Inflation, Large-Scale Structure and Inhomogeneous Cosmologies

Seshadri Nadathur
Merton College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy

Trinity Term 2011

To my parents
who will be pleased I have reached this point.

Acknowledgements

First of all, I must thank my supervisor Subir Sarkar for his support and encouragement during my time at Oxford, and the many opportunities he has provided me with. Without his help, there would obviously have been no thesis.

I am particularly grateful to Shaun Hotchkiss for collaborations and many illuminating discussions about physics, and also for providing me with an example of what to do—and occasionally what not to do!—in pursuit of a D.Phil. I am also grateful to Anupam Mazumdar for introducing me to many new ideas and for help with the joint work that is part of this thesis.

Outside the physics world, I certainly would not have got through this D.Phil., or indeed many other challenges over the last few years, without Iza, whose love and companionship have kept me sane, and whose patience inspires me—especially since I can't have been much fun to be with for the last few months!

Many thanks also to the members of the OUMC: for the friendships that made Oxford life so enjoyable, for the climbing weekends that kept me physically active, and for the mountaineering trips that put the D.Phil. worries in perspective. I have also been lucky to have some wonderful housemates over the last few years in Lena, Rob and Aeron. Thanks for the many shared meals, shared interests and great conversation, and thanks also to Aeron for proof-reading assistance with this document.

And finally, I'd like to thank every one of the many non-physicists—friends, family, acquaintances, or just chance encounters—who have at any point over the last four years asked me questions about black holes, time travel, string theory, the LHC, the nature of space-time, the origin of the Universe, or any other aspect of modern physics, for reminding me what an exciting and fascinating subject this is. I hope my answers didn't leave you more confused than before.

My work was supported by the Clarendon Fund and a Domus A scholarship from Merton College.

Publications

Chapter 2 is based on research coauthored with Shaun Hotchkiss and Anupam Mazumdar, which was published in *JCAP*, 1106 002 (2011). The contents of Chapter 3 have been modified from work coauthored with Subir Sarkar and published in *Phys.Rev.* **D83** 063506 (2011). Chapter 4 is based on a paper, coauthored with Shaun Hotchkiss and Subir Sarkar, which has been submitted for publication, and is available on arXiv.org as arXiv:1109.4162 [astro-ph.CO].

Abstract

Determining cosmological parameters from current observational data requires knowledge of the primordial density perturbations generated during inflation. We begin by examining a model of inflation along a flat direction of the minimal supersymmetric Standard Model (MSSM) and the power spectrum of perturbations it can produce. We consider the fine-tuning issues associated with this model and discuss a modification of the potential to include a hybrid transition that reduces the fine-tuning, without affecting the viability of the model. However, supersymmetric flat directions might play a role in other models of inflation as well. In particular, they may cause a feature in the primordial power spectrum of perturbations, unlike the scale-free spectrum assumed in the standard Lambda Cold Dark Matter (ΛCDM) cosmological model. We then show that in the presence of such a feature, an alternative cosmological model with a large local void and no dark energy provides a good fit to both Type Ia supernovae and the cosmic microwave background (CMB) data from the WMAP satellite. Constraints from the locally measured Hubble parameter, baryon acoustic oscillations and primordial nucleosynthesis are also satisfied. This degeneracy motivates a search for other independent observational tests of ΛCDM. The integrated Sachs-Wolfe (ISW) imprint of large-scale structure on the CMB is one such test. The ISW imprint of superstructures of size $\sim 100\ h^{-1}$Mpc at redshift $z \sim 0.5$ has been detected with $> 4\sigma$ significance, however it has been noted that the signal is much larger than expected. We revisit the calculation using linear theory predictions in a ΛCDM cosmology and find the theoretical prediction is inconsistent by $> 3\sigma$ with the observation. If the observed signal is indeed due to the ISW effect then huge, extremely underdense voids are far more common in the observed universe than predicted by ΛCDM.

Contents

List of Figures

Chapter 1

Introduction

The current standard model of cosmology is the Lambda Cold Dark Matter model (ΛCDM). It has been shown to be in impressive agreement with data from a wide range of different observational probes of our universe, and hence is sometimes also known as the "concordance" model. Although consistent in most of its details, it still leaves some questions unanswered, the most important of which are to do with the nature of the mysterious "dark energy" that appears to make up $\sim 70\%$ of the current energy budget of the universe. It is also true that although the theory of inflation has been accepted as the standard paradigm for the generation of the seed perturbations for the origin of structure in the universe, there is an abundance of well-founded inflationary models and as yet no unique underlying model can be said to be favoured.

This thesis relates to both inflation at early times and the make-up of our current universe. In fact neither can be determined uniquely without assumptions about the other, so in fact both directions of investigation are intimately linked. We start in this chapter by reviewing the standard theory of the cosmological model, the evolution of cosmological perturbations, and the generation of these perturbations during inflation. We will highlight some of the observational evidence in favour of ΛCDM, which serves as a test for any alternative model, and areas where future observations may constrain deviations from the standard view.

1.1 Overview of Big Bang Cosmology

The universe we observe today is expanding. We see evidence of this expansion in the cosmic microwave background (CMB) radiation and the relative abundances of light elements, from which we can deduce that the universe has evolved from a hot, dense state, with a temperature of at least $\sim 10^{10}$ K, the temperature at which nucleosynthesis occurred [4]. Since then it has been expanding and cooling. We can trace this expansion backwards in time to form a picture of the history of the universe. The main events along this timeline are as follows.

1. Quantum gravity era

   It is usually assumed that the universe extends back in time all the way to Planck scale energies, $M_P \sim 10^{19}$ GeV. At this scale quantum corrections to General Relativity become important and the notion of an underlying metric of space-time itself becomes unclear, so any theoretical picture of the universe is necessarily very speculative. This is usually taken as the 'beginning' of the universe, $t = 0$.

2. Inflation

   The history of the universe in the regime of energy densities between 10^{19} GeV and 100 GeV (the scale of the electroweak phase transition) is also speculative, because the laws of physics at these energies are currently unknown. However, certain reasonable assumptions may be made by extrapolating from known physical theories. The theory of inflation—a period of accelerated expansion of the universe, which sets the initial conditions for the hot Big Bang model—is now accepted as a standard paradigm.

3. The hot Big Bang

   At some point after the end of inflation, the universe will have become dominated by radiation and reached thermal equilibrium. This should have occurred by the time the temperature reached 100 GeV, and from this point on the physics is reasonably well understood. The electroweak phase transition occurs at ~ 100 GeV, and the QCD phase transition at temperatures of $100 - 300$ MeV. The baryon asymmetry of the universe must have been generated before the electroweak transition.

4. Nucleosynthesis

   When the temperature dropped to 1 MeV, neutrinos decoupled from the thermal bath of radiation. Subsequently, at temperatures of ~ 0.1 MeV, the light elements were formed from protons and neutrons in the process known as nucleosynthesis. The physics of nucleosynthesis is very well understood, and observation of the relative abundance of the light elements provides us with the earliest constraints on the standard cosmological model.

5. Matter-radiation equality

   As the universe expanded, the density of non-relativistic species (matter) will have grown relative to that of radiation. At temperatures of ~ 1 eV these densities became equal, and subsequently the matter-dominated epoch began, marking the start of the formation of structure.

6. Photon decoupling and formation of the CMB

   When the universe was $\sim 10^{13}$ seconds old and the temperature had dropped to 0.1 eV, free electrons bound with the atomic nuclei to form neutral atoms. This process is known as recombination. At this point Thomson scattering of photons off electrons can no longer occur and so the radiation field decoupled from matter and photons propagated freely through the universe. These photons are observed today as the CMB.

7. Reionization

   The radiation produced by the first massive structures formed in the early universe began to reionize the neutral hydrogen gas in the universe, possibly as early as 1 Gyr after the Big Bang. The process of reionization was complete by a redshift $z \approx 6$.

8. Structure formation

   The first large structures in the universe formed at around 1 Gyr. The non-linear growth of these structures under gravity leads to the formation of individual galaxies, clusters of galaxies, and voids.

1.1.1 The FRW metric

The starting point for a mathematical description of the universe is Einstein's General Theory of Relativity (GR). According to GR, the curvature of space-time is related to the matter content of the universe by the Einstein field equations

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = 8\pi G T_{\mu\nu} \, . \tag{1.1}$$

In this equation, $g_{\mu\nu}$ is the metric describing the space-time;[1] $R_{\mu\nu}$ is the Ricci tensor

$$R_{\mu\nu} = \frac{\partial \Gamma^{\lambda}_{\mu\nu}}{\partial x^{\lambda}} - \frac{\partial \Gamma^{\lambda}_{\lambda\mu}}{\partial x^{\nu}} + \Gamma^{\lambda}_{\mu\nu}\Gamma^{\sigma}_{\lambda\sigma} - \Gamma^{\lambda}_{\mu\sigma}\Gamma^{\sigma}_{\nu\lambda} \, , \tag{1.2}$$

where $\Gamma^{\lambda}_{\mu\nu}$ is the affine connection; $R \equiv g^{\mu\nu}R_{\mu\nu}$ is the Ricci scalar; G is Newton's gravitational constant and $T_{\mu\nu}$ is the energy-momentum tensor.

Solving the equations above in general is highly complex. In practice one must start with an ansatz for the metric $g_{\mu\nu}$ in order to simplify the task. We take as a starting point the observation that the visible universe is highly isotropic around us, at least on large enough scales. We then make the assumption that we are not at a special place in the universe, and that the universe must therefore appear isotropic to observers at all locations. This is known as the Cosmological Principle, and is equivalent to assuming that the universe is—at least on large enough scales—spatially homogeneous and isotropic. Although the Cosmological Principle is aesthetically appealing, it is important to remember that it is merely an assumption that is consistent with our prejudices, and should not be taken as fact in the absence of conclusive evidence for homogeneity. We shall discuss this issue again later.

The assumption of homogeneity and isotropy allows us to choose a particularly simple form for the space-time metric, for which the Friedmann-Robertson-Walker (FRW) line element is

$$ds^2 \equiv g_{\mu\nu}dx^{\mu}dx^{\nu} = -dt^2 + a^2(t)\left[\frac{dr^2}{1-kr^2} + r^2 d\theta^2 + r^2\sin^2\theta d\phi^2\right] \, , \tag{1.3}$$

in units where the speed of light is $c = 1$. The constant k can take the values 0, ±1. For $k = 0$ the 3-d space-like hypersurface will be flat ($\mathbb{R}^3$); $k = 1$ gives a space of positive curvature, $\mathcal{S}^3$; and $k = -1$ corresponds to the hyperbolic space of negative curvature, $\mathcal{H}^3$. Here $a(t)$ is an arbitrary function of time, known as the scale factor,

---

[1] We use the $- + ++$ convention for defining the metric.

and the coordinates x^i or $\vec{r}$ are called the spatial comoving coordinates. A particle at rest in these coordinates will remain at rest, and is thus "co-moving with the flow of the universe". Proper time along such a geodesic corresponds to the universal time, t.

In some situations it is preferable to use conformal time τ, defined by $d\tau = dt/a(t)$, in which case the line element becomes

$$ds^2 = a^2(t)\left[-d\tau^2 + \frac{dr^2}{1-kr^2} + r^2 d\theta^2 + r^2\sin^2\theta d\phi^2\right] \; . \tag{1.4}$$

Where necessary, we use an overdot to denote derivatives with respect to t and a prime ($'$) to denote derivatives with respect to conformal time.

However, while the comoving distance may remain constant, the proper distance between the origin and a comoving observer at radial coordinate r does change with time,

$$d_p(r,t) = a(t)\int_0^r \frac{dr}{\sqrt{1-kr^2}} \; . \tag{1.5}$$

Therefore even though the comoving coordinate of the object remains constant, the proper distance from us to the object increases (or decreases) with time. This is true of any two comoving objects: it is in this sense that the universe is expanding when $a(t)$ is increasing. The recessional velocity due to the expansion of the universe between any two galaxies that remain at fixed comoving coordinates is

$$v = \frac{\dot{a}}{a} d_p = H(t) d_p \; , \tag{1.6}$$

where $\dot{a} \equiv \frac{da}{dt}$ and $H(t) \equiv \frac{\dot{a}(t)}{a(t)}$ is the Hubble rate.

Let us consider an FRW coordinate system in which we are at the origin. A light ray passing through the origin in a radial ($d\theta = d\phi = 0$) direction travels along the geodesic $ds^2 = 0$ and so

$$dt = \pm a(t)\frac{dr}{\sqrt{1-kr^2}} \; , \tag{1.7}$$

where the plus (minus) sign indicates a photon moving away from (towards) the origin. Hence the light emitted at time t_1 from a source at comoving coordinate r_1 arrives at the origin at time t_0, where

$$\int_{t_1}^{t_0} \frac{dt}{a(t)} = \int_0^{r_1} \frac{dr}{\sqrt{1-kr^2}} \; . \tag{1.8}$$

Taking the differential form of this equation, we can show that the time interval δt_1 between subsequent wave crests on emission is related to the time interval δt_0 at observation by

$$\frac{\delta t_1}{a(t_1)} = \frac{\delta t_0}{a(t_0)} \, , \tag{1.9}$$

and hence if $a(t)$ is increasing, the observed frequency $\nu_0 = 1/\delta t_0$ is redshifted compared to the emitted frequency $\nu_1 = 1/\delta t_1$. The redshift, z, is defined to be

$$1 + z \equiv \frac{\nu_1}{\nu_0} = \frac{a(t_0)}{a(t_1)} \, . \tag{1.10}$$

By convention the scale factor is normalised such that $a(t_0) = 1$.

Two other important quantities remain to be defined. The particle horizon, $d_H(t)$, is the maximum proper distance that light can have travelled from the Big Bang at $t = t_{\mathrm{start}}$ to time t, and is given by

$$d_H(t) = a(t) \int_{t_{\mathrm{start}}}^{t} \frac{dt'}{a(t')} \, . \tag{1.11}$$

The Hubble distance, H^{-1}, also known as the horizon, provides a rough estimate of the distance that light can travel in the time that it takes the universe to expand appreciably; it is therefore an estimate of the scale over which causal processes can occur.

1.1.2 The expanding universe

At large enough scales, if we assume that the clustering of matter into structures and voids can be neglected, we may treat the contents of the universe as fluid. The assumption of homogeneity and isotropy imposes the following form for the energy-momentum tensor:

$$T^0{}_0 = -\rho(t), \quad T^i{}_j = p(t)\delta^i{}_j, \tag{1.12}$$

with all other components of $T_{\mu\nu}$ zero. Here $\rho(t)$ denotes the density of the fluid and $p(t)$ its pressure, both of which are constant along the space-like hypersurface. As we have noted in Sectionch1:sec:FRWmetric, the assumption of homogeneity and isotropy of the background $T_{\mu\nu}$ is an aesthetic choice that is not forced on us by observational data. In Chapter 3 we instead postulate a matter distribution that is isotropic but radially inhomogeneous, and show that, given other reasonable assumptions, such a model is also compatible with observational data. In Chapter 4 we approach the question for a different perspective and ask whether the predictions for the scale of homogeneity in the standard cosmological model are themselves consistent with observation. However, for the standard cosmological results presented in this chapter we shall continue to assume that the universe is, at least to a first approximation, homogeneous and isotropic.

Using the FRW metric to calculate the other quantities in Eqs. (1.1), we obtain the Friedmann equation for the evolution of the scale factor $a(t)$

$$H^2(t) \equiv \left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{k}{a^2}\ , \tag{1.13}$$

the Raychaudhuri equation

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3p)\ , \tag{1.14}$$

and the conservation law

$$\dot{\rho} + \frac{3\dot{a}}{a}(\rho + p) = 0\ . \tag{1.15}$$

Given an equation of state specifying p as a function of ρ, these equations may be solved to find ρ as a function of a and a as a function of t.

For the case of a perfect fluid, $p = w\rho$, where w is a constant, known as the equation of state parameter. Three interesting cases that we will have use for later are:

1. Non-relativistic matter (dust): In this case there is no pressure, $w = 0$ and so from Eq. (1.15), $\rho \propto a^{-3}$, i.e., the energy density is simply diluted by the expansion of the universe;

2. Relativistic matter (radiation): $p = \frac{1}{3}\rho \Rightarrow a^{-4}$, i.e., the energy density is diluted by both the expansion and the redshift;

3. Vacuum energy or cosmological constant: $p = -\rho \Rightarrow \rho = \text{const}$. Unlike dust or radiation, the effect of a cosmological constant does not get diluted with the expansion of the universe and hence if it is present, it will eventually come to dominate over other contributions at late times.

Returning to Eq. (1.13), for any current value of the Hubble constant $H_0 \equiv \dot{a}(t_0)/a(t_0)$, we may define the critical density as

$$\rho_c \equiv \frac{3H_0^2}{8\pi G} = 1.878 \times 10^{-29}\ h^2\ \mathrm{g/cm}^3\ , \tag{1.16}$$

where h is the Hubble constant in units of 100 km s^{-1} Mpc^{-1}. Then for a flat universe ($k = 0$) we may write the Friedmann equation as

$$\frac{H^2(t)}{H_0^2} = \frac{\rho(t)}{\rho_c}\ , \tag{1.17}$$

where the energy density ρ includes the contributions from all species: matter, radiation and vacuum energy. More generally, for arbitrary k and denoting the present energy densities in matter, radiation and vacuum energy as ρ_m, ρ_r and ρ_Λ, respectively, the Friedmann equation can be written as

$$H^2(t) = H_0^2 \left[\Omega_m \left(\frac{a_0}{a}\right)^3 + \Omega_r \left(\frac{a_0}{a}\right)^4 + \Omega_k \left(\frac{a_0}{a}\right)^2 + \Omega_\Lambda\right]\ , \tag{1.18}$$

where we have defined the density parameters

$$\Omega_m = \frac{\rho_m}{\rho_c},\ \Omega_r = \frac{\rho_r}{\rho_c},\ \Omega_\Lambda = \frac{\rho_\Lambda}{\rho_c},\ \Omega_k = -\frac{k}{a_0^2 H_0^2}\ . \tag{1.19}$$

By construction, the density parameters must satisfy the cosmic sum rule

$$\Omega_m + \Omega_r + \Omega_\Lambda + \Omega_k = 1\ . \tag{1.20}$$

At the present time, radiation makes up a very small fraction of the total energy density of the universe, $\Omega_r \approx 10^{-5}$, and hence can safely be neglected.

