(\gamma_{\text{eff}}-1)} \tag{A.3}$$

$$g(x) = g_0(x) + g_1 \tag{A.4}$$

$$g_0(x) = \frac{2(\gamma_{\text{eff}} - 1)}{\gamma_{\text{eff}}} \frac{\ln(1 + cx)}{\ln(1 + c) - c(1 + c)^{-1}} \frac{1}{x}. \tag{A.5}$$

Here,  $x \equiv r/r_{\text{vir}}$  is the ratio of the cluster-centric radius to the virial radius,  $\rho_g$  and  $g_1$  are constants of integration, which can be specified by the constraint that the baryonic mass inside the shock radius is equal to the total baryonic mass of the halo  $f_b M$  and the shock jump condition relating post-shock temperature to the incoming velocity, respectively.

In our implementation of this model we relax two particularly unjustified assumptions, that the

accretion shock is always strong, and that the constant  $g_1$  in Eq. (A.4) is negligible. To relax these assumptions we apply the further (better-founded) assumption that the gas accreting onto the halo is compressed and heated adiabatically as it makes its way to the shock radius from the mean density IGM. Using this assumption, and also assuming that the temperature of the mean IGM is 3000 K, we can calculate the temperature just outside the shock, from which we can obtain the Mach number of the shock  $\mathcal{M}$ .

With this prescription for finding the strength of the shock  $\mathcal{M}$ , where  $\mathcal{M} \rightarrow \infty$  is the strong-shock limit initially imposed by (88), we use the shock relations of (295) to obtain an equation for  $g_1$ :

$$g_1 = \frac{((2\gamma\mathcal{M}^2 - (\gamma - 1))((\gamma - 1)\mathcal{M}^2 + 2))}{\gamma(\mathcal{M}^2 - 1)^2} \frac{\xi}{x_{ac}} - \frac{2(\gamma_{\text{eff}} - 1)}{\gamma_{\text{eff}}} \frac{\ln(1 + cx_{ac})}{\ln(1 + c) - c(1 + c)^{-1}} \frac{1}{x_{ac}}. \quad (\text{A.6})$$

Here, the ratio of the accretion shock radius to the virial radius is shown as  $x_{ac}$ . We finally find that the equation for the accretion shock, Eq. (A12) in (88), is modified to

$$\left[ \frac{3\mathcal{C}}{4} \tilde{\rho} I(x_{ac}, g_1, \gamma_{\text{eff}}, c) \right]^2 x_{ac}^3 + x_{ac} - 2 = 0, \quad (\text{A.7})$$

where  $\mathcal{C}$  is the shock compression factor, the ratio of the densities internal and external to the accretion shock,

$$\mathcal{C} \equiv \frac{\rho_2}{\rho_1} = \frac{(\gamma + 1)\mathcal{M}^2}{(\gamma - 1)\mathcal{M}^2 + 2}, \quad (\text{A.8})$$

$I$  is the integral

$$I(x_{ac}, g_1, \gamma_{\text{eff}}, c) \equiv \frac{1}{x_{ac}^3} \int_0^{x_{ac}} \left[ \frac{g_0(x) + g_1}{g(x_{ac})} \right]^{1/(\gamma-1)} x^2 dx, \quad (\text{A.9})$$

and we use the variable  $\tilde{\rho}$  which is defined in (88) to be

$$\tilde{\rho} \equiv \frac{4}{3} \frac{2}{\Delta}^{1/2} (Ht)^{-1} \frac{d \ln M}{d \ln t}. \quad (\text{A.10})$$

There is a further subtlety in the calculation of the mass accretion rate onto halos, which has a strong effect on the final accretion shock radius through its appearance in Eq. (A.10). We have explored several possibilities, but in this work we adopt a spherically symmetrized model averaged over the dynamical time of the halo. More specifically, we calculate the average mass accretion rate over the past one dynamical time of the halo, including all of the mass added to the most massive progenitor of the current halo through mergers or accretion. Since we use this mass as an estimate of the spherically accreted mass of the halo, we are making the assumption that all of this added

mass, even that which was added through mergers, will distribute itself and cause accretion shocks much as the same mass of spherically symmetric accretion would.

# Bibliography

- [1] M. L. Balogh, J. F. Navarro, and S. L. Morris, *ApJ* **540**, 113 (2000).
- [2] M. Verdugo, B. L. Ziegler, and B. Gerken, *A&A* **486**, 9 (2008).
- [3] H. Butcher and A. Oemler, Jr., *ApJ* **226**, 559 (1978).
- [4] A. Dressler, *ApJ* **236**, 351 (1980).
- [5] B. C. Whitmore, D. M. Gilmore, and C. Jones, *ApJ* **407**, 489 (1993).
- [6] W. A. Barkhouse, H. K. C. Yee, and O. López-Cruz, *ApJ* **703**, 2024 (2009).
- [7] B. Moore, N. Katz, G. Lake, A. Dressler, and A. Oemler, *Nature* **379**, 613 (1996).
- [8] M. L. Balogh, S. L. Morris, H. K. C. Yee, R. G. Carlberg, and E. Ellingson, *ApJ* **527**, 54 (1999).
- [9] J. E. Gunn and J. R. I. Gott, *ApJ* **176**, 1 (1972).
- [10] M. G. Abadi, B. Moore, and R. G. Bower, *MNRAS* **308**, 947 (1999).
- [11] V. Quilis, B. Moore, and R. Bower, *Science* **288**, 1617 (2000).
- [12] J. D. P. Kenney *et al.*, in *The Evolving ISM in the Milky Way and Nearby Galaxies*, 2009.
- [13] K. Bekki, W. J. Couch, and Y. Shioya, *ApJ* **577**, 651 (2002).
- [14] I. G. McCarthy *et al.*, *MNRAS* **383**, 593 (2008).
- [15] A. I. Zabludoff and J. S. Mulchaey, *ApJ* **496**, 39 (1998).
- [16] S. L. McGee, M. L. Balogh, R. G. Bower, A. S. Font, and I. G. McCarthy, *MNRAS* **400**, 937 (2009).
- [17] D. Kawata and J. S. Mulchaey, *ApJ* **672**, L103 (2008).
- [18] Y. Fujita, *PASJ* **56**, 29 (2004).
- [19] F. G. Braglia, D. Pierini, A. Biviano, and H. Böhringer, *A&A* **500**, 947 (2009).