1.1.3 Distance measures

The distances between us and various astronomical or cosmological objects may be measured in different ways. One measure is the *comoving distance* between us and an object at scale factor a,

$$d_c(a) = \int_{t(a)}^{t_0} \frac{dt'}{a(t')} = \int_a^1 \frac{da'}{a'^2 H(a')}\ . \tag{1.21}$$

Another way to determine distances is to measure the angle θ subtended by an object of known physical size l, and define the *angular diameter distance* $d_A = l/\theta$. Inspection of the FRW metric shows that a source at comoving radial coordinate r that emits light at time t, and is observed today to subtend an angle θ on the sky, extends over a proper distance $l = a(t)r\theta$ normal to the line of sight. Therefore the angular diameter distance of the object is

$$d_A = a(t)r \; . \tag{1.22}$$

If the observed object is not of known physical size but of known absolute luminosity L, we instead define the *luminosity distance*, d_L, such that the observed flux from the object is

$$F(a) = \frac{L}{4\pi d_L(a)^2} \; . \tag{1.23}$$

It can easily be shown that this definition implies that the luminosity distance satisfies

$$d_L = a(t_0)r(1+z) = (1+z)\int_0^z \frac{dz}{H_0\sqrt{\Omega_m(1+z)^3 + \Omega_k(1+z)^2 + \Omega_\Lambda}} \; , \tag{1.24}$$

where the second equality holds only for a a flat Universe ($k = 0$). This means that the angular diameter distance and luminosity distance satisfy the following relationship:

$$\frac{d_L}{d_A} = (1+z)^2 \; . \tag{1.25}$$

For objects with $z \ll 1$, we can rewrite Eq. (1.10) as a power series expansion in the "look-back time" $t_0 - t_1$. On inverting this expression, after some manipulation, we obtain the following relationship between d_L and z, known as the Mattig formula:

$$H_0 d_L = z + \frac{1}{2}(1 - q_0)z^2 + \ldots \tag{1.26}$$

to second order. Here q_0 is the deceleration parameter, $q_0 = -(\frac{\ddot{a}}{aH^2})|_{t=t_0}$, whose value depends (via Eqs. (1.13) and (1.14)) on the contents of the universe

$$q_0 = \frac{\Omega_m}{2} - \Omega_\Lambda \; . \tag{1.27}$$

Thus if objects of known absolute luminosity can be found in the universe then from a plot of their measured redshifts and luminosity distances (known as a "Hubble diagram"), we can calibrate the value of H_0 and also determine the value of the deceleration parameter, thus constraining a linear combination of Ω_m and Ω_Λ. Such objects are known as "standard candles".

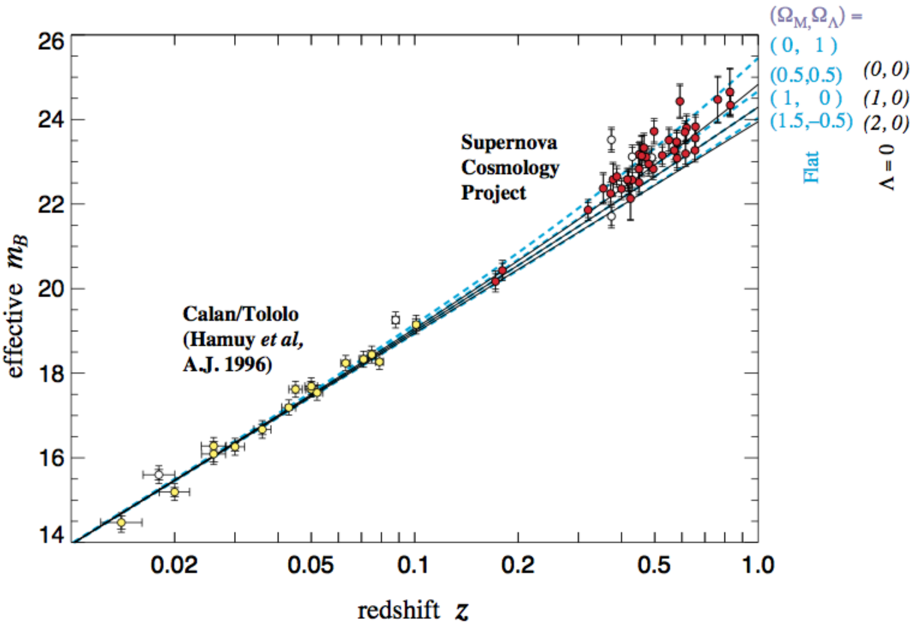

Figure 1.1: Evidence for dark energy found in 1998 by the Supernova Cosmology Project [1]. The figure shows the effective rest-frame B-band magnitude of 42 high-redshift SNe Ia observed by the Supernova Cosmology Project and 18 low-redshift SNe Ia from the Calán/Tololo Supernova Survey, after correcting for the SNe Ia lightcurve width-luminosity relation. Horizontal bars indicate the uncertainty in redshift due to an assumed peculiar velocity uncertainty of 300 km s^{-1}. Dashed and solid curves give the theoretical predictions for cosmological models with $\Omega_k = 0$ and $\Omega_\Lambda = 0$ respectively, for various values of Ω_m. Figure taken from [1].

1.1.4 Type Ia supernovae as standard candles

Supernovae of Type Ia (SNe Ia) are believed to occur when a white dwarf star in a binary system gains enough matter from its partner that its mass exceeds the Chandrasekhar limit, triggering a thermonuclear explosion. As the exploding star always has a mass near the Chandrasekhar limit, it is believed that there is little variation in the absolute luminosity of different SNe Ia, and therefore they can be treated as standard candles. As they are also extremely bright, this allows us to measure the luminosity distance-redshift relationship as far out as redshifts of $z \sim 2$.

By calibrating SNe Ia against Cepheid variable stars observed in the same galaxies, the Hubble Key Project [8] obtain a value for the Hubble constant of $H_0 = 71 \pm 2 \pm 6$ km s^{-1} Mpc^{-1}, and $H_0 = 72 \pm 8$ km s^{-1} Mpc^{-1} by combining different observational techniques. A more recent study [9] finds the value $H_0 = 74.8 \pm 3.1$ km s^{-1} Mpc^{-1}, a 3% determination. Other studies find values of h as low as 0.62—for a review,

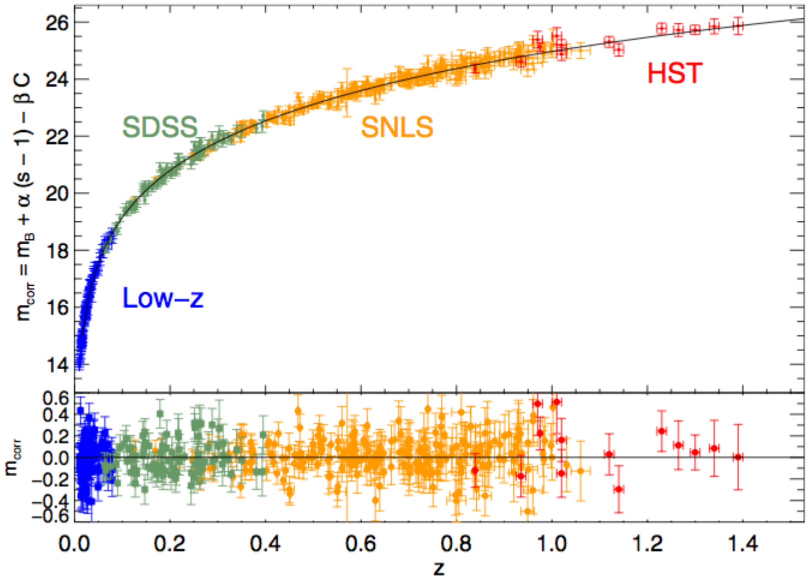

Figure 1.2: The Hubble diagram for a combined sample of Type Ia supernovae from the Supernova Legacy Survey (SNLS) data sample [5], the Sloan Digital Sky Survey (SDSS) [6], the Hubble Space Telescope (HST) [7] and from a combination of non-rolling surveys at low redshift. The combined sample contains a total of 472 supernovae. Figure taken from [2].

see [10].

Most excitingly, measurements of the luminosity distances of distant SNe Ia indicate that the deceleration parameter q_0 is *negative*, implying a late-time acceleration of the universe and a cosmological constant term that dominates the energy density today. The original studies were conducted by the Supernova Cosmology Project [1] and the High-z Supernova Search Team [11], who found that when the Hubble diagram for SNe Ia is compared with theoretical predictions for the luminosity distance as a function of redshift for different cosmological models, the data favour a model with $\Omega_\Lambda \sim 0.7$, $\Omega_m \sim 0.3$ and $q_0 \sim -0.5$. The equation of state for the dark energy component is consistent with $w = -1$, i.e., a cosmological constant. This basic picture remains consistent with recent, larger data sets [12, 13, 2].

A few caveats should be borne in mind. Firstly, at very low redshifts, there is significant scatter in the Hubble diagram due to peculiar motions of nearby supernovae. Secondly, it has become clear the SNe Ia are not simple standard candles [14], as their absolute luminosities are found to vary somewhat with environmental conditions, particularly the metallicity of their host galaxies, which in turn can depend on the epoch of the explosion. Hence, the distance to each supernova has to be determined individually, either through an empirical relationship between the maximum luminosity and the width of the light curve, or by normalising the peak luminosities and using a "corrected" apparent peak brightness [1]. We will discuss some of these issues later in Chapter 3.

Nevertheless, SNe Ia data clearly appear to favour an accelerating universe, and this, along with other evidence, has led to the adoption of the current standard cosmological model to which we now turn.

1.1.5 Dark energy and the standard cosmological model

From the Raychaudhuri equation, Eq. (1.14), we see that a necessary condition for the expansion of the universe to be accelerating, $\ddot{a} > 0$, as suggested by the SNe Ia data, is

$$p < -\frac{\rho}{3} \ . \tag{1.28}$$

For ordinary matter and radiation, as $w = 0$ and $1/3$ respectively, this condition is never satisfied and $q_0 \geq 0$. Thus we must postulate the existence of a new form of energy density, with negative pressure $w < -1/3$. As already mentioned, the energy density of the vacuum can provide exactly such a negative pressure. Lorentz invariance requires that the contribution of the vacuum to the energy-momentum tensor, $T_V^{\mu\nu}$, must be proportional to the metric $g^{\mu\nu}$. Comparing this with the form of the energy-momentum tensor for a perfect fluid, we see that the vacuum must have $p_V = -\rho_V$ and so $T_V^{\mu\nu} = -\rho_V g^{\mu\nu}$.

However, the principle of general covariance on which Einstein's theory is based, also allows a term proportional to the metric tensor on the left-hand side of Eq. (1.1):

$$R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} + \lambda g_{\mu\nu} = 8\pi G T_{\mu\nu} \ . \tag{1.29}$$

The contributions form the "geometrical" term and the energy density of the vacuum combine to give an effective cosmological constant

$$\Lambda = \lambda + \frac{\rho_V}{M_P^2} \ , \tag{1.30}$$

where we have substituted the Planck mass for Newton's constant, $M_P^2 = (8\pi G)^{-1} \simeq 2.4 \times 10^{18}$ GeV. This corresponds to an effective vacuum energy

$$\rho_\Lambda \equiv \Lambda M_P^2 \ , \tag{1.31}$$

and a vacuum energy density $\Omega_\Lambda = \Lambda/3H_0^2$.

As mentioned earlier, the standard cosmological model today is the flat ΛCDM model, with $\Omega_\Lambda \simeq 0.7$, $\Omega_m \simeq 0.7$ and $h \simeq 0.7$. Thus the cosmological constant dominates the energy density of the current universe and is responsible for the acceleration inferred from the SNe Ia data. Although we have motivated this model primarily from the SNe Ia observations, it has proven to be a remarkably good fit to other cosmological data, in particular the power spectrum of anisotropies in the CMB observed by WMAP [15], baryonic acoustic oscillations detected in the SDSS galaxy survey [16], and the lack of power on small scales in the angular power spectrum of the clustering of galaxies [17, 18] (however, for a critical appraisal of the data and the status of the standard model, see [19]).

Despite this impressive agreement with a wide range of observational data, however, we have no fundamental understanding of why the cosmological constant should take the value it does. In fact there are two troubling problems with the concept of a cosmological constant. The first is known as the *cosmological constant problem*, and was emphasized in the review by Weinberg [20], among other places. The energy density inferred from current observations constraining $\Omega_\Lambda \simeq 0.7$ is $\rho_\Lambda \approx 10^{-47}$ GeV^4. On the other hand, the vacuum energy density obtained by summing all the normal modes of some field with mass m up to a cutoff $M \gg m$ is

$$\rho_V = \int_0^M \frac{4\pi k^2 dk}{(2\pi)^3} \frac{1}{2}\sqrt{k^2+m^2} \simeq \frac{M^4}{16\pi^2} \, . \tag{1.32}$$

If we chose to believe GR right up to the Planck scale, we might choose the cutoff $M = M_P$, giving $\rho_V \simeq 10^{71}$ GeV^4, requiring the unrelated terms in Eq. (1.30) to cancel to nearly 120 decimal places. If we treat the standard model (SM) of particle physics as an effective field theory valid up to the energy scale of supersymmetry, the cutoff can at most be lowered from M_P down to the electroweak scale $M_{\rm EW} \sim 100$ GeV, but this still implies a minimum contribution to the vacuum energy of quantum fluctuations of $\mathcal{O}(M_{\rm EW}^4)$, which still leaves a discrepancy of at least 10^{60} between the theoretical expectation and the observed value.

There is at present no accepted solution to the cosmological constant problem. It may be hoped that in the future some new physical symmetry principle will be discovered to explain why the vacuum does not gravitate, which would mean Λ is exactly zero. However, the observation of a tiny but non-zero value for Λ creates a second problem, known as the *cosmic coincidence problem*: why is the value of Λ such that it is dynamically important *today*? In fact, the ratio Ω_Λ/Ω_m evolves as the cube of the scale factor $a(t)$, so the vacuum energy was negligible in the past and in the future matter energy density will be negligible. The effects of dark energy appear to become important at a value of the scale factor within a factor of about 2 of the current value, though it has increased by around 10^{28} since the end of inflation. So what is special about the present epoch that means that Ω_Λ/Ω_m is of order unity?

Partly as a result of these problems, there have been many attempts at building alternative theories to explain the apparent acceleration. One approach is to leave GR unchanged, but postulate new contributions to the energy-momentum tensor, arising from a dark energy field known as "quintessence" [21]. Another approach is to leave the field content unchanged but to modify GR [22]. Both of these approaches suffer from severe fine-tuning, however, and cannot be said to improve on the ΛCDM model. Some speculative ideas from string theory have been advanced, including a version of the anthropic argument [23]. It has also been argued that perhaps no explanation is required at all for a simple cosmological constant [24].

Cosmologists do invoke the effect of vacuum energy in order to drive inflation, as we shall discuss in Section 1.3. However, inflation occurs at a high energy scale, which appears natural in field theory. The problem with dark energy is that the implied fundamental energy scale, $\rho_\Lambda^{1/4} \sim \sqrt{H_0 M_P}$ is so unnaturally small. What is actually inferred from observations is however not an energy density, just a value of $\mathcal{O}(H_0^2)$ for the Λ term in the Friedmann equation. In the context of an FRW cosmological framework, $H_0 \sim 10^{-42}$ GeV is the *only* scale, and the inference of $\Lambda \sim \mathcal{O}(H_0^2)$ may be a result of mistakenly interpreting the inhomogeneous real universe in a homogeneous theoretical framework (e.g. [19]). We discuss an alternative inhomogeneous model in Chapter 3, and in Chapter 4 we examine an alternative observational test for ΛCDM through the dynamical effects of dark energy on the growth of structure.

In summary, despite the success of the ΛCDM model in matching observational data, it is fair to say that the nature of dark energy remains a major theoretical challenge, and so long as this is the case, alternative models should continue to be investigated.

1.2 Perturbations from homogeneity

In the previous section we have discussed the background solutions to the Einstein field equations assuming perfect homogeneity and isotropy. The real universe, however, is not perfectly homogeneous and in fact exhibits a rich structure at small enough scales. This structure has grown under the action of gravity, from very small matter perturbations in the early universe. We can study the evolution of these perturbations using linear perturbation theory, assuming that during most of the history of the universe deviations from the homogeneous and isotropic FRW background are small enough to be treated as first-order perturbations. A full treatment of cosmological perturbation theory is beyond the scope of this thesis, but in this section we will sketch out the basic elements of relevance to us. The details of the analysis can be found in [25], or in standard cosmology textbooks such as [26, 27].

1.2.1 Evolution of perturbations

The starting point for an analysis of perturbations are the Einstein field equations, Eq. (1.1). We account for deviations from the FRW background by writing the perturbed metric as

$$g_{\mu\nu} = \bar{g}_{\mu\nu} + h_{\mu\nu} \ , \tag{1.33}$$

where $\bar{g}_{\mu\nu}$ is the background metric and $h_{\mu\nu} = h_{\nu\mu}$ is a small perturbation (we use a bar over any quantity to denote its unperturbed value). The linearized Einstein equations can be written as

$$\delta R_{\mu\nu} - \frac{1}{2}\bar{g}_{\mu\nu}\delta R - \frac{1}{2}h_{\mu\nu}\bar{R} = 8\pi G \delta T_{\mu\nu} \ . \tag{1.34}$$

Here $\delta R_{\mu\nu}$, δR and $\delta T_{\mu\nu}$ denote the first order perturbations in $R_{\mu\nu}$, R and $T_{\mu\nu}$ respectively. In general, perturbations to the metric can be divided into scalar, vector and tensor types. The rotational and translational symmetry of the field equations means that each of these types of perturbation evolves independently and can be analysed separately. In general vector perturbations decay as $1/a^2$ with the expansion and are therefore usually ignored. Tensor perturbations correspond to gravitational waves and are important for some theories of inflation, but we will focus exclusively on scalar perturbations here.

Unlike in the unperturbed case, there is no uniquely preferred choice of coordinates for the perturbed metric, with the only requirement being that it reduce to the FRW background in the limit of infinitesimal perturbations. This also means that among the solutions of Eq. (1.34) will be unphysical modes that arise simply due to a change of coordinates of the unperturbed metric. In order to eliminate this problem we must make a gauge choice to fix the coordinate system, and evaluate all quantities in that gauge. There are several different possible gauge choices but we use the conformal Newtonian (or longitudinal) gauge, in which the line element is

$$ds^2 = a^2(\tau)\left[-(1+2\Psi)d\tau^2 + (1-2\Phi)dx^i dx^j\right] \ . \tag{1.35}$$

This gauge is restricted as it does not allow vector or tensor perturbations; however it is a convenient gauge in which to analyse scalar perturbations, which is what we are interested in here. Another advantage is that the quantity Ψ corresponds to the standard gravitational potential in the Newtonian limit, and Ψ and Φ are straightforwardly related to the gauge-independent variables Φ_A and Φ_H of Bardeen [28].

We utilise the symmetry of the field equations under translations in space by working in Fourier space. To linear order in the perturbations, there is no coupling between Fourier modes of different wavenumber $\mathbf{k}$, and the equations depend only on the magnitude and not the direction of $\mathbf{k}$. Some lengthy calculation (see [25] for details) starting from Eq. (1.34) leads to the following four equations:

$$\begin{aligned}
k^2\Phi + 3\frac{a'}{a}\left(\Phi' + \frac{a'}{a}\Psi\right) &= 4\pi G a^2 \delta T^0{}_0, && \text{(1.36a)}\\
\Phi'' + \frac{a'}{a}(\Psi' + 2\Phi') + \left(2\frac{a''}{a} - \frac{a'^2}{a^2}\right)\Psi + \frac{k^2}{3}(\Phi - \Psi) &= \frac{4\pi}{3} G a^2 \delta T^i{}_i, && \text{(1.36b)}\\
k^2\left(\Phi' + \frac{a'}{a}\Psi\right) &= 4\pi G a^2 (\bar\rho + \bar p)\theta, && \text{(1.36c)}\\
k^2(\Phi - \Psi) &= 12\pi G a^2 (\bar\rho + \bar p)\sigma, && \text{(1.36d)}
\end{aligned}$$

where the variables θ and σ are defined as

$$(\bar\rho + \bar p)\theta \equiv i k^j \delta T^0{}_j, \quad (\bar\rho + \bar p)\sigma \equiv -\left(\hat k_i \hat k_j - \frac{1}{3}\delta_{ij}\right)\Sigma^i{}_j, \tag{1.37}$$

with $\Sigma^i{}_j \equiv T^i{}_j - \delta^i{}_j T^k{}_k/3$. We have used the same notation to refer to the Fourier transforms of quantities as to the quantities themselves, and the $'$ denotes derivatives with respect to conformal time. Note that in the absence of any anisotropic stresses,

$\Phi = \Psi$ from Eq. (1.36). All the quantities $\delta T^0{}_0$, $\delta T^i{}_i$, θ and σ are to be evaluated in the conformal Newtonian gauge. The components of the energy-momentum tensor are subject to the conservation conditions $T^{\mu\nu}{}_{;\mu} = 0$ (these are not independent conditions as they can be derived from the field equations, but they may be usefully substituted in place of one or more of the field equations).

Evaluating the energy-momentum tensor is relatively simple for a non-relativistic fluid, for which we can write $\delta T^0{}_0 = -\delta\rho$, $\delta T^0{}_i = (\bar{\rho}+\bar{p})v_i$ and $\delta T^i{}_j = \delta p \delta^i{}_j$, where v^i is a small coordinate velocity $v^i = dx^i/d\tau$ and can be treated as a perturbation of the same order as $d\rho$, dp and the metric perturbations. Such a description is appropriate for CDM and for baryons, but for the photon and neutrino components one needs to evaluate the energy-momentum tensor through integrals over the momenta of the distribution functions in phase space. The general expression for the energy-momentum tensor is

$$T_{\mu\nu} = \int dp_1 dp_2 dp_3 (-g)^{1/2} \frac{p_\mu p_\nu}{p^0} f(x^i, p_j, \tau) \ , \tag{1.38}$$

where $f(x^i, p_j, \tau)$ is the phase space distribution and p_j is the conjugate momentum to x^i. The matter is further complicated by the fact that the different fluid components of the universe interact with each other non-gravitationally as well, for instance before recombination the photons and baryons are tightly coupled via Thomson scattering of photons off electrons. In order to describe this coupling we need the Liouville equation,

$$\frac{Df}{d\tau} = \frac{\partial f}{\partial \tau} + \frac{dx^i}{d\tau}\frac{\partial f}{\partial x^i} + \frac{dp_i}{d\tau}\frac{\partial f}{\partial p_i} = \left(\frac{\partial f}{\partial \tau}\right)_c \ . \tag{1.39}$$

The term $\left(\frac{\partial f}{\partial \tau}\right)_c$ describes the collisions affecting the particular species in question. The integral form of Eq. (1.39) is the Boltzmann equation.

All the information needed to track the evolution of the cosmological perturbations is contained in the Einstein equations, Eqs. (1.36) and the Boltzmann equations for the interaction of the various constituents of the universe, and all that is required is to specify the initial conditions. Clearly, however, these equations are quite complex, and an analytic solution will not be possible. Instead cosmologists use numerical packages to solve the equations in order to obtain the perturbations at the present (or any arbitrary) time. The package we use in this thesis is CAMB [29], which is based on an earlier package called CMBFAST [30], which in turn is based on the original code due to the authors of [25].[2] For most purposes we will be happy to treat this code as a black box.

1.2.2 Initial conditions

In order to be able to use the framework for tracking the evolution of perturbations outlined in the previous section, we need to specify the perturbations at some initial time. The perturbations are believed to have originated during an era of inflation, as will be discussed in Section 1.3; in fact, it is the ability the inflationary paradigm to explain the generation of these perturbations on scales that would not otherwise have been in causal contact that is the chief reason for its success. During the inflationary era, when the energy density is dominated by (at least in the simplest models) a single scalar field, we are able to make predictions about the perturbations that are generated. However, between the time of the end of inflation and the time of observation of the perturbations, the universe undergoes several phase transitions (such as the period of reheating after inflation, or the freezing out of certain particles as the universe cools) during which we do not know the equations governing the perturbations.

This being the case, we are only able to make any predictions about the observable perturbations based on inflationary theories because certain conservation laws apply to perturbation modes which have wavelengths long enough that they are outside the horizon, i.e., $k/a \ll H$. During inflation, H is almost constant and a increases almost exponentially with time, so all modes except those with the very shortest wavelengths exit the horizon, and the conservation laws apply. After inflation, a increases as $t^{1/2}$ or $t^{2/3}$ in radiation and matter-dominated eras respectively, whereas H decreases as $1/t$. Therefore the horizon grows and eventually all the modes of interest re-enter the horizon, though only at relatively recent times when we are once again able to track the evolution of perturbations without unknown effects. Thus the conservation laws for super-horizon perturbations allow us to connect theories of inflation with the observation of perturbations in the universe today.

---

[2] Although we have provided the equations in the conformal Newtonian gauge, most numerical solvers use the synchronous gauge as the equations are better behaved numerically.

For scalar modes, the quantity defined in the conformal Newtonian gauge as [31][3]

$$\zeta \equiv -\Phi - \frac{ik^i \delta T^0{}_i H}{k^2(\bar{\rho}+\bar{p})} \tag{1.40}$$

is conserved outside the horizon if the initial perturbations generated at the time of inflation are adiabatic—i.e., the different constituents are initially in thermal equilibrium, such that the scalar modes have equal values of $\delta\rho_\alpha/\dot{\bar{\rho}}_\alpha$ for each individual constituent α of the universe, whether or not energy is separately conserved for each constituent (this condition is always satisfied in models of inflation driven by a single scalar field as the pressure and density are unique functions of the field value). When this condition is met, we can make predictions for the perturbations in ζ based on the model of inflation, and then translate these back to perturbations of Φ, $\delta\rho$ etc. at the time that the relevant mode re-enters the horizon. This is used as the initial condition for the numerical Boltzmann solvers discussed above.

The quantity ζ is actually gauge-invariant. In the conformal Newtonian gauge it takes the value in Eq. (1.40), but during inflation it is more convenient to use a gauge with spatially flat slicing. We will discuss this further in Section 1.3.4.

Cosmological theories do not give definite predictions for perturbed quantities such as $\delta\rho(t)$. Instead the fluctuations are treated as stochastic random variables, and theoretical predictions are made about the probability distribution functions (PDFs) of the perturbations. The PDF in the cases of interest can be completely specified in terms of the spatial n-point correlation functions of the random field, and the assumption of homogeneity and isotropy for the unperturbed universe is replaced now by the assumption of statistical homogeneity and isotropy of the perturbations.

The simplest models of inflation predict Gaussian PDFs and observations of the CMB anisotropies indicate that any non-Gaussianity, if present, is small, so for the purposes of this thesis we shall consider only the Gaussian case. This simplifies the analysis, as for a scalar Gaussian random field with zero mean, we only require knowledge of the two-point correlation function, its derivatives and integrals in order to calculate any statistical property. The Fourier transform of the two-point function is the power spectrum, defined as

$$\langle \Phi(\mathbf{k},t)\Phi^*(\mathbf{k}',t)\rangle = (2\pi)^3 P_\Phi(k)\delta(\mathbf{k}-\mathbf{k}') \ . \tag{1.41}$$

---

[3] Other authors use the related quantity $\mathcal{R}(k)$. In the long-wavelength limit the definitions coincide, $\zeta = \mathcal{R}$. For a fuller discussion and historical context, see [32] and the references within.

In linear perturbation theory, an initially Gaussian perturbation will remain Gaussian at all times, and so the power spectrum remains enough to describe the statistical properties of the perturbation. Non-linear effects will in fact introduce some late-time non-Gaussianity even to a perfectly Gaussian initial fluctuation, but we do not consider these effects in this work.

An important property of Gaussian random processes is their ergodicity, which ensures that spatial averages taken over a large volume in our universe will be equal to expectations taken over an ensemble of universes, to which our universe belongs. This is fortunate, as of course we are unable to measure quantities over an ensemble of universes![4]

1.2.3 The matter power spectrum

It is of great interest to obtain theoretical predictions for the perturbations in matter density $\delta = \delta\rho/\bar{\rho}$ that can be compared with observation. Given the initial power spectrum generated by inflation (Section 1.3.4), it is possible to solve for the evolution of the Fourier modes $\delta(k,t)$ in order to construct the matter power spectrum today. To do so obviously requires the full perturbation treatment outlined above, but we present here a simplified discussion of the processes involved.

It is natural to divide up the evolution of cosmological perturbations into three distinct stages: early time, when the modes are outside the horizon, $k/a \ll H$; intermediate time, when the modes cross into the horizon and also the universe changes from radiation-domination to matter-domination; and late time, when modes are well inside the horizon, $k/a \gg H$, and the universe is matter-dominated (or dark energy-dominated). The evolution of modes that are well outside or well inside the horizon is independent of k, but there is a strong k-dependence during the intermediate stage. We therefore describe the evolution of modes in terms of a k-independent growth factor describing the growth of perturbations at late times, and a transfer function $T(k)$ describing the evolution through the epoch of horizon crossing and the radiation/matter transition.

---

[4] It is actually not a necessary condition that the fluctuations be strictly Gaussian, only that they be uncorrelated at distant arguments.

The transfer function is conventionally defined as

$$T(k) \equiv \frac{\Phi(k, a_{\mathrm{late}})}{\Phi_{\mathrm{Large-Scale}}(k, a_{\mathrm{late}})} \tag{1.42}$$

where a_{late} is the scale factor at some late time well after the regime in which the transfer function is applicable. $\Phi_{\mathrm{Large-Scale}}$ is the primordial perturbation evolved to the late time using only the large-scale (super-horizon) solution. This is somewhat less than the primordial perturbation Φ_i from inflation as even super-horizon modes evolve somewhat as the universe passes through the epoch of equality. In fact, it can be shown from the super-horizon solution that $\Phi_{\mathrm{Large-Scale}}(a_{\mathrm{late}}) = \frac{9}{10}\Phi_i$ [33]. The growth factor $D(a)$ is defined as

$$\frac{D(a)}{a} = \frac{\Phi(a)}{\Phi(a_{\mathrm{late}})} \qquad (a > a_{\mathrm{late}}) \; . \tag{1.43}$$

Solving the evolution equations for sub-horizon modes gives an expression for the growth factor:

$$D(a) = \frac{5\Omega_m}{2}\frac{H(a)}{H_0}\int_0^a \frac{da'}{(a'H(a')/H_0)^3} \; . \tag{1.44}$$

For a matter-dominated universe $\Omega_m = 1$, the growth factor is simply $D(a) = a$, i.e., the potential does not evolve (but the CDM density perturbations grow $\delta \propto D(a)$). In the presence of a cosmological constant the growth of density perturbations is slowed due to the accelerated expansion of the universe. For a ΛCDM cosmology, an approximate form, good to a few percent accuracy, is [34]:

$$D(a) \approx \frac{5}{2}\Omega_m \left[\Omega_m^{4/7} - \Omega_\Lambda + \left(1 + \frac{1}{2}\Omega_m\right)\left(1 + \frac{1}{70}\Omega_\Lambda\right)\right]^{-1} \; . \tag{1.45}$$

Using these definitions, the potential perturbation at late times can be written as

$$\Phi(k, a) = \frac{9}{10}\Phi_i(k)T(k)\frac{D(a)}{a} \qquad (a > a_{\mathrm{late}}) \; . \tag{1.46}$$

The matter density contrast today can be related to the potential at late time using Poisson's equation (which is obtained from a combination of the perturbation equations (1.36) in the limit that radiation density is negligible and $k/a \gg H$)

$$\Phi = \frac{4\pi G \rho_m a^2 \delta}{k^2} \; . \tag{1.47}$$

Inverting this and using Eq. (1.46) we can relate the density contrast to the initial perturbation Φ_i

$$\delta(k, a) = \frac{3}{5}\frac{k^2}{\Omega_m H_0^2}\Phi_i(k)T(k)D(a) \; , \tag{1.48}$$

which allows us to relate the power spectrum for δ to the primordial power spectrum for Φ

$$P_\delta(k,a) = \frac{9}{25}\frac{k^4}{\Omega_m^2 H_0^4} T^2(k) D^2(a) P_\Phi(k) \ . \tag{1.49}$$

This relationship will be generally valid, irrespective of whether the primordial perturbation Φ_i is generated from inflation or not. Note that our linear treatment will not be valid on small enough scales: for most models the scale at which non-linear corrections to the power spectrum need to be accounted for is $k_{\rm nl} \simeq 0.2\ h\ \mathrm{Mpc}^{-1}$.

The simplest single-field models of inflation motivate a form of the primordial power spectrum $P_\Phi(k) \propto k^{n_s-4}$, where n_s is the scalar spectral index, $n_s \approx 1$. The value $n_s = 1$ corresponds to a scale-invariant, Harrison-Zel'dovich spectrum [35]. Therefore on large scales, where $T(k) \approx 1$, the matter power spectrum increases as $P_\delta(k) \propto k^{n_s}$. On small scales the power spectrum turns over because smaller-scale modes that enter the horizon before matter-radiation equality decay during the radiation-dominated epoch. The point of turnover is clearly dependent on the horizon scale at matter-radiation equality, which in turn depends on the value of $\Omega_m h^2$.

The relative abundance of baryons also affects the matter power spectrum in two ways. Firstly, baryons reduce the growth rate between the matter-radiation epoch and the drag epoch when the baryons are decoupled from the photons, leading to a suppression of power on small scales relative to large scales. Secondly, the tight coupling between photons and baryons before recombination leads to sound waves in the primordial plasma seeded by the initial fluctuations in the dark matter density. When the photons and baryons decouple, there is a drop in pressure and the sound wave in the baryons is stalled leading to an excess of clustering at a radius of approximately the sound speed times the age of the universe at recombination, which is $\simeq 154$ Mpc today [36]. Although the baryons slowly move back into equilibrium with the dark matter, evidence for the stalled sound waves is expected to be visible today in an excess of clustering of galaxies at the acoustic scale. This is known as the baryon acoustic oscillation (BAO) peak.

In comparing the theoretical predictions for the CDM power spectrum to observation one has to take account of the fact that the galaxy distribution (which is what is actually observed) traces the baryonic or luminous matter, and that this may not correspond precisely to the distribution of dark matter. Naïvely, one expects that galaxy formation is enhanced in regions of greatest matter overdensity. This is particularly so of luminous red galaxies (LRGs), which are elliptical galaxies in massive galaxy clusters occurring in large dark matter halos [37]. Because LRGs are thought to be physically similar objects throughout the redshift range over which they are observed [38], they are often used in cosmological tests as a tracer of the matter density, as we do in Chapter 4. The simplest assumption is to take a constant bias $\delta_g = b\delta$, which leads to $P_g(k) = b^2 P_\delta(k)$.

The best measurements of the galaxy power spectrum currently come from the Sloan Digital Sky Survey (SDSS) collaboration, e.g. [39]. On the largest scales observed, the turnover point of $P_g(k)$ is still not observed, which appears to strongly favour low values of Ω_m. The best-fit cosmological parameters, assuming a flat ΛCDM model with a fixed spectral index n_s consistent with the CMB measurements, are $\Omega_m \sim 0.29$, $\Omega_\Lambda \sim 0.71$, $h \sim 0.69$ consistent with the "concordance" values obtained from other observations. A detection of the BAO peak in SDSS Data Release 7 is also reported [40], with constraints on Ω_m and h again consistent with the standard model, but it has been claimed that the statistical significance of this observation has been over-estimated due to incorrect treatment of errors [41]. Using the luminous red galaxy (LRG) sample of the SDSS DR3, a BAO feature was reported in the galaxy correlation function at high significance [16]; however the significance of this detection is lower in the DR7 LRG data, though still consistent with ΛCDM [42].

1.2.4 CMB anisotropies

After subtracting a dipole in order to correct for our relative motion with respect to the CMB rest-frame, we observe the CMB to have a distribution that is almost exactly a Planck spectrum [43] with a mean temperature $T_0 = 2.726 \pm 0.001$ K [44]. This temperature is extremely uniform, with fractional deviations in temperature $\Delta T/T$ with direction of observation being of $\mathcal{O}(10^{-5})$. Nevertheless, the temperature anisotropies $\Delta T(\hat{n}) \equiv T(\hat{n}) - T_0$ are of great importance to cosmology as they can provide information about both the nature of the primordial perturbations and also the content of the universe at late times.

It is convenient to expand the temperature anisotropies observed in direction $\hat{n}$ in spherical harmonics $Y_l^m(\hat{n})$:

$$\Delta T(\hat{n}) = \sum_{lm} a_{lm} Y_l^m(\hat{n}) \ , \tag{1.50}$$

with the sum over l running over all positive-definite integers, and the sum over m running from $-l$ to l. The reality of $\Delta T(\hat{n})$ imposes the condition $a^*_{lm} = a_{l\,-m}$ on the coefficients. The monopole term a_{00} is unobservable. The motion of the earth relative to the CMB rest frame contributes an anisotropy that is $\propto \mathcal{P}_1(\cos\theta) \propto Y_1^0(\theta, \phi)$ (where $\mathcal{P}_l$ denotes the Legendre polynomial), i.e., the dipole term. Multipoles with $l \geq 2$ then represent the intrinsic anisotropy of the CMB.