- [20] S. D. M. White and C. S. Frenk, *ApJ* **379**, 52 (1991).
- [21] E. Bertschinger, *ApJS* **58**, 39 (1985).
- [22] S. R. Furlanetto and A. Loeb, *ApJ* **611**, 642 (2004).
- [23] R. Barkana, *MNRAS* **347**, 59 (2004).
- [24] U. Keshet, E. Waxman, A. Loeb, V. Springel, and L. Hernquist, *ApJ* **585**, 128 (2003).
- [25] S. M. Molnar *et al.*, *ApJ* **696**, 1640 (2009).
- [26] Y. Birnboim and A. Dekel, *MNRAS* **345**, 349 (2003).
- [27] D. Kereš, N. Katz, D. H. Weinberg, and R. Davé, *MNRAS* **363**, 2 (2005).
- [28] A. J. Benson and R. Bower, 1004.1162.
- [29] S. Cole, C. G. Lacey, C. M. Baugh, and C. S. Frenk, *MNRAS* **319**, 168 (2000).
- [30] R. S. Somerville, J. R. Primack, and S. M. Faber, *MNRAS* **320**, 504 (2001).
- [31] C. M. Baugh *et al.*, *MNRAS* **356**, 1191 (2005).
- [32] G. Kauffmann, J. M. Colberg, A. Diaferio, and S. D. M. White, *MNRAS* **307**, 529 (1999).
- [33] A. J. Benson, C. S. Frenk, C. G. Lacey, C. M. Baugh, and S. Cole, *MNRAS* **333**, 177 (2002).
- [34] A. J. Benson, C. G. Lacey, C. M. Baugh, S. Cole, and C. S. Frenk, *MNRAS* **333**, 156 (2002).
- [35] R. S. Somerville, *ApJ* **572**, L23 (2002).
- [36] D. J. Croton *et al.*, *MNRAS* **365**, 11 (2006).
- [37] R. G. Bower *et al.*, *MNRAS* **370**, 645 (2006).
- [38] R. S. Somerville, P. F. Hopkins, T. J. Cox, B. E. Robertson, and L. Hernquist, *MNRAS* **391**, 481 (2008).
- [39] A. S. Font *et al.*, *MNRAS* **389**, 1619 (2008).
- [40] B. Lanzoni, B. Guiderdoni, G. A. Mamon, J. Devriendt, and S. Hatton, *MNRAS* **361**, 369 (2005).
- [41] S. M. Weinmann *et al.*, *MNRAS* **394**, 1213 (2009).
- [42] V. Springel *et al.*, *Nature* **435**, 629 (2005).
- [43] J. R. Bond, S. Cole, G. Efstathiou and N. Kaiser, *ApJ* **379**, 440 (1991).

- [44] C. G. Lacey and S. Cole, MNRAS **262**, 627 (1993).
- [45] A. J. Benson, M. Kamionkowski and S. H. Hassani, MNRAS **357**, 847 (2005) [arXiv:astro-ph/0407136].
- [46] R. K. Sheth, H. J. Mo and G. Tormen, MNRAS **323**, 1 (2001) [astro-ph/9907024].
- [47] A. J. Benson, C. Reichardt and M. Kamionkowski, MNRAS **331**, 71 (2002) [arXiv:astro-ph/0110299].
- [48] R. Scoccimarro, E. Sefusatti and M. Zaldarriaga, Phys. Rev. D **69**, 103513 (2004) [arXiv:astro-ph/0312286].
- [49] N. Dalal, O. Dore, D. Huterer and A. Shirokov, Phys. Rev. D **77**, 123514 (2008) [arXiv:0710.4560 [astro-ph]].
- [50] S. Chongchitnan and J. Silk, Phys. Rev. D **85**, 063508 (2012) [arXiv:1107.5617 [astro-ph.CO]].
- [51] F. Schmidt, A. Vikhlinin and W. Hu, Phys. Rev. D **80**, 083505 (2009) [arXiv:0908.2457 [astro-ph.CO]].
- [52] S. Ferraro, F. Schmidt and W. Hu, Phys. Rev. D **83**, 063503 (2011) [arXiv:1011.0992 [astro-ph.CO]].
- [53] M. D. Gladders and H. K. C. Yee, ApJS **157**, 1 (2005) [arXiv:astro-ph/0411075].
- [54] B. P. Koester *et al.*, ApJ **660**, 221 (2007) [arXiv:astro-ph/0701268].
- [55] Z. Staniszewski *et al.*, ApJ **701**, 32 (2009) [arXiv:0810.1578 [astro-ph]].
- [56] H. Ebeling, A. C. Edge, A. Mantz, E. Barrett, J. P. Henry, C. J. Ma and L. van Speybroeck, MNRAS **407**, 83 (2010).
- [57] J. Hao *et al.* [SDSS Collaboration], ApJS **191**, 254 (2010) [arXiv:1010.5503 [astro-ph.CO]].
- [58] K. Vanderlinde *et al.*, ApJ **722**, 1180 (2010) [arXiv:1003.0003 [astro-ph.CO]].
- [59] T. A. Marriage *et al.*, ApJ **737**, 61 (2011) [arXiv:1010.1065 [astro-ph.CO]].
- [60] N. Sehgal *et al.*, ApJ **732**, 44 (2011) [arXiv:1010.1025 [astro-ph.CO]].
- [61] N. Aghanim *et al.* [Planck Collaboration], A&A **536**, A10 (2011) [arXiv:1101.2043 [astro-ph.CO]].
- [62] A. E. Evrard, C. A. Metzler and J. F. Navarro, ApJ **469**, 494 (1996) [arXiv:astro-ph/9510058].