As mentioned above, for cosmological purposes we are interested in average quantities. Assuming that the universe is rotationally invariant, $\Delta T(\hat{n})$ is independent of $\hat{n}$ and $\langle \Delta T(\hat{n}) \rangle$ vanishes. The simplest non-vanishing quantity is therefore the power spectrum, which takes the form

$$\langle a_{lm} a^*_{l'm'} \rangle = \delta_{ll'} \delta_{mm'} C_l \tag{1.51}$$

as a result of the rotational invariance. If the temperature perturbation is Gaussian, this is the only relevant statistical quantity. Eq. (1.51) shows that the multipole coefficients C_l are real and positive. Since

$$\langle \Delta T(\hat{n}) \Delta T(\hat{n}') \rangle = \sum_{lm} C_l Y_l^m(\hat{n}) Y_l^{-m}(\hat{n}') = \sum_l C_l \left(\frac{2l+1}{4\pi} \right) \mathcal{P}_l(\hat{n} \cdot \hat{n}') \ , \tag{1.52}$$

the multipole coefficients can be written as

$$C_l = \frac{1}{4\pi} \int d^2\hat{n} \, d^2\hat{n}' \ \mathcal{P}_l(\hat{n} \cdot \hat{n}') \, \langle \Delta T(\hat{n}) \Delta T(\hat{n}') \rangle \ . \tag{1.53}$$

However, as we can only observe one universe, what is actually observed is not the true C_l but a quantity averaged over m but not over position of the observer,

$$C_l^{\mathrm{obs}} \equiv \frac{1}{2l+1} \sum_m a_{lm} a^*_{lm} = \frac{1}{4\pi} \int d^2\hat{n} \, d^2\hat{n}' \ \mathcal{P}_l(\hat{n} \cdot \hat{n}') \, \Delta T(\hat{n}) \Delta T(\hat{n}') \ . \tag{1.54}$$

For Gaussian perturbations, it can be shown that the mean square fractional difference decreases with l:

$$\left\langle \left(\frac{C_l - C_l^{\mathrm{obs}}}{C_l} \right)^2 \right\rangle = \frac{2}{2l+1} \ . \tag{1.55}$$

This quantity is known as the *cosmic variance*, and limits the useful information that can be obtained from the lowest multipoles. For the WMAP 7-year data, cosmic variance provides the largest source of error for multipoles with $l \leq 548$ [3].

We can expand the temperature fluctuation at a given space-time point (x^i, τ) in terms of its Fourier transform

$$\Delta T(\mathbf{x}, \hat{n}, \tau) = \int \frac{d^3k}{(2\pi)^3} e^{i\mathbf{k}\cdot\mathbf{x}} \Theta(\mathbf{k}, \hat{n}, \tau) \ . \tag{1.56}$$

$\Theta(\mathbf{k}, \hat{n}, \tau)$ does not depend on the direction $\hat{k}$. It can be further expanded in terms of its multipole moments Θ_l, so that the temperature anisotropy may be written as

$$\Delta T(\mathbf{x}, \hat{n}, \tau) = \int \frac{d^3k}{(2\pi)^3} e^{i\mathbf{k}\cdot\mathbf{x}} \sum_{l=0}^{\infty} (-i)^l (2l+1) \mathcal{P}_l(\hat{k}\cdot\hat{n}) \Theta_l(k, \tau) \ . \tag{1.57}$$

At our location ($\mathbf{x} = 0, \tau = \tau_0$) we may compare Eq. (1.57) with Eq. (1.50) and obtain the coefficients a_{lm}

$$a_{lm} = (-i)^l 4\pi \int \frac{d^3k}{(2\pi)^3} \, Y_l^{-m}(\hat{k}) \, \Theta_l(k, \tau_0) \ . \tag{1.58}$$

$\Theta_l(k, \tau_0)$ is determined from the evolution of the initial temperature perturbation. For adiabatic initial perturbations, the initial conditions can be specified entirely in terms of the potential Φ in the conformal Newtonian gauge, and we can write

$$\Theta_l(k, \tau_0) = \Delta_l(k, \tau_0)\Phi(k) \ , \tag{1.59}$$

where $\Delta_l(k, \tau_0)$ is the radiation transfer function. Using this and Eq. (1.58) it follows that the temperature angular power spectrum is

$$C_l = \frac{2}{\pi} \int_0^{\infty} k^2 P_\Phi(k) \Delta_l^2(k, \tau_0) \, dk \ , \tag{1.60}$$

where $P_\Phi(k)$ is the power spectrum of the initial perturbations in Φ, Eq. (1.41). Thus, for a given l, C_l is an integral over all Fourier modes of the variance of $\Theta_l(k)$, written now in terms of the transfer function and the initial power spectrum of Φ.

In order to calculate $\Delta_l^2(k, \tau_0)$ one needs to solve the full set of Boltzmann equations starting with an initial perturbation $\Phi_i = 1$. This can be done accurately by the numerical Boltzmann solvers such as CAMB, and we rely on them for detailed results. Approximate analytic calculations are performed in, e.g. [45], but this is also beyond the scope of this thesis. Instead we present here a simple heuristic description of the physics affecting the anisotropy spectrum.

To a first approximation, we can assume that the CMB photons have travelled freely since last scattering at $z \sim 1100$, only redshifting due to the expansion of the universe (there is in fact a small effect due to reionization). Therefore the anisotropies we observe today arise from the following sources:

1. Intrinsic temperature fluctuations in the photon-baryon plasma at the time of last scattering at a redshift of $z \sim 1100$.

2. The Doppler effect due to bulk velocities of the photon-baryon plasma at last scattering.

3. The gravitational redshift/blueshift of photons due to fluctuations in the gravitational potential Φ at last scattering (known as the Sachs-Wolfe effect [46]).

4. The gravitational redshift/blueshift of photons due to *time-dependent* fluctuations in Φ along the line of sight between the last scattering surface and us (known as the integrated Sachs-Wolfe effect [46]).

The Sachs-Wolfe effect itself is the net result of two contributions: the gravitational redshift of photons due to Φ, and the change in the rate of expansion of the universe due to Φ, which shifts the value of the redshift at which the universe reaches the decoupling temperature $\simeq 3000$ K. The total effect is

$$\left(\frac{\Delta T(\hat{n})}{T}\right)_{\mathrm{SW}} = \frac{1}{3}\Phi \ . \tag{1.61}$$

By Fourier transforming Φ and using a Legendre expansion, this can be written as

$$\left(\frac{\Delta T(\hat{n})}{T}\right)_{\mathrm{SW}} = \frac{1}{3}\sum_{l=0}^{\infty}(2l+1)i^l \int \frac{d^3k}{(2\pi)^3}\Phi(k)\, j_l(kr_L)\mathcal{P}_l(\hat{k}\cdot\hat{n}) \ , \tag{1.62}$$

where j_l is the spherical Bessel function and r_L is the coordinate distance to the surface of last scattering. From this it is possible to obtain the result that

$$C_{l,\mathrm{SW}} \propto \frac{1}{l(l+1)} \tag{1.63}$$

if the power spectrum of the curvature fluctuation is scale-invariant, $P_\Phi \propto k^{-3}$. Pressure gradients are important in the coupled photon-baryon plasma, so the differential

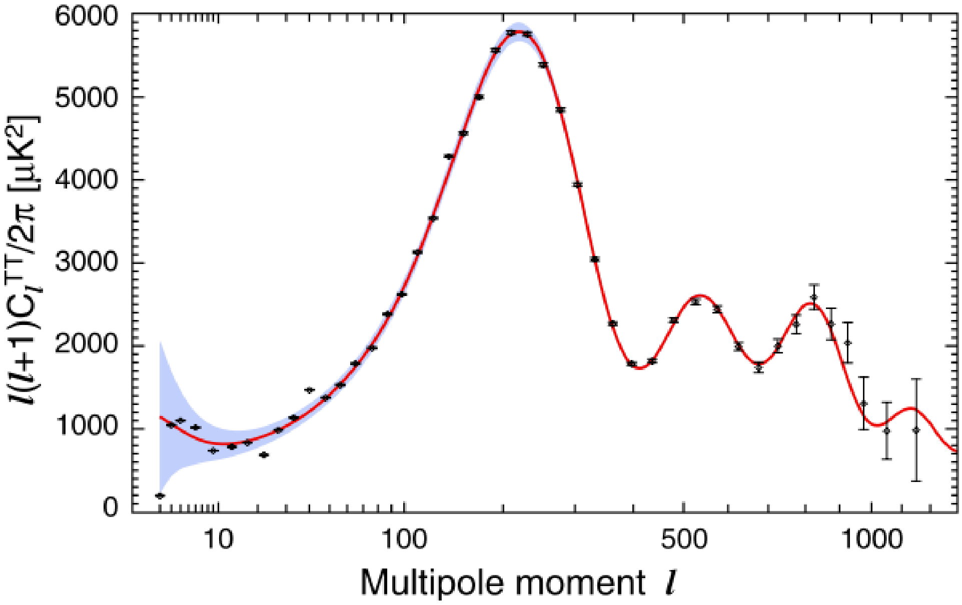

Figure 1.3: Binned WMAP 7-year data for the CMB temperature power spectrum. The solid line shows the prediction for the best-fitting flat ΛCDM model with 6 parameters. The error bars depicted include errors due to instrument noise, and the grey band shows the cosmic variance. Figure taken from [3].

equations describing the dynamics of the plasma will depend on the wavenumber k. These equations need to be solved in full in order to obtain the contributions from the Doppler effect and the intrinsic temperature fluctuations. However, without solving them we can note that the perturbation in Fourier space can always be expanded in terms of Legendre polynomials with coefficients proportional to $j_l(kr_L)$, and hence for $l \gg 1$, the integral over k will be dominated by values of $k \approx l/r_L$. Note that the physical wavenumber at time t_L is $\mathbf{k}_L = \mathbf{k}/a(t_L)$, so for large l the contribution to C_l reflects the behaviour of the perturbations at $k_L \approx l/d_A$, where $d_A = r_L a(t_L)$ is the angular diameter distance to the last scattering surface.

For perturbations that are outside the horizon at the time of last scattering, no evolution has occurred, so the dependence on k comes entirely from the initial conditions. The horizon distance in question is not the same as that defined in Section 1.1.1 but rather the *sound horizon distance* $d_{\rm hor}$, as the perturbations in the photon-baryon plasma are in fact sound waves. Therefore gradients become important when $k_L \approx 1/d_{\rm hor}$, and since the contribution to C_l comes from $k_L \approx l/d_A$, gradients will be important when $l \gtrsim l_{\rm hor} \equiv d_A/d_{\rm hor}$. For the cosmological parameters of the ΛCDM model, $l_{\rm hor} \approx 100$. For multipoles of order $l \ll l_{\rm hor}$, the contribution of the Doppler effect is less than that of the Sachs-Wolfe effect by a factor $\sim l^2/l_{\rm hor}^2$. The interference of the contribution from intrinsic temperature fluctuations with the Sachs-Wolfe effect is also of the same order. Therefore at multipoles $l \lesssim 100$, the Sachs-Wolfe effect dominates and the shape of the power spectrum is given by Eq. (1.63). This is known as the Sachs-Wolfe plateau.

At higher multipoles the Doppler effect and the intrinsic temperature anisotropy imprint a characteristic oscillation on the CMB, which appears as a series of acoustic peaks in the angular power spectrum. The peaks decrease in amplitude with increasing l due to photon diffusion which increases isotropy at small scales (Silk damping). A number of other factors dependent on the physics at last scattering affect the details of the shape of the peaks. Baryon loading enhances the amplitude of odd-numbered peaks, and the baryon density also affects the sound speed, and hence the sound horizon and peak spacing. The amplitudes of the peaks are also sensitive to the cold dark matter-radiation ratio at recombination, because the radiation energy density drives the oscillations by contributing to the time evolution of the potentials. The position of the peak corresponds to the wavelength of the last acoustic wave created just before decoupling, which basically corresponds to the sound horizon at decoupling. However, the angle subtended on the sky by the sound horizon is dependent on the angular diameter distance to the last scattering surface and can therefore constrain Ω_k or Ω_Λ.

Figure 1.3 shows the most recent 7-year data from the WMAP satellite measurement of the power spectrum, along with the theoretical prediction for the best-fit flat ΛCDM model, which has six parameters: Ω_b, Ω_m, h, the amplitude A_s and spectral index n_s of the primordial scalar power spectrum, and the optical depth to reionization, τ.

1.3 Inflation

1.3.1 Shortcomings of the Big Bang model

There are three main problems that the Big Bang scenario discussed so far does not explain, known as the horizon, flatness, and monopole problems, which are commonly cited as the motivation for a period of inflation in the early universe.

Horizon problem

The high degree of isotropy observed in the CMB poses a significant problem for the Big Bang scenario. This is because the particle horizon size in both the matter- and radiation-dominated eras decreases as we go back in time. In fact, at the time of last scattering $d_H \approx H_0^{-1}(1+z_L)^{-3/2}$, where $z_L \sim 1100$ is the redshift of last scattering: this corresponds to an angle on the sky today of only $\approx 1.6°$. Thus regions on the sky that are separated by more than a few degrees should not have been in causal contact at the time of last scattering, and no physical influence could have brought them into thermal equilibrium with each other, in apparent contradiction with the near-perfect isotropy that is actually observed.

Flatness problem

As we have discussed, current observations from SNe Ia surveys are consistent with FRW models with zero spatial curvature. The WMAP data, when combined with a prior on the value of H_0, favours models with $\Omega_k = 0$. There is some degeneracy in these results between the values of Ω_k and Ω_Λ, but at the very least we may conclude that $|\Omega_k| < 1$. However, $|\Omega_k|$ is just the current value of the dimensionless curvature parameter $|k|/a^2H^2$, and since the scale factor grows as $a \propto t^\alpha$ with $\alpha < 1$ for both radiation- and matter-domination, $|k|/a^2H^2$ grows as $t^{2-2\alpha}$. Therefore, for $|\Omega_k| < 1$ today, at the time of nucleosynthesis (a temperature of $\sim 10^{10}$ K), the value of k/a^2H^2 must have been at most $\approx 10^{-16}$.

Although this is not impossible, the question of why the curvature should have been so small at early times is a puzzle that we would like to be able to answer.

Monopole problem

In grand unified theories of particle physics, a symmetry-breaking phase transition is expected to occur at an energy of $\approx 10^{16}$ GeV. The process of symmetry breaking is expected to result in the creation of relics that carry a non-zero magnetic charge, or monopoles. As the field responsible for breaking the symmetry would have to be uncorrelated on distances larger than the particle horizon at the time, these monopoles are expected to have been produced with a number density comparable with the nucleon number density. This is in gross disagreement with the current limits from searches for monopoles of fewer than about 10^{-30} monopoles per nucleon [47].

Inflation—defined very generically as a period of almost exponential expansion—solves or alleviates all three of these problems. To see how, let us assume that the scale factor of the universe increased by some factor $e^{\mathcal{N}}$ during the inflationary period, where $\mathcal{N}$ is known as the number of e-foldings of inflation. From Eq. (1.11) we see that the particle horizon at the time of last scattering is

$$d_H(t_L) \equiv a(t_L) \int_{t_*}^{t_L} \frac{dt}{a(t)} \, , \tag{1.64}$$

where t_* is the beginning of the inflationary era (t_* is possibly equal to $-\infty$). This integral will be dominated by the contribution from the inflationary period, as we have already seen that the radiation- and matter-dominated eras give a horizon that is far too small. Given that the expansion was exponential, we may write $a(t) = a_I \exp\left(-H_I(t_I - t)\right)$ for $t < t_I$, where the subscript I denotes quantities at the end of inflation. Then $\mathcal{N} = H_I(t_I - t_*)$, and

$$d_H(t_L) = \frac{a(t_L)}{a_I H_I}\left[e^{\mathcal{N}} - 1\right] \simeq \frac{a(t_L)}{a_I H_I} e^{\mathcal{N}} \, . \tag{1.65}$$

In order for inflation to be able to explain the horizon problem, we require that $d_H(t_L) > d_A(t_L)$, the angular diameter distance of the surface of last scattering. Given the definition of the angular diameter distance (1.22) and the fact that for a nearly flat universe $r_L \approx 1/a_0 H_0$, this condition amounts to

$$e^{\mathcal{N}} \gtrsim \frac{a_I H_I}{a_0 H_0} \, , \tag{1.66}$$

which gives the minimum number of e-folds of inflation required to solve the horizon problem.

The flatness problem is similarly alleviated, because $|k|/a^2H^2$ decreases by a factor of $e^{-2\mathcal{N}}$ during inflation and increases by a factor of $\left(\frac{a_I H_I}{a_0 H_0}\right)^2$ between the end of inflation and today. Therefore a value of $|k|/a^2H^2$ of order unity at the start of inflation is compatible with the current limits on the curvature if the horizon condition (1.66) holds. Estimating the number of e-folds required to solve the monopole problem is a model-dependent problem. It will be enough to require that the universe is homogeneous over horizon scales at the unification scale, and since this horizon is smaller than the current horizon, a number of e-folds satisfying the condition (1.66) will also satisfy the monopole condition.

The most compelling of these problems is the horizon problem. It is possible to imagine scenarios in which the flatness and monopole problems are resolved without requiring inflation—for instance if the curvature is for some reason exactly zero and always has been, or if the physics beyond the Standard Model does not in fact produce exotic relics—and any number of e-folds of inflation that solve the horizon problem will also solve the other two. The main reason to believe in the inflationary paradigm is that it is at present the best known method for generating the correlated superhorizon density perturbations, the presence of which can be inferred from the observed CMB angular power spectrum.

1.3.2 Slow-roll inflation

As we saw in Section 1.1.2, in order to obtain a period of accelerated expansion of the universe, we require the universe to be dominated by some form of energy with a sufficiently negative equation of state, such as vacuum energy. Guth [48] first attempted to implement this using the scalar field responsible for breaking the SU(5) "grand unified" symmetry trapped in a false minimum of its potential, thus contributing a constant vacuum energy density while the universe expands exponentially with a constant Hubble rate.[5] However, this model of "old inflation" suffered from the problem that it was not possible to construct a model that successfully exited this inflationary phase to the Big Bang phase while simultaneously solving the horizon and flatness problems. Most inflationary models today achieve nearly exponential expansion through having one or more scalar fields slowly roll down a potential hill such that their vacuum energy dominates over the kinetic energy. At the end of inflation, this field (or fields) releases its energy into other degrees of freedom, thus reheating the universe.

The energy density ρ and pressure p of a spatially homogeneous scalar field $\phi(t)$ (known as the 'inflaton' field) with potential $V(\phi)$ in an FRW background are

$$\rho = \frac{1}{2}\dot{\phi}^2 + V(\phi) \, , \qquad p = \frac{1}{2}\dot{\phi}^2 - V(\phi) \, . \tag{1.67}$$

---

[5] In fact other authors [49] had already considered the possibility of exponential expansion, but it was Guth's work that brought inflation to the attention of the wider community.

The energy conservation condition, Eq. (1.15), for the scalar field becomes

$$\ddot{\phi} + 3H\dot{\phi} + V'(\phi) = 0 \ , \tag{1.68}$$

where $'$ indicates differentiation with respect to ϕ. While the scalar field dominates the energy density of the universe, the Hubble parameter is

$$H = \sqrt{\frac{1}{3M_P^2}\left(\frac{1}{2}\dot{\phi}^2 + V(\phi)\right)} \ , \tag{1.69}$$

where as in Section 1.1.5 we use the Planck mass M_P instead of Newton's constant. By combining Eqs. (1.68) and (1.69), we obtain the time evolution of the Hubble parameter

$$\dot{H} = -\frac{\dot{\phi}^2}{2M_P^2} \ . \tag{1.70}$$

For inflation we require a period of almost exponential growth of the scale factor, in which the Hubble parameter will be almost constant. Quantitatively, we require that the fractional change in H in one expansion time $(1/H)$ should be much less than one, which means $|\dot{H}| \ll H^2$. From Eqs. (1.69) and (1.70), this condition translates to

$$\dot{\phi}^2 \ll |V(\phi)| \ , \tag{1.71}$$

and hence

$$H \simeq \sqrt{\frac{V(\phi)}{3M_P^2}} \ . \tag{1.72}$$

Another condition that is often imposed is that the fractional change in $\dot{\phi}$ in one expansion time should also be small, which means $|\ddot{\phi}| \ll H|\dot{\phi}|$. If this condition is satisfied, we can drop the second derivative term from Eq. (1.68), thus obtaining

$$\dot{\phi} = -\frac{V'}{3H} \ . \tag{1.73}$$

Under this approximation, we can see that the condition for inflation to occur is that $(V'/V)^2$ be small compared to M_P^2. This motivates the definition of the first slow-roll parameter

$$\epsilon \equiv \frac{M_P^2}{2}\left(\frac{V'}{V}\right)^2 \ , \tag{1.74}$$

which must be small compared to unity for as many e-folds of expansion as we require inflation to generate. The condition $|\ddot{\phi}| \ll H|\dot{\phi}|$ implies that $|V''/V| \ll M_P^2$, so we define the second slow-roll parameter

$$\eta \equiv M_P^2 \left(\frac{V''}{V} \right) \ . \tag{1.75}$$

For the slow-roll approximation to be valid, we require $\epsilon \ll 1$ and $\eta \ll 1$. These are not strictly necessary conditions—it is possible for the slow-roll approximation to be violated without bringing inflation to an end, but such violations must be brief. For inflation to last for many e-folds, the slow-roll conditions must be satisfied for most of that epoch.

Alternative slow-roll parameters may be defined in terms of the actual dynamic variables during inflation, rather than the inflaton potential [50]. Following from the condition $|\ddot{\phi}| \ll H|\dot{\phi}|$, we define

$$\epsilon_H \equiv - \left(\frac{\dot{H}}{H^2} \right) \ , \tag{1.76}$$

and the second slow-roll parameter is defined in order to quantify the rate of change of ϵ_H:

$$\eta_H \equiv \frac{\dot{\epsilon}_H}{\epsilon_H H} . \tag{1.77}$$

In terms of these slow-roll parameters, the condition for inflation to occur is exactly $\epsilon_H < 1$. If the slow-roll approximation is valid, the two sets of slow-roll parameters are related to each other by

$$\epsilon = \epsilon_H \ , \quad \eta = \frac{1}{2}\eta_H + 2\epsilon_H \ . \tag{1.78}$$

The number of e-folds of inflation that occur between time t and some later time $t_{\rm end}$ at which slow-roll inflation ends is given by $\mathcal{N}(t) = \int_t^{t_{\rm end}} H(t)dt$, which can be written as

$$\mathcal{N}(\phi) = \int_{\phi}^{\phi_{\rm end}} \frac{H}{\dot{\phi}} d\phi \ , \tag{1.79}$$

where $\phi_{\rm end}$ denotes the field value at which slow-roll inflation ends. In many models, though not all, inflation itself will end very soon after slow-roll is violated, so this may be taken to be the actual number of e-foldings of inflation.

1.3.3 The number of e-folds of slow-roll inflation

The observations of the perturbations that we make today probe a region of the inflationary potential corresponding to the location of the inflaton at the time that these modes left the horizon during inflation. It is therefore useful to quantify the number of e-foldings before the end of inflation at which the modes that we currently observe in the CMB left the horizon. In doing so we follow closely the argument of [51]. We assume that after the end of inflation there is a period of reheating, during which the universe expands as matter-dominated, followed by radiation-domination until a redshift of a few thousand and the subsequent matter-domination (and possibly Λ-domination) of the current time. We further assume sharp transitions between these epochs.

A mode with wavenumber k leaves the horizon when the scale factor and Hubble parameter have values a_k and H_k respectively, where $k = a_k H_k$, and by definition $a_k = e^{-\mathcal{N}(k)} a_{\rm end}$. Therefore

$$\frac{k}{a_0 H_0} = \frac{a_k H_k}{a_0 H_0} = e^{-\mathcal{N}(k)} \frac{a_{\rm end}}{a_{\rm reh}} \frac{a_{\rm reh}}{a_{\rm eq}} \frac{H_k}{H_{\rm eq}} \frac{a_{\rm eq} H_{\rm eq}}{a_0 H_0} \ , \tag{1.80}$$

where the subscripts 'reh' and 'eq' denote values at the end of reheating and at matter-radiation equality respectively. Using standard cosmological relations (see [51] for details) this can be rearranged to give

$$\mathcal{N}(k) = -\ln \frac{k}{a_0 H_0} + \frac{1}{3} \ln \frac{\rho_{\rm reh}}{\rho_{\rm end}} + \frac{1}{4} \ln \frac{\rho_{\rm eq}}{\rho_{\rm reh}} + \ln \sqrt{\frac{V_k}{3M_P^2}} \frac{1}{H_{\rm eq}} + \ln 219 \Omega_m h \ . \tag{1.81}$$

If we further assume that there is no significant drop in energy density in the final stages of inflation, so that $V_k = \rho_{\rm end}$, then the number of e-folds that could have occurred between the relevant mode leaving the horizon and the end of inflation is maximal for the case of instantaneous reheating $\rho_{\rm end} = \rho_{\rm reh}$. This upper bound is

$$\mathcal{N}_{\rm max}(k) = 66.9 - \ln \frac{k}{a_0 H_0} + \frac{1}{4} \ln \sqrt{\frac{V_k}{M_P^4}} \ , \tag{1.82}$$

after substituting numerical values for known quantities. Normally we wish to know this result for either the horizon scale $k_0 = a_0 H_0$ or for the pivot scale, $k_{\rm pivot}$, at which the WMAP values for the power spectrum amplitude and spectral index are quoted. Either can be substituted in to obtain the desired bound.

1.3.4 Scalar fluctuation power spectrum

As discussed in Section 1.2.2, models of inflation can make predictions for the observed power spectrum of perturbations Φ, δ, etc. In this section we will discuss the prediction for the simplest slow-roll models of inflation and show that they generically give a power spectrum that is close to scale-invariant. In discussing the connection to the power spectrum of observable quantities, as in Sections 1.2.3 and 1.2.4 and in subsequent chapters, we assume that it is only the fluctuations of the inflaton field that are responsible for generating the primordial density perturbation, i.e., that there are no contributions from topological defects, isocurvature perturbations and so on. Also, in order to make predictions about the curvature perturbation generated by inflation we make the assumption that the inflaton field perturbation $\delta\phi$ has negligible interaction with other fields, so that linear perturbation theory is valid during inflation. We will restrict ourselves to discussion of the scalar power spectrum, as for the inflation models considered in this thesis, the tensor signal is always small.

In discussing the perturbations during inflation, it is most convenient to use a gauge with spatially flat slicing so that the line element is

$$ds^2 = -(1+2A)dt^2 - 2aB_{,i}dx^i dt + \delta_{ij}a^2 dx^i dx^j \ , \tag{1.83}$$

and the perturbations are specified by the functions A and B. In this gauge the perturbations of the scalar field do not couple to those of the metric in the slow-roll limit [52, 26]. Working in Fourier space, it can be shown that the linearly perturbed part of the $\nu = 0$ component of the conservation equation $T^{\mu}{}_{\nu;\mu} = 0$ is

$$\ddot{\delta\phi} + 3H\dot{\delta\phi} + \left[V'' + \left(\frac{k}{a}\right)^2\right]\delta\phi = 0 \ . \tag{1.84}$$

The assumption of slow-roll ensures that the potential is close to flat, and so V'' will be negligible until at least a few Hubble times after horizon exit. Thus, to a good approximation, we may treat the inflaton as a massless, minimally coupled field, so its quantum-mechanical fluctuations are Gaussian random fields with [53]

$$\frac{k^3}{2\pi^2}\langle|\delta\phi|^2\rangle = \left(\frac{H}{2\pi}\right)^2 \ . \tag{1.85}$$

Once the relevant modes have exited the horizon, the fluctuations can be treated as a classical quantity with power spectrum $P_{\delta\phi} = (2\pi^2/k^3)(H/2\pi)^2$.

The gauge-invariant quantity ζ introduced in Section 1.2.2 can be written as

$$\zeta = -\frac{H}{\dot{\phi}}\delta\phi \tag{1.86}$$

in spatially flat slicing, independently of any assumptions about slow-roll. Therefore the power in ζ can easily be related to the power in $\delta\phi$:

$$P_\zeta = \left(\frac{H}{\dot{\phi}}\right)^2 P_{\delta\phi} \ . \tag{1.87}$$

Using Eqs. (1.73)-(1.74), we may write this explicitly as

$$P_\zeta(k) = \left(\frac{2\pi^2}{k^3}\right)\frac{1}{24\pi^2 M_P^4}\frac{V}{\epsilon} \ , \tag{1.88}$$

where the right hand side is to be evaluated at horizon exit for the mode, $k = aH$.[6] The final piece of the jigsaw is to note that, after the end of inflation, ζ can be related to the metric perturbation in the conformal Newtonian gauge again:

$$\zeta = -\frac{3}{2}\Phi \ . \tag{1.89}$$

As ζ is conserved outside the horizon for adiabatic perturbations, we can then simply use $P_\Phi = \frac{4}{9}P_\zeta$ in order to obtain the power spectrum of fluctuations in Φ, which we have used in Sections 1.2.3 and 1.2.4.

Unfortunately there are a few different ways of defining the power spectrum in the literature. The definition used by the WMAP team, which is what we will need to compare our predictions to, differs from the definition used above by a factor of $(k^3/2\pi^2)$. Therefore, in order to enable comparison, we shall henceforth use the quantity $\mathcal{P}_\mathcal{R} \equiv (k^3/2\pi^2)P_\zeta$ instead of P_ζ to refer to the curvature power spectrum.[7]

The variation of $\mathcal{P}_\mathcal{R}$ with wavenumber k is of great importance when comparing model predictions to observation. We are interested in the *spectral index* n_s, defined by $n_s(k) - 1 \equiv d\ln\mathcal{P}_\mathcal{R}/d\ln k$. The motivation for this definition is clear in the case that n_s itself is independent of k, as then the power spectrum has the pure power-law form

---

[6] Strictly speaking it should be evaluated a few Hubble times *after* horizon exit, but as the right hand side is slowly varying, the fractional error introduced by taking $k = aH$ is of order the slow-roll parameters.

[7] As noted earlier, the quantity $\mathcal{R}$ is used instead of ζ by many authors. However, in the long-wavelength limit the two are equal and we do not wish to suggest any different. Our choice of subscript $\mathcal{R}$ is merely intended as an aid to easily distinguish the two definitions of the power spectrum, while still corresponding to notation used in the literature.

$\mathcal{P}_\mathcal{R} = Ak^{n_s-1}$. Since H is slowly varying we have $d\ln k = d(\ln(aH)) \simeq d\ln a = Hdt$. Therefore, after a bit of algebra, Eqs. (1.73) and (1.88) and the definition of $\mathcal{P}_\mathcal{R}$ lead to

$$n_s - 1 = 2\eta - 6\epsilon \ . \tag{1.90}$$

The running of the spectral index, $n_{\rm run}$, can be obtained by a similar calculation:

$$n_{\rm run} \equiv \frac{dn_s}{d\ln k} = -16\epsilon\eta + 24\epsilon^2 + 2\xi^2 \ , \tag{1.91}$$

where $\xi \equiv M_P^4(V'V'''/V^2)$.

In single-field slow-roll models, this calculation suffices to specify the power spectrum. As ϵ, $|\eta| \ll 1$ and ξ is of second order in $\{\epsilon,\ \eta\}$, deviations from scale-invariance are small, $n_s \simeq 1$. This is the form of the spectrum assumed in the ΛCDM cosmological model. It is found to be *consistent* with the current WMAP observations [15], though it is not possible to determine the shape of the primordial power spectrum *independently* of the other assumptions of the ΛCDM model. The single-field slow-roll description provided here is certainly the simplest, but not the only, model of inflation. Single- and multiple-field models which incorporate brief departures from slow-roll will tend to break scale-invariance and generate distinct features in the primordial power. In such a case, an entirely different cosmology may be equally consistent with the CMB data. We shall return to this point in Chapter 3.

1.3.5 Building inflationary models

When speaking of a “model” of inflation, at the most basic level one could mean nothing more than a form for the potential $V(\phi)$ for the inflaton field, perhaps including the effect of a second field ψ if considering a hybrid inflation model, such that a satisfactory power spectrum of density perturbations can be generated to match the constraints on $\mathcal{P}_\mathcal{R}$, n_s and $n_{\rm run}$ obtained from WMAP, while also matching the constraints on the number of e-folds discussed above.

For most theorists, however, this alone is not satisfactory. A “model” of inflation, in the true sense of the word, should make contact with particle physics theory, by identifying the field ϕ with one of the scalar fields in a relevant extension of the Standard Model of particle physics. In this case, depending on the identification of the field it will have various interactions with other fields, and various terms in the

Lagrangian that are allowed by the symmetries of the theory. Some of the parameters of the Lagrangian might be already constrained from other particle physics considerations. Radiative corrections to the potential might be important and need to be accounted for in determining its flatness. We also want inflation to end in the right vacuum state so that the decay of the inflaton can excite Standard Model degrees of freedom for nucleosynthesis, and generate the required abundance of cold dark matter. One of the problems for many models of inflation, especially those based on branes in a higher-dimensional space, is to ensure that the inflaton decays into Standard Model end products. In addition, some supersymmetric models of inflation can generate their own exotic relics (gravitinos) on which there are stringent constraints from the standard cosmology [54], and avoiding observational bounds may place restrictions on the reheating temperature.

All these considerations act as extra constraints on the types of potentials that it is possible to obtain. The art of model-building is to choose a scalar field for which it is possible to obtain a potential along which inflation is possible, while simultaneously satisfying all of these constraints.

From this viewpoint, models which cannot be easily connected with particle physics in this manner—for example, monomial potentials such as $V(\phi) \propto \phi^2$ or $\propto \phi^4$ [55] which require the field to have values $\phi \gtrsim M_P$, so that stability of the potential to radiative corrections is an issue [56, 57]—are justifiably considered less desirable. Needless to say, however, there are still many models of inflation which *can* be embedded in some extension of the Standard Model, though none in the Standard Model itself. Within these models, one may distinguish between “natural” models and those that are in some way fine-tuned.

Fine-tuning issues

In Section 1.3.1 we invoked several cosmological fine-tuning problems in order to justify the need for inflation. It would therefore be undesirable for the model of inflation itself to require fine-tuning in order to be successful. However, in considering fine-tuning issues we must separate the question of fine-tuning of the parameters of the Lagrangian in order to obtain the inflationary potential from that of the fine-tuning of initial conditions.

A model of inflation may be considered fine-tuned in the first sense if, in order to satisfy the required cosmological conditions (that the potential be sufficiently flat and stable to radiative corrections, etc.), some correlations need to be imposed between the terms of the Lagrangian over and above those that are imposed by the symmetries of the particle physics model in question (for a discussion, see [58]). The degree of fine-tuning required is often quantified by a "fine-tuning parameter", the value of which reflects the accuracy to which the relevant terms must be tuned against each other. If a high degree of tuning is required for successful inflation, the particle physics embedding of the model must be regarded as less satisfactory, despite any other virtues it may possess.

The question of the initial conditions required for inflation is a more difficult one. In fact it has been shown that with Einstein gravity, inflation can only start in a patch of the universe that is already homogeneous over scales larger than the horizon scale [59]. Therefore inflation is in fact not a true solution to the problem of the observed homogeneity of the universe (though it remains the best known mechanism for generating primordial perturbations with the observed form of super-horizon correlations) as it itself appears to require a fine-tuned initial condition to start! There is also the question of how the energy density of the pre-inflationary universe came to be dominated by a single scalar field in the first place. In some sense this is unsurprising: entropy increases with time, so unless our universe has infinite entropy, as we go to epochs further and further back in time we expect to find initial states that have lower entropy and appear more and more fine-tuned. The merit of the inflationary paradigm is therefore not that it solves these fine-tuning problems, but that it provides genuine quantitative predictions about the perturbations in the early universe, which may be compared to observations.

On top of this, some models of inflation require that the initial location of the field ϕ on the potential be fine-tuned to a greater or lesser degree for the flatness condition of the potential to be satisfied.[8] Obviously it is preferable to remove the need for this fine-tuning to the extent possible, but this will not address the larger problem of unlikely initial conditions for inflation.

---

[8] Chaotic inflation models [55] based on monomial potentials are sometimes cited as examples which require only generic initial conditions, but this depends on different interpretations of the measure of likelihood [60]. In any case, as mentioned above, these models pose their own problems because of super-Planckian VEVs.

1.4 Outline of the thesis

In the course of this chapter, we have discussed a variety of different aspects of modern cosmology. In this concluding section we would like to tie together the different threads and provide a coherent link between the topics discussed in greater detail in subsequent chapters.

We take as our starting point the standard ΛCDM cosmological model, and the constraints on the parameters that are derived from fitting this model to data, in particular to the WMAP observations of the CMB anisotropies. The latest data from the WMAP 7-year data release provides constraints on the amplitude and spectral index of the primordial power spectrum of the nearly scale-invariant form obtained from slow-roll inflation [15]. Given this information, the first task is to ask what models of inflation—in the broader particle physics sense discussed in Section 1.3.5—can meet these constraints. In Chapter 2 we discuss one particular model of this class, which is based on flat directions in the minimal supersymmetric extension to the Standard Model (MSSM). We consider the issue of fine-tuning, which is a common problem with viable inflationary models. We then propose a mechanism for greatly reducing the required fine-tuning in this case, which requires an extension of the particle physics embedding from the MSSM. This is based on work carried out with Shaun Hotchkiss and Anupam Mazumdar, and published in [61].