- [63] A. Mantz, S. W. Allen, H. Ebeling and D. Rapetti, MNRAS **387**, 1179 (2008) [arXiv:0709.4294 [astro-ph]].
- [64] S. Borgani *et al.*, ApJ **561**, 13 (2001) [arXiv:astro-ph/0106428].
- [65] A. Mahdavi, H. Hoekstra, A. Babul and J. P. Henry, MNRAS **384**, 1567 (2008) [arXiv:0710.4132 [astro-ph]].
- [66] A. Vikhlinin *et al.*, ApJ **692**, 1060 (2009) [arXiv:0812.2720 [astro-ph]].
- [67] H. J. Eckmiller, D. S. Hudson and T. H. Reiprich, AAP **535**, A105 (2011) [arXiv:1109.6498 [astro-ph.CO]].
- [68] C. M. Bird, ApJL **445**, L81 (1995) [arXiv:astro-ph/9503038].
- [69] S. Borgani, A. Gardini, M. Girardi and S. Gottlober, New Astron. **2**, 119 (1997).
- [70] M. Meneghetti, M. Bartelmann and A. Jenkins, MNRAS **381**, 171 (2007) [astro-ph/0509323].
- [71] A. Saro, G. Bazin, J. Mohr and K. Dolag, [arXiv:1203.5708 [astro-ph.CO]].
- [72] R. A. Sunyaev and Y. B. Zeldovich, Comments Astrophys. Space Phys. **4**, 173 (1972).
- [73] M. Birkinshaw, Phys. Rept. **310**, 97 (1999) [astro-ph/9808050].
- [74] J. E. Carlstrom, G. P. Holder and E. D. Reese, Ann. Rev. A&A **40**, 643 (2002) [astro-ph/0208192].
- [75] P. Draper, S. Dodelson, J. Hao and E. Rozo, Phys. Rev. D **85**, 023005 (2012) [arXiv:1106.2185 [astro-ph.CO]].
- [76] S. Dodelson, Amsterdam, Netherlands: Academic Pr. (2003) 440 p.
- [77] N. Kaiser and G. Squires, ApJ **404**, 441 (1993).
- [78] D. Clowe, N. Kaiser, G. Luppino, J. P. Henry and I. M. Gioia, astro-ph/9801208.
- [79] E. S. Sheldon *et al.* [SDSS Collaboration], ApJ **703**, 2217 (2009) [arXiv:0709.1153 [astro-ph]].
- [80] M. Bartelmann, Class. Quant. Grav. **27**, 233001 (2010) [arXiv:1010.3829 [astro-ph.CO]].
- [81] S. LaRoque, M. Bonamente, J. Carlstrom, M. Joy, D. Nagai, E. Reese and K. Dawson, ApJ **652**, 917 (2006) [astro-ph/0604039].
- [82] A. Kravtsov, A. Gonzalez, A. Vikhlinin, D. Marrone, A. Zabludoff, D. Nagai, M. Markevitch and B. Benson *et al.*, [arXiv:0903.0388 [astro-ph.CO]].

- [83] S. Ettori, A. Morandi, P. Tozzi, I. Balestra, S. Borgani, P. Rosati, L. Lovisari and F. Terenziani, *A&A* **501**, 61 (2009) [arXiv:0904.2740 [astro-ph.CO]].
- [84] S. W. Allen, A. E. Evrard and A. B. Mantz, *Ann. Rev. A&A* **49**, 409 (2011) [arXiv:1103.4829 [astro-ph.CO]].
- [85] J. C. Helly *et al.*, *MNRAS* **338**, 903 (2003).
- [86] M. Stringer, A. Brooks, A. Benson, and F. Governato, *MNRAS* submitted (2010).
- [87] A. J. Benson, *MNRAS* **358**, 551 (2005).
- [88] G. M. Voit, M. L. Balogh, R. G. Bower, C. G. Lacey, and G. L. Bryan, *ApJ* **593**, 272 (2003).
- [89] C. S. Frenk *et al.*, *ApJ* **525**, 554 (1999).
- [90] S. Borgani and A. Kravtsov, 0906.4370.
- [91] H. K. C. Yee, E. Ellingson, and R. G. Carlberg, *ApJS* **102**, 269 (1996).
- [92] T. C. Beers, K. Flynn, and K. Gebhardt, *AJ* **100**, 32 (1990).
- [93] G. De Lucia and J. Blaizot, *MNRAS* **375**, 2 (2007).
- [94] R. G. Bower, I. G. McCarthy, and A. J. Benson, *MNRAS* **390**, 1399 (2008).
- [95] W. H. Press and P. Schechter, *ApJ* **187**, 425 (1974).
- [96] R. K. Sheth and G. Tormen, *MNRAS* **329**, 61 (2002).
- [97] A. J. Benson, M. Kamionkowski, and S. H. Hassani, *MNRAS* **357**, 847 (2005).
- [98] J. Zhang, O. Fakhouri, and C. Ma, *MNRAS* **389**, 1521 (2008).
- [99] E. Neistein and A. Dekel, *MNRAS* **383**, 615 (2008).
- [100] H. Parkinson, S. Cole, and J. Helly, *MNRAS* **383**, 557 (2008).
- [101] M. Kuhlen, J. Diemand, P. Madau, and M. Zemp, *J. Phys. Conf. Ser.* **125**, 012008 (2008).
- [102] M. Boylan-Kolchin, V. Springel, S. D. M. White, A. Jenkins, and G. Lemson, *MNRAS* **398**, 1150 (2009).
- [103] J. Stadel *et al.*, *MNRAS* **398**, L21 (2009).
- [104] A. Klypin, S. Trujillo-Gomez, and J. Primack, arXiv:1002.3660, 2010.
- [105] J. F. Navarro, C. S. Frenk, and S. D. M. White, *ApJ* **490**, 493 (1997).

- [106] J. F. Navarro *et al.*, MNRAS **349**, 1039 (2004).
- [107] J. F. Navarro *et al.*, MNRAS **402**, 21 (2010).
- [108] J. Dubinski and R. G. Carlberg, ApJ **378**, 496 (1991).
- [109] H. J. Mo, S. Mao, and S. D. M. White, MNRAS **295**, 319 (1998).
- [110] A. J. Benson and R. Bower, MNRAS, 600 (2010).
- [111] O. Y. Gnedin, A. V. Kravtsov, A. A. Klypin, and D. Nagai, ApJ **616**, 16 (2004).
- [112] M. Gustafsson, M. Fairbairn, and J. Sommer-Larsen, Phys. Rev. D **74**, 123522 (2006).
- [113] E. Romano-Díaz, I. Shlosman, C. Heller, and Y. Hoffman, ApJ **687**, L13 (2008).
- [114] E. Romano-Díaz, I. Shlosman, C. Heller, and Y. Hoffman, ApJ **702**, 1250 (2009).
- [115] M. G. Abadi, J. F. Navarro, M. Fardal, A. Babul, and M. Steinmetz, arXiv:0902.2477, 2009.
- [116] P. B. Tissera, S. D. M. White, S. Pedrosa, and C. Scannapieco, arXiv:0911.2316, 2009.
- [117] A. R. Duffy *et al.*, “Impact of baryon physics on dark matter structures: a detailed simulation study of halo density profiles”, MNRAS, 624 (2010).
- [118] F. Governato *et al.*, Nature **463**, 203 (2010).
- [119] R. B. Tully and J. R. Fisher, A&A **54**, 661 (1977).
- [120] S. Hatton *et al.*, MNRAS **343**, 75 (2003).
- [121] G. De Lucia, G. Kauffmann, and S. D. M. White, MNRAS **349**, 1101 (2004).
- [122] P. Monaco, F. Fontanot, and G. Taffoni, MNRAS **375**, 1189 (2007).
- [123] A. A. Dutton, F. C. van den Bosch, A. Dekel, and S. Courteau, ApJ **654**, 27 (2007).
- [124] A. A. Dutton and F. C. van den Bosch, MNRAS **396**, 141 (2009).
- [125] D. N. Spergel and P. J. Steinhardt, PRL **84**, 3760 (2000).
- [126] K. Abazajian, G. M. Fuller, and M. Patel, Phys. Rev. D **64**, 023501 (2001).
- [127] J. L. Feng, A. Rajaraman, and F. Takayama, PRL **91**, 011302 (2003).
- [128] M. Kaplinghat, Phys. Rev. D **72**, 063510 (2005).
- [129] J. D. Simon, A. D. Bolatto, A. Leroy, L. Blitz, and E. L. Gates, ApJ **621**, 757 (2005).
- [130] R. Kuzio de Naray, S. S. McGaugh, and W. J. G. de Blok, ApJ **676**, 920 (2008).

- [131] W. J. G. de Blok *et al.*, AJ **136**, 2648 (2008).
- [132] N. Padmanabhan *et al.*, New Astronomy **9**, 329 (2004).
- [133] A. B. Newman *et al.*, ApJ **706**, 1078 (2009).
- [134] T. Treu and L. V. E. Koopmans, ApJ **611**, 739 (2004).
- [135] R. Gavazzi *et al.*, ApJ **667**, 176 (2007).
- [136] A. E. Schulz, R. Mandelbaum, and N. Padmanabhan, arXiv:0911.2260, 2009.
- [137] F. C. van den Bosch, A. Burkert, and R. A. Swaters, MNRAS **326**, 1205 (2001).
- [138] F. C. van den Bosch, T. Abel, R. A. C. Croft, L. Hernquist, and S. D. M. White, ApJ **576**, 21 (2002).
- [139] F. C. van den Bosch, T. Abel, and L. Hernquist, MNRAS **346**, 177 (2003).
- [140] J. Binney and S. Tremaine, *Galactic Dynamics: Second Edition* (Princeton University Press, 2008).
- [141] G. R. Blumenthal, S. M. Faber, R. Flores, and J. R. Primack, ApJ **301**, 27 (1986).
- [142] J. G. Stadel, *Cosmological N-body simulations and their analysis*, PhD thesis, AA(UNIVERSITY OF WASHINGTON), 2001.
- [143] Hernquist, L., Bouchet, F. R., & Suto, Y. 1991, ApJS, 75, 231
- [144] L. Hernquist, F. R. Bouchet and Y. Suto, ApJS **75**, 231 (1991).
- [145] D. N. Spergel *et al.*, ApJS **170**, 377 (2007).
- [146] N. Katz and S. D. M. White, ApJ **412**, 455 (1993).
- [147] A. M. Brooks, F. Governato, T. Quinn, C. B. Brook, and J. Wadsley, ApJ **694**, 396 (2009).
- [148] A. Pontzen *et al.*, MNRAS **402**, 1523 (2010).
- [149] J. I. Read, L. Mayer, A. M. Brooks, F. Governato, and G. Lake, MNRAS **397**, 44 (2009).
- [150] A. Zolotov *et al.*, ApJ **702**, 1058 (2009).
- [151] F. Governato *et al.*, MNRAS **398**, 312 (2009).
- [152] D. Reed *et al.*, MNRAS **346**, 565 (2003).
- [153] F. Governato *et al.*, MNRAS **374**, 1479 (2007).