Flat direction fields are in fact generically present in supersymmetric theories, and can have many important cosmological consequences, as detailed in [62]. In particular, if flat directions that are coupled to the inflaton field undergo phase transitions, the slow-roll assumptions made in Section 1.3.4 may be violated, generating a feature in the primordial power spectrum. This naturally raises the question of whether such features are constrained by the data. In Chapter 3, which is based on [63] published with Subir Sarkar, we relax the standard assumptions of ΛCDM and show that in fact a specific model with a feature in the primordial power allows an inhomogeneous cosmology with a metric different to FRW and no dark energy to simultaneously fit data from the CMB anisotropies, SNe Ia luminosity distances, local Hubble measurements, measurements of the BAO scale and nucleosynthesis constraints. This demonstrates that although ΛCDM provides a consistent fit to the available data, it is not necessarily the only model which can do so.

In light of this, it is then pertinent to ask whether there are any other observations which can be used to test the consistency of the ΛCDM paradigm. In Chapter 4 we consider the imprint that large-scale structure leaves on the CMB through the integrated Sachs-Wolfe effect. We show that a recent observation of this effect is indeed in tension with the theoretical expectation from ΛCDM, suggesting that the real universe may be substantially more inhomogeneous than is assumed when using an FRW background treatment. This chapter is based on work with Shaun Hotchkiss and Subir Sarkar [64] that has been submitted for publication.

Chapter 2

Fine tuning in MSSM inflation

2.1 Introduction

As discussed in Chapter 1, an outstanding goal for the theory of primordial inflation is to connect it to particle physics, and in particular to the Standard Model (SM) and its extensions (for reviews of attempts to embed inflation in physics beyond the SM, see [32, 57]). For this to happen it would appear that a low inflationary scale and sub-Planckian values of ϕ are necessary, so that an effective field theory description of the inflaton potential is possible.

These features are naturally present when inflation is generated about a point of inflection in the potential (i.e., a point where the second derivative V'' vanishes), due to the flatness of the potential at the inflection point. Such a point of inflection can be found in models based on gauge invariant D-flat directions of the minimal supersymmetric Standard Model (MSSM), as was first pointed out in [65, 66]. In these examples, the inflaton was postulated to be a gauge-invariant combination of either squark or slepton fields (the supersymmetric partners of the quarks and leptons). In the limit of an unbroken supersymmetry (SUSY) the potential is exactly vanishing, but accounting for soft SUSY-breaking terms and non-renormalizable terms leads to a potential with an inflection point which can support slow-roll inflation.

Subsequently, another model was proposed, based on a flat direction of the MSSM with the addition of right-handed (RH) neutrino supermultiplets, with a purely renormalizable potential [67, 68].[1] Both of these models were shown to have regions of parameter space where it was possible to produce a power spectrum of perturbations

---

[1] In this case the SM gauge group is extended by an additional $U(1)_{B-L}$, where B and L denote the baryon and lepton numbers respectively.

consistent with observational constraints, while also having interesting implications for dark matter searches and particle physics phenomenology [68, 69, 70].

The advantage of such a particle physics embedding is that the model parameters are motivated by low scale SUSY within a *visible* sector. Therefore, the couplings of the inflaton to (MS)SM degrees of freedom are determined, and—at least in principle—measurable in ground-based laboratory experiments such as the Large Hadron Collider. This provides the opportunity to track the thermal history of the universe after inflation, as details of reheating, thermalization and generation of the baryon asymmetry are all sensitive to these couplings (for an example of such a calculation, see [71]). This is in contrast to many other particle physics models of inflation, in which the inflaton is either an *ad hoc* gauge singlet, or a hidden sector field with unknown couplings.

This significant advantage of such models is unfortunately offset by the fact that in order to maintain sufficient flatness of the potential in the face of radiative corrections, the parameters of the potential need to be fine-tuned to several orders of perturbation theory [66, 72]. At low scales this fine-tuning is very acute and a significant challenge to overcome. For high or intermediate scale inflation the required tuning can be reduced to some extent but remains problematic. Although the string landscape can perhaps naturally account for the fine-tuning of soft SUSY breaking terms from degenerate vacua [73], a dynamical solution to the problem is desirable.[2]

In this chapter we attempt to address this issue for the model with the purely renormalizable potential introduced in [67]. The potential in this case has the form

$$V(\phi) = A\phi^2 - C\phi^3 + B\phi^4 \ , \tag{2.1}$$

where the parameters A, B and C are determined by the particle physics embedding as discussed below. The inflaton potential is required to meet two sets of cosmological constraints, as discussed in Sections 1.3.3 and 1.3.4: inflation must generate a primordial power spectrum matching observational constraints, and it must last for an appropriate number of e-folds after the CMB scales have left the horizon. We will discuss the implications of these constraints on the allowed regions of parameter

---

[2] In the context of a non-renormalizable potential for inflection point inflation, a different perspective on the fine-tuning problem has been presented in [70]. Using the renormalization group equations, the tuning of the ratio of SUSY breaking terms can instead be viewed as an equal tuning of the non-renormalizable coupling.

space for the model, and the fine-tuning. The fine-tuning required to fit the spectrum constraints can be reduced by simply raising the scale of inflation, as first pointed out in [74]. However, on its own this violates the requirement to generate a suitable number of e-folds. We propose to include a new hybrid scalar field which provides a vacuum energy while it remains trapped in a false minimum. On being released from this false minimum, the field rolls quickly to its true minimum and brings slow-roll to a premature end, in exactly the same manner as hybrid inflation [75, 76]. This extension can significantly reduce the amount of fine-tuning required in the model, while still matching observations and the e-fold constraint, without ruining any of the attractive features of the MSSM embedding.

In Section 2.2 we briefly discuss the aspects of the SUSY embedding of Eq. (2.1), define the measure of fine-tuning of the potential, and obtain expressions for the slow-roll parameters that are used later on. The bulk of Section 2.2 is however devoted to a calculation of the e-fold number that corresponds to the observed CMB scales. This calculation is not original but it provides important results that are used in Sections 2.3 and 2.4, which contain the new results of this chapter. In Section 2.3 we investigate the region of parameter space that allows for a period of inflation that is consistent with the e-fold constraint and constraints from the WMAP 7-year power spectrum [15]. In Section 2.4, we introduce the hybrid extension to the model and show how it reduces the required fine-tuning.

2.2 Slow-roll parameters and the e-folding number

Let us consider a re-parameterisation of Eq. (2.1):

$$V\left(|\phi|\right) = \frac{m_\phi^2}{2}|\phi|^2 + \frac{h^2}{12}|\phi|^4 - \frac{Ah}{6\sqrt{3}}|\phi|^3 \, , \tag{2.2}$$

The potential will have a region suitable for inflation if the mass term satisfies the condition $A \approx 4m_\phi$. This form of the potential was first motivated in [67, 68] from a low-scale extension of the MSSM with an additional $U(1)_{\mathrm{B-L}}$, which can relate the light neutrino masses to the flatness of the inflaton potential. Here ϕ contains the RH sneutrino field $\tilde{N}$, Higgs H_u, and the left handed slepton field $\tilde{L}$. The potential can be derived from the superpotential term

$$W \supset h\mathbf{N}\mathbf{H_u}\mathbf{L} \, , \tag{2.3}$$

where $\mathbf{N}$, $\mathbf{L}$ and $\mathbf{H_u}$ are superfields, h is the Yukawa coupling, and generation indices have been omitted. In the basis where neutrino masses are diagonalized, $\mathbf{N_3H_uL_3}$ is a D-flat direction under the extended gauge group. Then the scalar field corresponding to the flat direction (the inflaton) is

$$\phi = \frac{\tilde{N}_3 + H_u^2 + \tilde{L}_3^1}{\sqrt{3}} \,, \tag{2.4}$$

where the superscripts refer to the lower and upper weak isospin components of $\mathbf{H_u}$ and $\mathbf{L_3}$ respectively.

Although $\mathbf{N_3H_uL_3}$ is a D-flat direction, it contributes an F-term obtained from $F = -(\partial W/\partial\phi)^*$, which lifts the potential through $V = V_F = |F|^2$. This accounts for the quartic term in Eq. (2.2). The potential is also lifted by the soft SUSY-breaking terms, i.e., the quadratic and cubic terms in Eq. (2.2). The mass of the inflaton is given in terms of the soft masses of $\tilde{N}_3$, H_u and $\tilde{L}_3$ by:

$$m_\phi^2 = \frac{m_{\tilde{N}}^2 + m_{H_u}^2 + m_{\tilde{L}}^2}{3} \,. \tag{2.5}$$

Therefore, the mass of the Higgs field is tied to the mass of the inflaton.

If $h \sim 10^{-12}$, it is possible to relate the smallness of the observed neutrino masses to the flatness of the inflaton potential [67, 68]. We stress these points because their implications are re-considered in Sections 2.3 and 2.4.

Following convention, we parameterize the fine-tuning of the potential as

$$\delta \equiv \frac{A^2}{16 m_\phi^2} \equiv 1 - \frac{\beta^2}{4} \,, \tag{2.6}$$

where β may be either real or imaginary, corresponding to $\delta < 1$ or $\delta > 1$ respectively. The ratio of SUSY-breaking terms $A/4m_\phi$ must be tuned to accuracy $\mathcal{O}(|\beta|^2) \ll 1$. Under this condition, the potential has a point of inflection at $\phi = \phi_0$ such that $V''(\phi_0) = 0$, where the point of inflection is[3] $\phi_0 = \sqrt{3} m_\phi / h$. In the vicinity of ϕ_0 the potential can be written as the truncated Taylor expansion:

$$V(\phi) = V_0 + \alpha(\phi - \phi_0) + \frac{\gamma}{6}(\phi - \phi_0)^3 \,, \tag{2.7}$$

where $V_0 \equiv V(\phi_0)$, $\alpha \equiv V'(\phi_0)$ and $\gamma \equiv V'''(\phi_0)$. It is worth emphasising that although the potential in Eq. (2.1) or Eq. (2.2) superficially resembles the monomial

---

[3] In fact there are two solutions for the inflection point ϕ_0 but we shall only consider the larger one. Our analysis is valid as long as the the field is always in the vicinity of this point.

forms $\propto \phi^2$ or $\propto \phi^4$ which have been extensively studied, in the region that inflation occurs the potential is actually the Taylor expansion, Eq. (2.7), which accounts for the difference in parameter constraints. This expansion will be valid provided that

$$|\alpha| \gg \left|\frac{d^m V}{d\phi^m}(\phi_0)\right| |\phi_e - \phi_0|^{m-1} \tag{2.8}$$

and

$$|\gamma| \gg \left|\frac{d^m V}{d\phi^m}(\phi_0)\right| |\phi_e - \phi_0|^{m-3} \tag{2.9}$$

for $m \geq 4$, where ϕ_e is the value of the field at the end of inflation. The terms in this expansion are [66] $V_0 \propto m_\phi^2 \phi_0^2 \propto m_\phi^4/h^2$; $\gamma \propto m_\phi^2/\phi_0 \propto m_\phi h$ and

$$\alpha = \frac{\sqrt{3} m_\phi^2}{4h} \beta^2 \ + \mathcal{O}(\beta^4)\ . \tag{2.10}$$

Note that the Hubble rate during inflation is $H_{\rm inf} \approx V_0^{1/2}/\sqrt{3} M_P$.

From the form of the potential in Eq. (2.7) we may write the slow-roll parameters $\epsilon \equiv (M_P^2/2)(V'/V)^2$ and $\eta \equiv M_P^2(V''/V)$ explicitly as:

$$\begin{aligned} \epsilon(\phi) &= \frac{M_P^2}{2V_0^2}\left(\alpha + \frac{\gamma}{2}(\phi - \phi_0)^2\right)^2 & (2.11)\\ \eta(\phi) &= -\frac{\gamma M_P^2}{V_0}(\phi_0 - \phi)\,. & (2.12) \end{aligned}$$

If inflation ends at field value $\phi = \phi_e$, the number of e-folds of inflation produced as the field rolls from ϕ to ϕ_e is given by

$$\mathcal{N}(\phi) = \int_{\phi}^{\phi_e} \frac{H d\phi}{\dot{\phi}} = \frac{V_0}{M_P^2}\sqrt{\frac{2}{\alpha\gamma}}\left[F(\phi_e) - F(\phi)\right] \tag{2.13}$$

where $F(z) \equiv \mathrm{arccot}\left(\sqrt{\frac{\gamma}{2\alpha}}(z - \phi_0)\right)$. When β is imaginary ($\alpha < 0$) the corresponding expression is found by analytic continuation. While the Taylor expansion of Eq. (2.7) is not strictly necessary to calculate $\mathcal{N}$ it does allow for this closed-form, analytic expression for $\mathcal{N}$. Calculating $\mathcal{N}$ numerically gives equivalent results but significantly lengthens the computation time. This is particularly true of the calculation in Section 2.4, where (2.13) must be inverted. This would make the production of Figs. 2.1 and 2.2 unfeasible.

Slow-roll ends at the field value ϕ_e at which $|\eta| \sim 1$. This is determined by Eq. (2.12):

$$\phi_e \sim \phi_0 - \frac{V_0}{\gamma M_P^2}. \tag{2.14}$$

Inflation itself will end very soon after the violation of slow-roll, within a fraction of an e-fold. Under the assumption that the energy scale of inflation is roughly constant during inflation (which is valid as $\epsilon \ll |\eta| \ll 1$) we can make use of Eq. (1.82) for the maximum number of e-foldings between the time when the observed perturbations were generated and the end of inflation. Using the WMAP pivot scale $k_{\text{pivot}} = 0.002\ \text{Mpc}^{-1}$ and the current best-fit values of the cosmological parameters, we find this to be[4]

$$\mathcal{N} \leq \mathcal{N}_{\text{pivot}} \equiv 64.7 + \ln\left(\frac{V_0^{1/4}}{M_P}\right). \tag{2.15}$$

Note that $\mathcal{N}_{\text{pivot}}$ is dependent on the scale of inflation and therefore on the parameters m_ϕ and h.

In fact, if the origin of the inflaton is given by Eq. (2.3), the time scale for transferring the inflaton energy into relativistic species is negligible compared to the Hubble time [66] so we can treat this process to be instantaneous and the inequality in Eq. (2.15) is saturated.[5] In order to account for errors in the value of $\mathcal{N}_{\text{pivot}}$ we can take the constraint to be $\mathcal{N} \simeq \mathcal{N}_{\text{pivot}}$. In fact, the error introduced due to uncertainties in the cosmological parameter values, the slow-roll approximation and the assumption of instantaneous transition from radiation- to matter-domination of the universe is very small, of order at most one e-fold. We can therefore safely require the e-fold constraint to be satisfied within ±5 e-folds. As shown in the next section, even this conservative assumption imposes tight constraints on the allowed values of m_ϕ and h.

2.3 WMAP constraints on (m_ϕ, h) and fine-tuning of parameters

If the perturbations relevant to the CMB spectrum observed today were generated at a field value $\phi = \phi_{\text{CMB}}$, the amplitude of the power spectrum and the scalar spectral

---

[4] We match our power spectrum constraints to the WMAP observational results, which are quoted at the pivot scale k_{pivot}, not the horizon scale $a_0 H_0$. The difference between these two scales corresponds to ~ 2 e-folds.

[5] Note that to determine the necessary number of e-folds we only need to know the point at which the equation of state changes from inflaton- to radiation- or matter-domination [77]. The time scale for all the relativistic species generated to subsequently attain thermal equilibrium could be different.

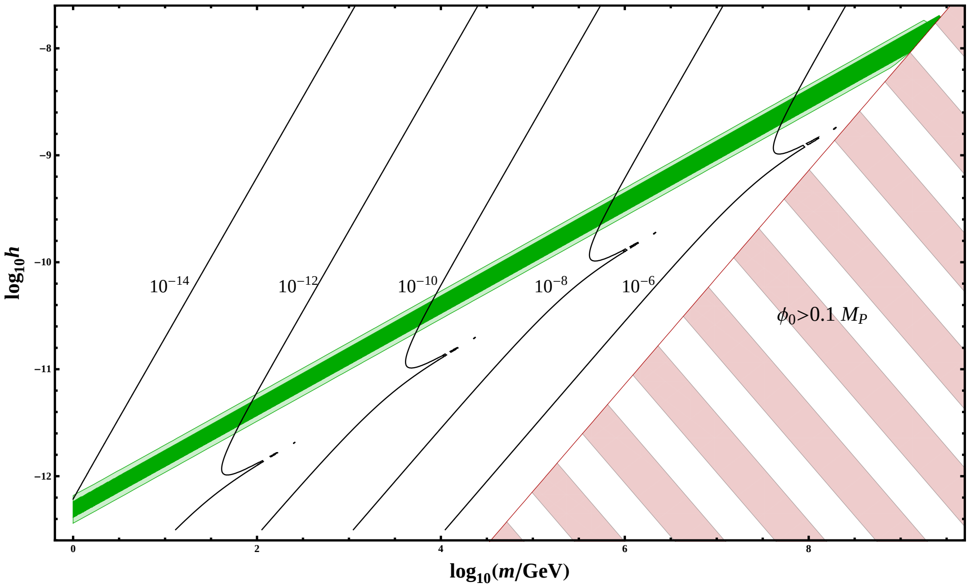

Figure 2.1: Regions of parameter space for the potential in Eq. (2.2) that satisfy the WMAP 7-year constraints on the amplitude and spectral index of the power spectrum and also match the e-fold constraint. We minimize $|\mathcal{N} - \mathcal{N}_{\mathrm{pivot}}|$ over the range of β allowed by the 95% C.L. constraints on $\mathcal{P}_{\mathcal{R}}$ and n_s. The dark green central contour shows the region for which $\left(|\mathcal{N} - \mathcal{N}_{\mathrm{pivot}}|\right)_{\mathrm{min}} \leq 1$, the light green contour is for $\left(|\mathcal{N} - \mathcal{N}_{\mathrm{pivot}}|\right)_{\mathrm{min}} \leq 5$. Contour lines of $|\beta|$ are shown in black, for the values of $|\beta|$ indicated. The red striped region is excluded by the requirement that the inflection point be much less than the Planck scale.

index are given by:

$$\begin{aligned} \mathcal{P}_{\mathcal{R}} &= \frac{1}{24\pi^2 M_P^4}\frac{V_0}{\epsilon\left(\phi_{\mathrm{CMB}}\right)} &\quad (2.16)\\ n_s &= 1 + 2\eta\left(\phi_{\mathrm{CMB}}\right) - 6\epsilon\left(\phi_{\mathrm{CMB}}\right). &\quad (2.17)\end{aligned}$$

As $\epsilon \ll |\eta|$ we can approximate the spectral index as

$$n_s = 1 + 2\eta\left(\phi_{\mathrm{CMB}}\right). \tag{2.18}$$

The latest data from the WMAP 7-year release suggest a power spectrum with $\mathcal{P}_{\mathcal{R}} = (2.43 \pm 0.11) \times 10^{-9}$ and a spectral index of $n_s = 0.967 \pm 0.014$ for models with no 'running' of the spectral index [15]. (In all the cases considered in this analysis, the running is found to be negligible, so we do not mention it explicitly.) Given these constraints Eqs. (2.16) and (2.18) may be inverted to obtain the values ϵ_{best} and η_{best} that will produce the required best-fit power spectrum, and the range of ϵ and η

values that lie within the 95% C.L. The field value $\phi_{\rm CMB}$ at which the perturbations are generated is then chosen to be

$$\phi_{\rm CMB} = \phi_0 - \frac{V_0}{\gamma M_P^2} |\eta_{\rm best}| \ . \tag{2.19}$$

Equation (2.11) then requires that the first derivative of the potential satisfy

$$\alpha_{\rm CMB} = \frac{\sqrt{2\epsilon_{\rm best}} V_0}{M_P} - \frac{V_0^2 \eta_{\rm best}^2}{2\gamma M_P^4} \ , \tag{2.20}$$

from which we obtain the required value of the fine-tuning parameter $\beta_{\rm CMB}$ using Eq. (2.10).