- [154] S. P. D. Gill, A. Knebe, and B. K. Gibson, *MNRAS* **351**, 399 (2004).
- [155] S. R. Knollmann and A. Knebe, *ApJS* **182**, 608 (2009).
- [156] M. A. K. Gross, *Cdm-Variant Cosmological Models: Nonlinear Comparisons Including Large-Scale Constrained Realizations*, PhD thesis, AA(UNIVERSITY OF CALIFORNIA, SANTA CRUZ), 1997.
- [157] V. R. Eke, S. Cole, and C. S. Frenk, *MNRAS* **282**, 263 (1996).
- [158] J. S. Bullock *et al.*, *ApJ* **555**, 240 (2001).
- [159] S. D. M. White, *ApJ* **286**, 38 (1984).
- [160] J. Zavala, T. Okamoto, and C. S. Frenk, *MNRAS* **387**, 364 (2008).
- [161] A. H. G. Peter, C. E. Moody, and M. Kamionkowski, *Phys. Rev. D* **81**, 103501 (2010).
- [162] M. Valluri, V. P. Debattista, T. Quinn, and B. Moore, *MNRAS* **403**, 525 (2010).
- [163] A. Buonanno, 0709.4682.
- [164] T. L. Smith, E. Pierpaoli, and M. Kamionkowski, *PRL* **97**, 021301 (2006).
- [165] F. A. Jenet *et al.*, *ApJ* **653**, 1571 (2006).
- [166] B. P. Abbott and *et al.*, at **460**, 990 (2009).
- [167] M. Kamionkowski, A. Kosowsky, and A. Stebbins, *PRL* **78**, 2058 (1997).
- [168] U. Seljak and M. Zaldarriaga, *PRL* **78**, 2054 (1997).
- [169] M. D. Niemack *et al.*, in volume 7741 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, 2010, 1006.5049.
- [170] R. W. Ogburn, IV *et al.*, in volume 7741 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, 2010.
- [171] C. Bischoff *et al.*, *ApJ* **684**, 771 (2008).
- [172] R. Bustos, in *The Cool Universe: Observing Cosmic Dawn*, edited by C. Lidman & D. Alloin, volume 344 of *Astronomical Society of the Pacific Conference Series*, pp. 37–+, 2005.
- [173] C. E. North *et al.*, 0805.3690.
- [174] A. A. Fraisse *et al.*, 0811.3920.

- [175] P. Oxley *et al.*, edited by M. Strojnik, in volume 5543 of *Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference*, pp. 320–331, 2004, arXiv:astro-ph/0501111.
- [176] J. R. Hinderks *et al.*, *ApJ* **692**, 1221 (2009).
- [177] D. T. Chuss *et al.*, in volume 7741 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, 2010.
- [178] K. Arnold *et al.*, in volume 7741 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, 2010.
- [179] H. Kudoh, A. Taruya, T. Hiramatsu, and Y. Himemoto, *Phys. Rev. D* **73**, 064006 (2006).
- [180] P. Amaro-Seoane, S. Aoudia, S. Babak, P. Binetruy, E. Berti, A. Bohe, C. Caprini and M. Colpi *et al.*, arXiv:1201.3621 [astro-ph.CO].
- [181] S. Kawamura, M. Ando, N. Seto, S. Sato, T. Nakamura, K. Tsubono, N. Kanda and T. Tanaka *et al.*, *Class. Quant. Grav.* **28**, 094011 (2011).
- [182] G. M. Harry [LIGO Scientific Collaboration], *Class. Quant. Grav.* **27**, 084006 (2010).
- [183] M. Punturo, M. Abernathy, F. Acernese, B. Allen, N. Andersson, K. Arun, F. Barone and B. Barr *et al.*, *Class. Quant. Grav.* **27**, 194002 (2010).
- [184] N. Seto and A. Cooray, *Phys. Rev. D* **73**, 023005 (2006).
- [185] B. Allen, in *Relativistic Gravitation and Gravitational Radiation*, edited by J.-A. Marck & J.-P. Lasota, pp. 373–+, 1997.
- [186] M. Maggiore, *Physics Reports* **331**, 283 (2000).
- [187] “GAIA Concept and Technology Study Report 2000”, ESA-SCI **4** (2000), Available at [http://www.rssd.esa.int/SA-general/Projects/GAIA\\_files/FILES/report-march.pdf](http://www.rssd.esa.int/SA-general/Projects/GAIA_files/FILES/report-march.pdf).
- [188] V. B. Braginsky, N. S. Kardashev, A. G. Polnarev, and I. D. Novikov, *Nuovo Cimento B Serie* **105**, 1141 (1990).
- [189] T. Damour and G. Esposito-Farese, *Phys. Rev.* **D58**, 044003 (1998).
- [190] R. Fakir, *ApJ* **426**, 74 (1994).
- [191] E. E. Flanagan, *Phys. Rev. D* **48**, 2389 (1993), Note that the RHS of Eq. (2.8) should be divided by  $\pi$ .

- [192] C. R. Gwinn, T. M. Eubanks, T. Pyne, M. Birkinshaw, and D. N. Matsakis, *ApJ* **485**, 87 (1997).
- [193] A. H. Jaffe, *New Astron. Rev.* **48**, 1483 (2004).
- [194] N. Kaiser and A. Jaffe, *ApJ* **484**, 545 (1997).
- [195] E. V. Linder, *Phys. Rev. D* **34**, 1759 (1986).
- [196] V. V. Makarov, in *IAU Symposium*, volume 261 of *IAU Symposium*, pp. 345–349, 2010.
- [197] F. Mignard and S. A. Klioner, in *IAU Symposium*, volume 261 of *IAU Symposium*, pp. 306–314, 2010.
- [198] T. Pyne, C. R. Gwinn, M. Birkinshaw, T. M. Eubanks, and D. N. Matsakis, *ApJ* **465**, 566 (1996).
- [199] B. F. Schutz, American Astronomical Society, IAU Symposium #261. Relativity in Fundamental Astronomy: Dynamics, Reference Frames, and Data Analysis 27 April - 1 May 2009 Virginia Beach, VA, USA, #12.01; *Bulletin of the American Astronomical Society*, Vol. 41, p.888 **261**, 1201 (2009).
- [200] E. Fomalont and M. Reid, *New Astronomy Reviews* **48**, 1473 (2004).
- [201] O. Titov, S. B. Lambert, and A. Gontier, “VLBI measurement of the secular aberration drift”, 1009.3698.
- [202] E. V. Linder, *ApJ* **328**, 77 (1988).
- [203] S. Detweiler, *ApJ* **234**, 1100 (1979).
- [204] M. Anholm, S. Ballmer, J. D. E. Creighton, L. R. Price, and X. Siemens, *Phys. Rev. D* **79**, 084030 (2009).
- [205] J. D. E. Creighton, private communication, 2010.
- [206] J. Yoo, A. L. Fitzpatrick, and M. Zaldarriaga, *Phys. Rev. D* **80**, 083514 (2009).
- [207] R. M. Wald, *General relativity* (Chicago, University of Chicago Press, 1984, 504 p., 1984).
- [208] Fixsen, D. J., Cheng, E. S., Gales, J. M., et al. 1996, *ApJ*, 473, 576
- [209] G. F. Smoot *et al.*, *ApJ* **396**, L1 (1992).
- [210] A. Lewis and A. Challinor, *Phys. Rept.* **429**, 1 (2006). [arXiv:astro-ph/0601594].
- [211] S. M. Carroll, G. B. Field and R. Jackiw, *Phys. Rev. D* **41**, 1231 (1990).

- [212] R. R. Caldwell, V. Gluscevic and M. Kamionkowski, arXiv:1104.1634 [astro-ph.CO].
- [213] A. de Oliveira-Costa, M. Tegmark, M. Zaldarriaga and A. Hamilton, Phys. Rev. D **69**, 063516 (2004) [arXiv:astro-ph/0307282].
- [214] F. K. Hansen, A. J. Banday and K. M. Gorski, MNRAS **354**, 641 (2004) [arXiv:astro-ph/0404206].
- [215] D. Hanson and A. Lewis, Phys. Rev. D **80**, 063004 (2009) [arXiv:0908.0963 [astro-ph.CO]].
- [216] L. Verde, L. M. Wang, A. Heavens and M. Kamionkowski, MNRAS **313**, L141 (2000) [arXiv:astro-ph/9906301].
- [217] E. Komatsu and D. N. Spergel, Phys. Rev. D **63**, 063002 (2001) [arXiv:astro-ph/0005036].
- [218] M. Kunz, A. J. Banday, P. G. Castro, P. G. Ferreira and K. M. Gorski, arXiv:astro-ph/0111250.
- [219] S. Furlanetto, S. P. Oh and F. Briggs, Phys. Rept. **433**, 181 (2006) [arXiv:astro-ph/0608032].
- [220] J. D. Bowman and A. E. E. Rogers, at **468**, 796 (2010); G. Paciga *et al.* (GMRT-EoR Collaboration), Mon. Not. R. Astron. Soc. **413**, 1174 (2011); [www.lofar.org](http://www.lofar.org); [www.mwatelescope.org](http://www.mwatelescope.org); [web.phys.cmu.edu/past](http://web.phys.cmu.edu/past); [www.phys.unm.edu/lwa](http://www.phys.unm.edu/lwa); [astro.berkeley.edu/dbacker/eor](http://astro.berkeley.edu/dbacker/eor); [www.haystack.mit.edu/ast/arrays/Edges](http://www.haystack.mit.edu/ast/arrays/Edges).
- [221] S. Jester and H. Falcke, New Astron. Rev. **53**, 1 (2009) [arXiv:0902.0493 [astro-ph.CO]].
- [222] A. Loeb and M. Zaldarriaga, PRL **92**, 211301 (2004) [arXiv:astro-ph/0312134].
- [223] M. F. Morales and J. S. B. Wyithe, Annu. Rev. Astro. Astrophys. **48**, 127 (2010) [arXiv:0910.3010 [astro-ph.CO]].
- [224] D. Walsh, R. F. Carswell and R. J. Weymann, Nature **279**, 381 (1979).
- [225] M. Limousin, E. Jullo, J. Richard, R. Cabanac, J. -P. Kneib, R. Gavazzi and G. Soucail, arXiv:0906.4118 [astro-ph.CO].
- [226] E. S. Cypriano, G. B. L. Neto, L. Sodre, Jr. and J. P. Kneib, ApJ **630**, 38 (2005). [astro-ph/0504036].
- [227] M. Hattori, J. P. Kneib and N. Makino, Prog. of Theor. Phys. Supp. **133**, 1 (1999) [astro-ph/9905009].
- [228] M. Postman, D. Coe, N. Benitez, L. Bradley, T. Broadhurst, M. Donahue, H. Ford and O. Graur *et al.*, ApJS **199**, 25 (2012) [arXiv:1106.3328 [astro-ph.CO]].
- [229] C. Alcock *et al.* [MACHO Collaboration], ApJ **542**, 281 (2000) [astro-ph/0001272].