Equation (2.20) is the condition for a given potential of the form of (2.7) to have a region in which perturbations with the best-fit power spectrum can be produced. Equation (2.19) identifies this region. For any combination of parameters m_ϕ and h it is always possible to make this choice, but the value of $\beta_{\rm CMB}$ thus obtained may not simultaneously satisfy $\mathcal{N} \simeq \mathcal{N}_{\rm pivot}$, with $\mathcal{N}$ calculated using Eqs. (2.13), (2.14), (2.19) and (2.20). However, one has the freedom to choose a range of values for β around $\beta_{\rm CMB}$, governed by the acceptable range in the slow-roll parameters. If for any value of β in this range $\mathcal{N} \simeq \mathcal{N}_{\rm pivot}$ to within the desired uncertainty then that combination of m_ϕ and h can be consistent with current WMAP limits.

In Fig. 2.1 we show the allowed range of parameters m_ϕ and h. We impose the conservative constraint that the inflaton field value should be sub-Planckian, $\phi_0 < 0.1 M_P$, and require that the conditions in Eqs. (2.8) and (2.9) be satisfied. Note that if β is imaginary the potential develops a false minimum separated from the true minimum by a barrier. Near the peak of this barrier a self-reproduction regime exists where quantum diffusion dominates over the classical evolution [66]. We have also checked that none of the allowed regions in Fig. 2.1 are in this regime. Two different contours are shown, using the criteria $|\mathcal{N} - \mathcal{N}_{\rm pivot}|_{\rm min} \leq 5$ and $|\mathcal{N} - \mathcal{N}_{\rm pivot}|_{\rm min} \leq 1$ over the allowed range of β to define consistency with the WMAP data.[6] Also plotted are contour lines of $|\beta_{\rm CMB}|$ — the kink in the contours indicates the transition from imaginary to real $\beta_{\rm CMB}$.

---

[6] These results may be compared with those obtained in [65, 78]. The small difference probably arises because the authors of these papers have taken $\mathcal{N} = 50$ throughout the range of the plot, without accounting for changes in $\mathcal{N}_{\rm pivot}$ due to the change in the energy scale of inflation.

Note that for *all* (m_ϕ, h) coordinates in Fig. 2.1 it is possible to generate perturbations with the desired power spectrum on *some* scales. However for coordinates lying below the contours shown, these scales are too large and have not re-entered our horizon yet, i.e., they correspond to $\mathcal{N} > \mathcal{N}_{\rm pivot}$. The opposite is true for the region above the allowed contour, where not enough e-folds are generated. This region could be accessible if the period of reheating can be delayed by several e-folds [79, 80].

2.3.1 Low scale inflation and fine-tuning

It can be seen that for low inflaton masses, the coupling h is required to be very small for the model to simultaneously match the spectrum and e-fold constraints. Such a low value of h can in fact provide an explanation for the low masses of the neutrinos if the origin of the inflaton is given by Eq. (2.3). If the neutrino is of Dirac type its mass is given by $m_\nu = h\,\langle H_u \rangle$ where $\langle H_u \rangle \simeq 174$ GeV is the Higgs VEV. For $m_\phi \sim 10$ GeV, $h \sim 10^{-12}$, this gives $m_\nu \sim 0.1$ eV. This is in line with current constraints from cosmology and atmospheric neutrino oscillations as detected by the Super-Kamiokande experiment. Note that the inflaton mass, Eq. (2.5), is constrained to be small at the scale of inflation when $\phi \sim \phi_0 \sim 10^{14}$ GeV, during which the relevant perturbations are created. However the inflaton mass evolves with the field value (and the energy scale) due to the SM gauge interactions, and at the LHC scale is much higher [68].

The fine-tuning required for such low scale inflation is very acute: note that for inflaton mass around the TeV scale, we have $\beta_{\rm CMB} \lesssim 10^{-10}$, which corresponds to a severe tuning of the ratio of the SUSY breaking terms.

2.3.2 High and intermediate scale inflation

The inflaton mass can vary over a wide range of scales, subject to an absolute upper bound of $m_\phi \sim \mathcal{O}(10^9)$ GeV obtained from the condition that ϕ during inflation be much less than the Planck scale.

For the model based on the superpotential term in Eq. (2.3) it can be seen that at large m_ϕ the Yukawa coupling h becomes much larger than the upper bound $h \sim 10^{-12}$ for the Dirac mass of the neutrino to not exceed current constraints. If the neutrino is not of Dirac type this does not impose a constraint on the model.

In fact the RH neutrinos can obtain a Majorana mass M through the breaking of the $U(1)_{\rm B-L}$ symmetry, which gives the active neutrino masses by the seesaw relation $h^2 \langle H_u \rangle^2 / M$ [81]. For reasonable values of $M \gtrsim \mathcal{O}(1 \text{ TeV})$ the active neutrino mass remains $h^2 \langle H_u \rangle^2 / M \ll 0.1$ eV, which is acceptable for the lightest neutrino species.

Raising the scale of inflation can somewhat reduce the fine-tuning, but it can be seen that the largest value of $\beta_{\rm CMB}$ is bounded by:

$$\beta_{\rm CMB} \lesssim 10^{-6} \, . \tag{2.21}$$

This still represents a serious problem. We address how to improve on this in the next section.

2.4 A dynamical approach to the fine-tuning

In this section, we will take the origin of the inflaton to be as in Eq. (2.3). The fine-tuning of soft SUSY-breaking parameters is required in order make the potential sufficiently flat. Raising the potential during inflation increases the flatness of the potential, thus ameliorating the tuning [74]. However, this also leads to too many e-folds of inflation, $\mathcal{N} > \mathcal{N}_{\rm pivot}$ and thus cannot by itself provide a consistent solution. When viewed from this perspective, it is clear that both problems can be solved simultaneously by introducing a hybrid mechanism [75, 76] to bring a premature end to inflation.

We therefore introduce a new scalar field, S, which acquires a VEV and couples to the inflaton, thus altering the basic potential in Eq. (2.2) to the form

$$V(|\phi|, S) = \frac{1}{2} m^2 |\phi|^2 + \frac{h^2}{12} |\phi|^4 - \frac{Ah}{6\sqrt{3}} |\phi|^3 + \left(S^2 - V_c^{1/2} \right)^2 + \frac{1}{2} g^2 |\phi|^2 S^2 \, . \tag{2.22}$$

During slow-roll the S field obtains a large effective mass from the VEV of ϕ, i.e., $m_{{\rm eff},S} \gg H_{\rm inf}$, and rolls quickly (within one Hubble time) to get trapped at its local minimum $S \approx 0$. The effect on the potential is to add a constant term V_c which flattens the potential and reduces the required fine-tuning in β. The Taylor expansion of Eq. (2.7) is the same, except for the modification

$$V_0 \to \tilde{V}_0 = V_c + \frac{m_\phi^4}{4h^2} + \mathcal{O}(\beta^2) \, , \tag{2.23}$$

and provided that the conditions in Eqs. (2.8) and (2.9) are satisfied, the analysis of Section 2.3 remains unchanged, except that an extra degree of freedom is now introduced via V_c. This situation persists until ϕ rolls to a critical value ϕ_c such that

$$\phi_c = \frac{2V_c^{1/4}}{g} \, , \tag{2.24}$$

at which point S is released from the origin and rolls to its global minimum, acquiring a VEV $\langle S \rangle = V_c^{1/4}$. Provided that $V_c > V_0$, this leads to a sudden steepening of the potential and brings an end to slow-roll inflation. For reasonable V_c, $\eta_S \equiv m_S^2 M_P^2 / V_c \gg 1$ where $m_S^2/2 = 2V_c^{1/2}$ is the negative mass term for S, and so inflation will end promptly [76]. The VEV acquired by the S field gives an effective mass to the inflaton, $m_\phi^2 \to \tilde{m}_\phi^2 = m_\phi^2 + g^2 V_c^{1/2}$.

From Eqs. (2.20) and (2.10) we see that for any given values of the parameters (m_ϕ, h), increasing V_c—and thus $\tilde{V}_0$—increases $|\beta|$ and alleviates the fine-tuning of the potential. Clearly there is a threshold above which increasing V_c requires β to be imaginary. We check that the field is never required to be in the self-reproduction region of the potential, therefore this does not pose a problem. However, V_c cannot be increased without limit, as with increasing $|\beta|$ there will come a point when the condition in Eq. (2.8) is no longer satisfied and our inflection point analysis will no longer be valid.

For any given parameter combination (m_ϕ, h), after having chosen some value of V_c and the corresponding β_{CMB} which will produce exactly the best-fit WMAP power spectrum amplitude and spectral index, we can invert Eq. (2.13) with $\mathcal{N}$ set equal to $\mathcal{N}_{\mathrm{pivot}}$ to find the appropriate field value ϕ_e for the end of inflation such that scale at which the perturbations with this power spectrum are produced should be exactly the scale observed by WMAP today. Setting the critical value $\phi_c = \phi_e$ we can then find the coupling strength g for which the hybrid scalar field will bring inflation to an end at exactly this field value. In this way we can choose the parameters V_c and g appearing in the potential in Eq. (2.22) while still ensuring that the constraint on the number of e-folds of inflation since the generation of the observed perturbations is still satisfied. This gives us the power to increase the value of β and thus ameliorate the fine-tuning problem.

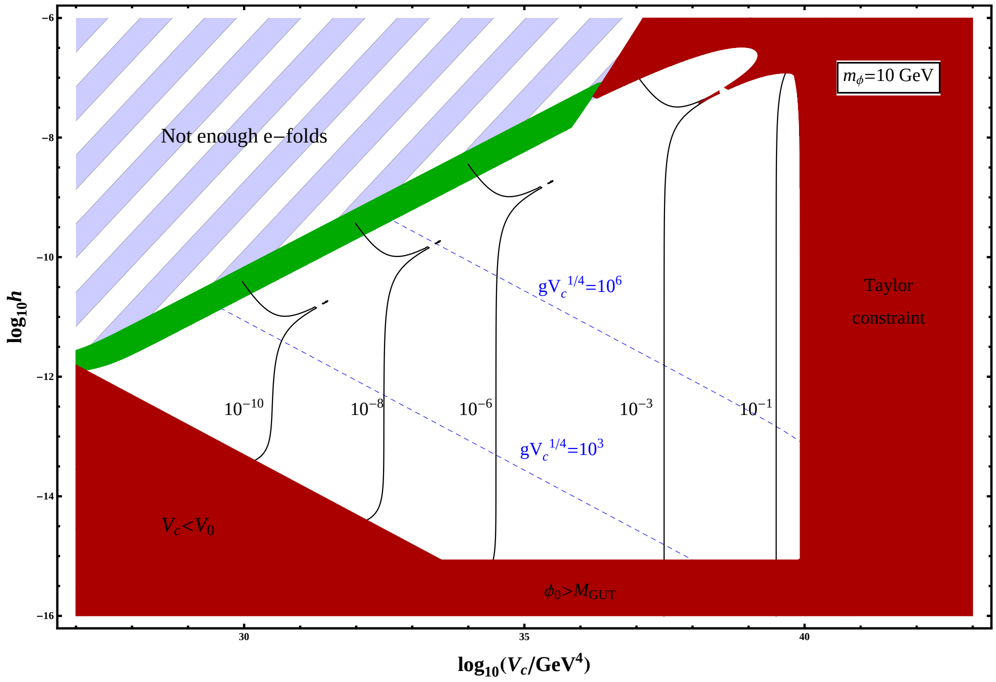

Figure 2.2: Regions of parameter space for the potential in Eq. (2.22) for the choice $m_\phi(\phi_0) = 10$ GeV. The solid red region is inaccessible to our analysis, for the reasons indicated by the text and explained more fully in Section 2.4. In the blue striped region, the slow-roll parameters become large and end inflation with number of e-folds $\mathcal{N} < \mathcal{N}_{\rm pivot} - 5$. The diagonal green contour indicates the region where the number of e-folds of slow-roll inflation is within ± 5 of $\mathcal{N}_{\rm pivot}$ and no hybrid mechanism is needed to end inflation. In the rest of the plot the WMAP constraints can be satisfied by the choice $\beta = \beta_{\rm CMB}$ when the S field ends slow-roll inflation through the hybrid mechanism. Contour lines of $\beta_{\rm CMB}$ are shown in black (solid) and those of $gV_c^{1/4}$ are shown in blue (dashed). For other values of m_ϕ, a qualitatively similar figure is obtained, with appropriate scaling of the axes.

In this analysis for simplicity we disregard any uncertainty in the power spectrum amplitude and spectral index, and in $\mathcal{N}_{\rm pivot}$. Accounting for these will provide a range of acceptable values of the coupling g.

2.4.1 WMAP constraints and solution to the fine tuning problem

Let us consider the case depicted in Fig. 2.2, which shows the $h - V_c$ plane, where we have taken the inflaton mass to be $m_\phi(\phi_0) = 10$ GeV. We impose the following constraints in deciding which regions of the $h - V_c$ plane are accessible to our theory: the field value at the end of inflation, ϕ_e, should be such that the constraints of

Eqs. (2.8) and (2.9) are satisfied and our inflection point analysis is valid; ϕ_0 should not exceed the GUT scale; and $V_c > V_0$ in order for the hybrid mechanism to be able to bring inflation to an end. The solid red contour around the boundary to the bottom and right shows the regions that are excluded by any of these criteria. For some values of V_c and h, the choice of $\beta_{\rm CMB}$ from WMAP spectrum constraints results in a number of e-folds $\mathcal{N}$ that is too low: this region is indicated by the blue striped contour. Clearly the hybrid mechanism cannot make this region accessible to successful inflation but subsequent periods of inflation could do so by reducing $\mathcal{N}_{\rm pivot}$.

The solid green contour extending diagonally upwards shows the region in which the WMAP power spectrum constraints and the e-fold constraint can be simultaneously satisfied, *without* the need for a hybrid end to inflation (as in Fig. 2.1, we use the conservative criterion $\left(|\mathcal{N} - \mathcal{N}_{\rm pivot}|\right)_{\rm min} \leq 5$ when β is allowed to vary over the range allowed by the WMAP constraints). That is, in this region we require the coupling to satisfy $g \geq 2V_c^{1/4}/\phi_e$ with ϕ_e given by Eq. (2.14), but it is otherwise unconstrained. In the rest of the $h - V_c$ plane (left white) the WMAP spectrum constraints and the e-fold constraint can be satisfied if the S field brings inflation to an early end through its coupling to the inflaton as described above. This is the region of interest to us.

In the region of interest we plot contour lines of $|\beta_{\rm CMB}|$ and of $gV_c^{1/4}$. The value of $gV_c^{1/4}$ gives an indication of the effective mass of the inflaton at the end of inflation. It can be seen that, for $V_c \sim 10^{36} - 10^{37}$ (GeV)^4 and Yukawa coupling $h \sim 10^{-14} - 10^{-15}$, we can obtain values as large as

$$|\beta_{\rm CMB}| \sim 10^{-3} \, , \tag{2.25}$$

with the inflaton mass around the TeV scale. This value should be compared with $|\beta_{\rm CMB}| \lesssim 10^{-12}$ for the equivalent case without the action of the S field.

This represents a great reduction in the amount of fine-tuning required in the theory. In fact, a further reduction of the fine-tuning can be achieved, with

$$|\beta_{\rm CMB}| \sim 10^{-1} \tag{2.26}$$

or even larger. For this case the effective mass of the inflaton becomes significantly larger than the TeV scale.

By lifting the scale of inflation, the mechanism proposed in this paper broadens the slow-roll region of the potential, thus reducing the problem of the fine-tuning of the initial value of ϕ. The slow-roll region of the potential, where $|\eta| < 1$, is of width $\Delta\phi \sim \mathcal{O}(|\beta|\phi_0)$. Without the action of the S field, as $\beta \lesssim 10^{-12}$, this region is exceedingly narrow and the initial location of the inflaton on the potential appears to be fine-tuned. However the values of $|\beta_{\rm CMB}| \sim 10^{-1}$ obtained with addition of the hybrid mechanism greatly broaden the slow-roll region and simultaneously reduce the fine-tuning of initial conditions of ϕ. (In [82, 73] a mechanism by which a prior period of false vacuum inflation may resolve the problem of initial conditions without fine-tuning is discussed; the broadening of the slow-roll region achieved here will make this mechanism more easily achievable.)

Note also that $h \sim 10^{-14} - 10^{-15}$ leads to an acceptable Dirac mass $m_\nu \sim 10^{-3} - 10^{-4}$ eV for the lightest neutrino species, and that all values of $h \lesssim 10^{-12}$ can also be obtained. Therefore the explanation for the small neutrino masses made in [67, 68] is not spoiled by our extension of the model.