- [230] P. Tisserand *et al.* [EROS-2 Collaboration], *A&A* **469**, 387 (2007) [astro-ph/0607207].
- [231] M. Moniez, *Gen. Rel. Grav.* **42**, 2047 (2010) [arXiv:1001.2707 [astro-ph.GA]].
- [232] F. Li, A. L. Erickcek and N. M. Law, arXiv:1202.1284 [astro-ph.CO].
- [233] R. D. Blandford *et al.*, *MNRAS* **251**, 600 (1991); J. Miralda-Escudé, *ApJ* **380**, 1 (1991); N. Kaiser, *ApJ* **388**, 272 (1992); M. Bartelmann and P. Schneider, *Astron. Astrophys.* **259**, 413 (1992); M. Bartelmann and P. Schneider, *Phys. Rept.* **340**, 291 (2001) [astro-ph/9912508]; P. Schneider, *MNRAS*, 283, 837 (1996).
- [234] D. J. Bacon, A. R. Refregier and R. S. Ellis, *MNRAS* **318**, 625 (2000) [astro-ph/0003008]; N. Kaiser, G. Wilson and G. A. Luppino, *ApJ* **556**, 601 (2001) [astro-ph/0003338].  
at **405**, 143 (2000) [astro-ph/0003014]; L. van Waerbeke *et al.*, *A&A* **358**, 30 (2000) [astro-ph/0002500].
- [235] The Dark Energy Survey Collaboration 2005, arXiv:astro-ph/0510346.
- [236] Schrabback, T., Hartlap, J., Joachimi, B., et al., *A&A*, 516, A63 (2010).
- [237] Huff, E. M., Eifler, T., Hirata, C. M., et al. (2011), arXiv:1112.3143.
- [238] K. M. Smith, O. Zahn and O. Dore, *Phys. Rev. D* **76**, 043510 (2007) [arXiv:0705.3980 [astro-ph]]; C. M. Hirata *et al.*, *Phys. Rev. D* **78**, 043520 (2008) [arXiv:0801.0644 [astro-ph]].
- [239] S. Das *et al.* (ACT Collaboration), *PRL* **107**, 021301 (2011) [arXiv:1103.2124 [astro-ph.CO]].
- [240] M. Zaldarriaga, U. Seljak, *Phys. Rev.* **D59**, 123507 (1999) [astro-ph/9810257]; U. Seljak and M. Zaldarriaga, *PRL* **82**, 2636 (1999) [astro-ph/9810092]; W. Hu, *Phys. Rev.* **D64**, 083005 (2001) [astro-ph/0105117]. *ApJ* **557**, L79 (2001) [astro-ph/0105424]; W. Hu and T. Okamoto, *ApJ* **574**, 566 (2002) [astro-ph/0111606]; M. H. Kesden, A. Cooray and M. Kamionkowski, *Phys. Rev.* **D67**, 123507 (2003) [astro-ph/0302536]; A. Lewis and A. Challinor, *Phys. Rept.* **429**, 1 (2006) [astro-ph/0601594]; S. Dodelson, F. Schmidt and A. Vallinotto, *Phys. Rev. D* **78**, 043508 (2008) [arXiv:0806.0331 [astro-ph]].
- [241] N. Kaiser and A. H. Jaffe, *ApJ* **484**, 545 (1997) [arXiv:astro-ph/9609043]; S. Dodelson, E. Rozo and A. Stebbins, *PRL* **91**, 021301 (2003) [arXiv:astro-ph/0301177]; C. Li and A. Cooray, *Phys. Rev. D* **74**, 023521 (2006) [arXiv:astro-ph/0604179]; D. Sarkar *et al.*, *Phys. Rev.* **D77**, 103515 (2008) [arXiv:0803.1490 [astro-ph]]; S. Dodelson, *Phys. Rev. D* **82**, 023522 (2010) [arXiv:1001.5012 [astro-ph.CO]]; T. Namikawa, D. Yamauchi and A. Taruya, arXiv:1110.1718 [astro-ph.CO].

- [242] A. Cooray, M. Kamionkowski and R. R. Caldwell, Phys. Rev. D **71**, 123527 (2005) [arXiv:astro-ph/0503002].
- [243] L. G. Book, M. Kamionkowski and T. Souradeep, Phys. Rev. D, in press, arXiv:1109.2910 [astro-ph.CO].
- [244] A. Stebbins, [astro-ph/9609149]; M. Kamionkowski *et al.*, MNRAS **301**, 1064 (1998) [astro-ph/9712030].
- [245] A. Hajian and T. Souradeep, Astrophys.J. **597**, L5,(2003) [arXiv:astro-ph/0308001].
- [246] A. Hajian and T. Souradeep, [arXiv:astro-ph/0501001].
- [247] N. Joshi, S. Jhingan, T. Souradeep and A. Hajian, Phys. Rev. D **81**, 083012 (2010) [arXiv:0912.3217 [astro-ph.CO]].
- [248] A. Hajian and T. Souradeep, [arXiv:astro-ph/0301590].
- [249] A. R. Pullen and M. Kamionkowski, Phys. Rev. D **76**, 103529 (2007) [arXiv:0709.1144 [astro-ph]].
- [250] A. Hajian, T. Souradeep and N. J. Cornish, ApJ **618**, L63 (2004) [arXiv:astro-ph/0406354].
- [251] T. Souradeep and A. Hajian, Pramana **62**, 793 (2004) [arXiv:astro-ph/0308002].
- [252] T. Souradeep, A. Hajian and S. Basak, New Astron. Rev. **50**, 889 (2006) [arXiv:astro-ph/0607577].
- [253] A. Hajian and T. Souradeep, Phys. Rev. D **74**, 123521 (2006) [arXiv:astro-ph/0607153].
- [254] T. Ghosh, A. Hajian and T. Souradeep, Phys. Rev. D **75**, 083007 (2007) [arXiv:astro-ph/0604279].
- [255] A. Hajian, T. Souradeep and N. J. Cornish, ApJ **618**, L63 (2004) [arXiv:astro-ph/0406354].
- [256] D. Hanson, A. Lewis and A. Challinor, Phys. Rev. D **81**, 103003 (2010) [arXiv:1003.0198 [astro-ph.CO]].
- [257] C. L. Bennett *et al.*, [arXiv:1001.4758 [astro-ph.CO]].
- [258] S. Basak, A. Hajian and T. Souradeep, Phys. Rev. D **74**, 021301 (2006) [arXiv:astro-ph/0603406].
- [259] M. Kamionkowski, PRL **102**, 111302 (2009) [arXiv:0810.1286 [astro-ph]].
- [260] V. Gluscevic, M. Kamionkowski and A. Cooray, Phys. Rev. D **80**, 023510 (2009) [arXiv:0905.1687 [astro-ph.CO]].

- [261] A. P. S. Yadav, R. Biswas, M. Su and M. Zaldarriaga, Phys. Rev. D **79**, 123009 (2009) [arXiv:0902.4466 [astro-ph.CO]].
- [262] T. Souradeep, Indian J. Phys. **80**, 1063 (2006) [arXiv:gr-qc/0609026].
- [263] M. Aich and T. Souradeep, Phys. Rev. D **81**, 083008 (2010) [arXiv:1001.1723 [astro-ph.CO]].
- [264] T. Souradeep and A. Hajian, [arXiv:astro-ph/0502248].
- [265] L. Ackerman, S. M. Carroll and M. B. Wise, Phys. Rev. D **75**, 083502 (2007) [arXiv:astro-ph/0701357].
- [266] W. Hu, Phys. Rev. D **62**, 043007 (2000) [arXiv:astro-ph/0001303].
- [267] A. Stebbins, [arXiv:astro-ph/9609149].
- [268] N. Kaiser and A. H. Jaffe, ApJ **484**, 545 (1997) [arXiv:astro-ph/9609043].
- [269] S. Dodelson, E. Rozo and A. Stebbins, PRL **91**, 021301 (2003) [arXiv:astro-ph/0301177].
- [270] S. Dodelson, Phys. Rev. D **82**, 023522 (2010) [arXiv:1001.5012 [astro-ph.CO]].
- [271] C. Li and A. Cooray, Phys. Rev. D **74**, 023521 (2006) [arXiv:astro-ph/0604179].
- [272] M. Kamionkowski, A. Kosowsky and A. Stebbins, Phys. Rev. D **55**, 7368 (1997) [arXiv:astro-ph/9611125].
- [273] M. Kamionkowski, A. Kosowsky and A. Stebbins, PRL **78**, 2058 (1997) [arXiv:astro-ph/9609132].
- [274] L. G. Book, E. E. Flanagan, Phys. Rev. **D83**, 024024 (2011) [arXiv:1009.4192 [astro-ph.CO]].
- [275] J. M. Bardeen, J. R. Bond, N. Kaiser, A. S. Szalay, ApJ **304**, 15-61 (1986).
- [276] E. Komatsu *et al.* [ WMAP Collaboration ], ApJS **192**, 18 (2011) [arXiv:1001.4538 [astro-ph.CO]].
- [277] A. Lue, L. M. Wang and M. Kamionkowski, PRL **83**, 1506 (1999) [arXiv:astro-ph/9812088].
- [278] C. R. Contaldi, J. Magueijo and L. Smolin, PRL **101**, 141101 (2008) [arXiv:0806.3082 [astro-ph]].
- [279] T. Takahashi and J. Soda, PRL **102**, 231301 (2009) [arXiv:0904.0554 [hep-th]].
- [280] V. Gluscevic and M. Kamionkowski, Phys. Rev. D **81**, 123529 (2010) [arXiv:1002.1308 [astro-ph.CO]].

- [281] M. Zaldarriaga and U. Seljak, Phys. Rev. D **55**, 1830 (1997) [arXiv:astro-ph/9609170]; U. Seljak and M. Zaldarriaga, PRL **78**, 2054 (1997) [arXiv:astro-ph/9609169].
- [282] P. Cabella and M. Kamionkowski, [arXiv:astro-ph/0403392].
- [283] J. R. Pritchard and M. Kamionkowski, Annals Phys. **318**, 2 (2005) [arXiv:astro-ph/0412581].
- [284] M. Kamionkowski and T. Souradeep, Phys. Rev. D **83**, 027301 (2011) [arXiv:1010.4304 [astro-ph.CO]].
- [285] <http://camb.info/>
- [286] K. W. Masui and U. L. Pen, PRL **105**, 161302 (2010) [arXiv:1006.4181 [astro-ph.CO]].
- [287] K. Sigurdson and A. Cooray, PRL **95**, 211303 (2005) [arXiv:astro-ph/0502549]; O. Zahn and M. Zaldarriaga, ApJ **653**, 922 (2006) [arXiv:astro-ph/0511547].
- [288] N. Kogo and E. Komatsu, Phys. Rev. D **73**, 083007 (2006) [arXiv:astro-ph/0602099].
- [289] J. Bock *et al.* [EPIC Collaboration], arXiv:0906.1188 [astro-ph.CO].
- [290] M. Zaldarriaga and U. Seljak, Phys. Rev. **D58**, 023003 (1998) [astro-ph/9803150]; A. Lewis, A. Challinor and N. Turok, Phys. Rev. **D65**, 023505 (2002) [astro-ph/0106536].
- [291] M. Kesden, A. Cooray and M. Kamionkowski, PRL **89**, 011304 (2002) [arXiv:astro-ph/0202434].
- [292] L. Knox and Y. S. Song, PRL **89**, 011303 (2002) [arXiv:astro-ph/0202286].
- [293] M. Kamionkowski, A. Babul, C. M. Cress and A. Refregier, MNRAS **301**, 1064 (1998) [arXiv:astro-ph/9712030].
- [294] J. F. Navarro, C. S. Frenk, and S. D. M. White, “A Universal Density Profile from Hierarchical Clustering”, ApJ **490**, 493 (1997).
- [295] L. D. Landau and E. M. Lifshitz, *Fluid Mechanics* (Pergamon, London, 1959).