2.4.2 Motivation from particle physics

The introduction of the singlet field may look like an *ad hoc* choice and would be a backward step if it could not be linked to a particle physics model. In fact such a field can generate an effective mass for the right handed (s)neutrino field through its VEV via an additional superpotential term: $g\mathbf{SNN}$, and can naturally occur within the NMSSM (next to Minimal Supersymmetric Standard Model) [83], where the same scalar S could be responsible for generating an effective μ-term, $\kappa\mathbf{SH_uH_d}$, where $\kappa\langle S\rangle \sim 100$ GeV. The superpotential terms are

$$W = h\mathbf{NH_uL} + g\mathbf{SNN} + W_{NMSSM}\ . \tag{2.27}$$

The required vacuum energy, V_c, can be obtained when the singlet field is settled near its local minimum, $\langle S\rangle \approx 0$, during inflation by virtue of its coupling $g\mathbf{SNN}$. Note that ϕ near inflection point is quite large, $\phi_0 \sim \tilde{N} \sim 10^{14}$ GeV. This induces an effective mass for S. For instance, for the values of parameter which leads to Eq. (2.25), the coupling $g \sim 10^{-6}$, and $m_S \sim g\langle\tilde{N}\rangle \sim 10^8$ GeV, which is much larger than the Hubble expansion rate during inflation. For $V_c \sim 10^{37}$ (GeV)^4, $H_{\rm inf} \sim 10$ GeV. Thus the S field settles down in its local minimum within one Hubble time during inflation. The value of $g \sim 10^{-6}$ is also adequate to generate the right handed (s)neutrino masses after inflation within an expected range: $gV_c^{1/4} \sim 10^3$ GeV. Finally the phase transition happens when $\phi \to \phi_c$, Eq. (2.24).

For larger values of m_ϕ during inflation, achieving a similar increase in $|\beta_{\rm CMB}|$ requires the correction to the inflaton mass after inflation to be greater than $\sim \mathcal{O}(1 \text{ TeV})$. This is acceptable in principle, though requiring the mass to be much larger than the TeV scale is somewhat undesirable, and offsets the reduction of the fine-tuning of the inflation potential.

For the above parameters, the reheat temperature of the universe is sufficiently high to generate the required thermal dark matter abundance. One particularly nice candidate is the lightest right handed sneutrino [68, 84]. However further investigation of how all the MSSM degrees of freedom achieve thermal equilibrium is needed, as in [71]. Preliminary results suggests that the gauge-invariance of the inflaton generates VEV-dependent masses for the $SU(2)$ gauge bosons which delay the process of thermalization and avoid overproduction of thermal gravitinos [85, 79, 86].

2.5 Conclusions

We have discussed the very generic renormalizable inflationary potential introduced in Eq. (2.1) and have analyzed the conditions required for this potential to generate a power spectrum of density perturbations compatible with the latest WMAP 7-year constraints. We have also presented a particular particle physics model for the potential, where the origin of the inflaton is the gauge-invariant flat direction introduced in Eq. (2.3), with a minimal extension of the standard model gauge group. We obtained the constraints on the parameter space for this model from cosmology, which are depicted in Fig. 2.1. The allowed region of parameter space also allows an explanation for the small neutrino mass.

The potential Eq. (2.1) (or equivalently Eq. (2.2)) can be realised in many different theoretical models. As long as the theory allows for rapid transfer of inflaton energy into radiation at the end of inflation, the constraints on parameter space presented in Fig. 2.1 will apply. Therefore our results are quite general.

Although in general the soft SUSY-breaking terms in this potential must be highly tuned against each other, we have shown that this is primarily a result of the requirement to produce the correct number of e-folds of inflation, rather than a constraint imposed by the observed power spectrum. With this perspective, we have presented a simple extension of the model to include a new scalar field which brings an end to inflation through the hybrid mechanism. We showed that this reduces the fine-tuning to very manageable levels, even achieving $|\beta| \sim 10^{-1}$. This result is significant as the required fine-tuning was one of the main objections to the original model. Finally, we argued that this hybrid extension of the model also reduces any need for fine-tuning of the initial value of the inflaton field, which also makes the hybrid extension more attractive.

Chapter 3

Fitting a local void model to cosmological data

3.1 Introduction

So far in this thesis we have considered only the standard ΛCDM cosmological model, in which the universe is spatially flat and dominated by a cosmological constant, with a matter density of only $\Omega_m \sim 0.3$. In Chapter 1, we argued that the low matter density shifts the peak in the matter power spectrum to smaller scales and hence allows a spatially flat model to fit the observed spectrum. Indeed, it was this observation that first prompted a cosmological model including Λ [17]. Subsequently, observations of Type Ia supernovae [1, 11], then measurements of galaxy clustering in the 2dF survey [18] and of CMB anisotropies by WMAP [87] made this the accepted cosmological model. In fact, the ΛCDM model has continued to prove a consistent fit to more recent observations, including the SNe Ia luminosity distance-redshift relation (e.g. [2, 12, 13]), the CMB anisotropies [15], the value of the local Hubble parameter (e.g. [9]), baryon acoustic oscillations (e.g. [16, 40]) and so on.

However, as discussed in Section 1.1.5, the cosmological constant problem and the coincidence problem remain. We do not yet have any fundamental physical explanation of the cosmological constant, and until such an explanation is found, it is certainly worthwhile to ask if alternative models can be found which also fit the data. If such models exist, they may help to illuminate the degeneracies in the data and may lead us to new observational tests of ΛCDM. They serve as a reminder of the assumptions that are (often implicitly) made in theoretical interpretation of cosmological observations. Even within the ΛCDM framework, relaxing some of these assumptions may have profound implications for the estimation of cosmological parameters.

In fact the WMAP results alone do not require a cosmological constant even for a spatially flat cosmology if the assumption of a scale-invariant primordial power spectrum is relaxed.[1] Although the simplest slow-roll inflation models do give such a spectrum, given our present ignorance of the details of the physics underlying inflation, it is well justified to consider alternatives. In [89, 90] it was shown that the WMAP 3-year data can be matched by a flat $\Omega_m = 1$ Einstein-de Sitter (EdS) universe with $h \simeq 0.44$ if the primordial power spectrum contains a feature at scales of $k \sim 0.01 - 0.1\ h\text{Mpc}^{-1}$. Other EdS models with a broken power-law spectrum [91] have also been shown to fit the WMAP data. Moreover, an EdS universe can fit measurements of the galaxy power spectrum if it includes a $\sim 10\%$ component of hot dark matter in the form of massive neutrinos of mass ~ 0.5 eV [89, 90, 91].

The measurements of apparent magnitudes and redshifts of SNe Ia appear to provide clearer evidence of acceleration and therefore the existence of dark energy. However, this relies on interpreting the data in an assumed homogeneous model universe, i.e. in an FRW background. This evidence is *geometrical* in nature, as it is based on interpretations of the distances to the supernovae and the expansion rate of the universe. (The same is true of the evidence from WMAP and the BAO, which are interpretations of another distance scale — the sound horizon at last scattering — projected on the sky.) In homogeneous models based on ordinary matter and gravity, these observations indicate that the predicted distances and expansion rate are too small by a factor of ~ 2 [92].

One approach to this problem is to study light propagation and observables in general inhomogeneous backgrounds, or models with only statistical homogeneity and isotropy—see, e.g., [93]. An alternative approach is to consider specific inhomogeneous toy models in which the effect of the inhomogeneity on observables can be calculated directly. This is mostly done using a Lemaítre-Tolman-Bondi (LTB) metric, an exact solution of Einstein's equations, to describe spherically symmetric inhomogeneities. In some studies, voids modelled by patches of LTB space-time are distributed in a homogeneous background to create a "Swiss-cheese" universe. The

---

[1] The WMAP data alone are consistent with $\Omega_\Lambda = 0$ even with a power-law primordial spectrum if the assumption of spatial flatness is dropped. This degeneracy is broken by the prior on H_0 [88].

effects on light propagation depend on how the voids are distributed, with some authors claiming it can partly mimic the effects of dark energy [94], and others finding a negligible change [95].

In this chapter we consider a "local void" model (sometimes referred to as a "Hubble bubble") as an explanation for dark energy. An underdense void expands faster than its surroundings, thus younger supernovae inside the void would be observed to be receding more rapidly than older supernovae outside the void. Under the assumption of homogeneity this would lead to the mistaken conclusion that the expansion rate of the universe is accelerating, although both the void and the global universe are actually decelerating. To prevent an excessive CMB dipole moment due to our peculiar velocity we must be located near the centre of the void. The local void scenario has been investigated by several authors using a variety of methods, including among others [96, 97, 98, 99, 100, 101, 102, 103].

Essentially, the local void model described using an LTB framework provides enough freedom through the choice of density profile of the void to allow a luminosity distance-redshift relation that cannot easily be distinguished from that of ΛCDM on the basis of current data [100]. As the Hubble parameter inside the void is higher inside the void than the global value in the background, it is also possible simultaneously to fit the location of the first acoustic peak in the CMB [98, 102]. It has however been argued that it is not possible for void models to provide a good fit to the *entire* CMB power spectrum [104, 105] unless the homogeneous background has significant curvature [106, 107]. These results are based on the assumption that the primordial power spectrum has the near-scale-invariant form that is consistent with the ΛCDM model.

In this work we argue that in fact a Gpc-sized void in an EdS background can simultaneously fit the SNe Ia data as well as the full CMB power spectrum, while also satisfying constraints from local Hubble measurements, primordial nucleosynthesis and the BAO data, if the primordial power spectrum is *not* assumed to be nearly scale-invariant. In Section 3.2 we summarize the general relativistic framework for LTB models and describe the characterization of the void. In Section 3.3 we discuss the form of the primordial power, and present a physical model with a primordial power spectrum that is not scale-free. In order to compare observables in the void model to existing cosmological data, some formalism needs to be developed. This is done in Section 3.4, and the statistical approach is discussed in Section 3.5. Finally Section 3.6 presents the main results.

3.2 LTB void models

3.2.1 The metric and solution

We model the void as an isotropic, radially inhomogeneous universe described by the LTB metric:

$$ds^2 = -c^2dt^2 + \frac{A'^2(r,t)}{1+K(r)}dr^2 + A^2(r,t)d\Omega^2, \tag{3.1}$$

where a prime denotes the partial derivative with respect to coordinate distance r, and the curvature $K(r)$ is a free function, bounded by $K < 1$. This reduces to the usual FRW metric in the limit where $A(r,t) \to a(t)r$ and $K(r) \to kr^2$.

We define two Hubble rates:

$$H_\perp \equiv \frac{\dot{A}(r,t)}{A(r,t)}, \;\; H_\parallel \equiv \frac{\dot{A}'(r,t)}{A'(r,t)}, \tag{3.2}$$

where an overdot denotes the partial derivative with respect to t. The analogue of the Friedmann equation is

$$H_\perp^2 = \frac{F(r)}{A^3(r,t)} + \frac{c^2K(r)}{A^2(r,t)}, \tag{3.3}$$

where $F(r) > 0$ is another free function which determines the local energy density through

$$8\pi G\rho(r,t) = \frac{F'(r)}{A^2(r,t)A'(r,t)}. \tag{3.4}$$

We define dimensionless density parameters $\Omega_M(r)$ and $\Omega_K(r)$ such that

$$F(r) = H_0^2(r)\Omega_M(r)A_0^3(r), \tag{3.5}$$

and

$$c^2K(r) = H_0^2(r)\Omega_K(r)A_0^2(r), \tag{3.6}$$

where $H_0(r)$ and $A_0(r)$ are the values of $H_\perp(r,t)$ and $A(r,t)$ respectively at the present time $t = t_0$. The Friedmann equation then becomes [108]:

$$H_\perp^2 = H_0^2\left[\Omega_M\left(\frac{A_0}{A}\right)^3 + \Omega_K\left(\frac{A_0}{A}\right)^2\right], \tag{3.7}$$

so $\Omega_M(r) + \Omega_K(r) = 1$. This equation can be integrated from the time of the Big Bang, $t_\text{B} = t_\text{B}(r)$, to yield the age of the universe at any given (r, t):

$$t - t_\text{B}(r) = \frac{1}{H_0(r)} \int_0^{A/A_0} \frac{dx}{\sqrt{\Omega_M(r)x^{-1} + \Omega_K(r)}}. \tag{3.8}$$

We thus have two functional degrees of freedom, in $\Omega_M(r)$ and $t_\text{B}(r)$, which can be chosen as desired. (The third function, $A_0(r)$, corresponds to a gauge mode and we choose to set $A_0(r) = r$.) A spatially varying t_B corresponds to a decaying mode [109], so for simplicity we set $t_\text{B} = 0$ everywhere, so that at the current time t_0:

$$H_0(r) = \begin{cases} \dfrac{-\sqrt{-\Omega_K} + \Omega_M \sin^{-1}\sqrt{-\frac{\Omega_K}{\Omega_M}}}{t_0\left(-\Omega_K\right)^{3/2}}, & \Omega_K < 0; \\ \dfrac{2}{3t_0}, & \Omega_K = 0; \\ \dfrac{\sqrt{\Omega_K} - \Omega_M \sinh^{-1}\sqrt{\frac{\Omega_K}{\Omega_M}}}{t_0\Omega_K^{3/2}}, & \Omega_K > 0. \end{cases} \tag{3.9}$$

The void model can then be specified by the choice of one free function, which we take to be $\Omega_M(r)$, and a constant $H \equiv H_0(0)$ which determines the *local* Hubble rate at the centre (and is equivalent to choosing t_0). Note that both at the centre of the void and far outside the void, the definition of $\Omega_M(r)$ reduces to the standard FRW density parameter Ω_m.

The solution to Eq. (3.7) for general r and t can be given in parametric form for the different values of $\Omega_K(r)$ (or $K(r)$) as follows [97]:

- for $\Omega_K(r) > 0$:

$$A = \frac{\Omega_M(r)A_0(r)}{2\Omega_K(r)}\left(\cosh\eta - 1\right), \tag{3.10a}$$

$$H_0 t = \frac{\Omega_M(r)}{2\Omega_K^{3/2}(r)}\left(\sinh\eta - \eta\right). \tag{3.10b}$$

- for $\Omega_K(r) = 0$:

$$A = \frac{1}{2}\left(18\Omega_M(r)\right)^{1/3}\left(H_0 t\right)^{2/3} A_0(r). \tag{3.11}$$

- for $\Omega_K(r) < 0$:

$$A = \frac{\Omega_M(r)A_0(r)}{2|\Omega_K(r)|}\left(1 - \cos u\right), \tag{3.12a}$$

$$H_0 t = \frac{\Omega_M(r)}{2|\Omega_K(r)|^{3/2}}\left(u - \sin u\right). \tag{3.12b}$$

Light travels to an observer at the centre of the void along null radial incoming geodesics described by [97]:

$$\frac{dt}{dz} = -\frac{1}{(1+z)\,H_{\|}(z)}, \tag{3.13a}$$

$$\frac{dr}{dz} = \frac{c\sqrt{1+K(r)}}{(1+z)\,A'(z)H_{\|}(z)}, \tag{3.13b}$$

where $H_{\|}(z) = H_{\|}\left(r(z), t(z)\right)$ *etc.* The angular diameter distance at redshift z is then given by

$$d_A(z) = A\left(r(z), t(z)\right), \tag{3.14}$$

and the luminosity distance by

$$d_L(z) = (1+z)^2\,A\left(r(z), t(z)\right). \tag{3.15}$$

All observable quantities along the light cone can be calculated from these equations.

3.2.2 Void profile

We can now choose the void profile by specifying $\Omega_M(r)$ and the local value of the Hubble rate by specifying H. Although the void profile may have any shape, we restrict ourselves to the simple Gaussian form:

$$\Omega_M(r) = \Omega_{\text{out}} - (\Omega_{\text{out}} - \Omega_{\text{in}}) \exp\left[-\left(\frac{r}{r_0}\right)^2\right], \tag{3.16}$$

where Ω_{in} and Ω_{out} correspond to the matter density parameter at the centre of the void and at infinity, respectively, and r_0 characterizes the width of the void. We wish to look only at voids that are asymptotically EdS, so we set $\Omega_{\text{out}} = 1$. Thus distances in the void model are completely specified by the three parameters Ω_{in}, r_0 and H. Throughout this chapter, we use $h_0 \equiv H_0(0)/(100\ \text{km s}^{-1}\ \text{Mpc}^{-1})$ to refer to the local Hubble rate at the centre of the void.

We wish to stress that in restricting ourselves to voids which have a Gaussian profile, we may be missing the model that fits the data best. In principle we could sample a wider class of profiles and choose the form that gives the best fit. However here we are concerned mainly with providing a counter-example of a void which can simultaneously fit both the SNe Ia magnitudes and the CMB spectrum, so we do not perform this search.

3.3 Primordial power spectra

The observed power spectrum of CMB anisotropies in any cosmological model is a convolution of three unknowns. The first two are the local physics at the time of recombination, which is dependent on the composition of the universe at that time, and the angular diameter distance to the last scattering surface (LSS), which depends on the geometry of the universe. Both of these are completely specified by the choice of the void model as described in Section 3.2.2, along with the further specification of the baryon fraction, Ω_b, the baryon-to-photon ratio $\eta \equiv n_b/n_\gamma$, and the choice $\Omega_\Lambda = 0$.

The third unknown is the shape of the primordial power spectrum of density perturbations. In the simplest single-field slow-roll models of inflation this is close to scale-invariant and featureless, as discussed in Chapter 1. Under the assumption that the primordial spectrum is described by a simple power law, the standard ΛCDM concordance cosmological model fits the observed angular power spectrum reasonably well [15]. However, there is no independent evidence for this form of the primordial power and as the observed anisotropies arise as a convolution of the *assumed* primordial spectrum with the transfer function of the *assumed* cosmological model, it is clear that we cannot determine one without making assumptions about the other. (In fact, there are indications that the primordial spectrum is *not* scale-free, even when a ΛCDM cosmology is assumed [110].)

In [106, 104, 105, 107] it is argued that, *assuming near-scale invariance of the primordial power*, void models cannot simultaneously provide an explanation for supernovae magnitudes and fit the observed CMB spectrum, unless the void is extremely deep [106] or embedded in a universe with large non-zero overall curvature [107]. However, in considering void models as an alternative to dark energy, we are in any case departing from the concordance cosmology, so there is no need to retain the assumption that the primordial power spectrum is scale-free. If this assumption is relaxed then even an EdS cosmology *without* dark energy can fit the CMB data [91, 89]. In the following we consider a particular alternative form of the primordial $\mathcal{P}_\mathcal{R}(k)$ that is not scale-free. It is of course possible to simply postulate an arbitrary alternative scale-dependent form for the primordial power without reference to a physical model.

The form of $\mathcal{P}_{\mathcal{R}}(k)$ that we describe below is in fact motivated by a particular physical theory of inflation, which is briefly described below, but this is not a necessary condition.

3.3.1 Bump model

Whereas the simplest toy models of inflation contain only a single scalar field which rolls slowly down its potential, physical models generically contain other fields, whose evolution is typically not slow-roll. These fields may couple to the inflaton and affect its evolution, thus breaking the scale-free nature of the primordial power, with important consequences.

An example of such a physical model is "multiple inflation" [111] in the framework of $N = 1$ supergravity, the locally realised version of supersymmetry (SUSY). This model includes flat-direction fields ψ which have gauge and Yukawa couplings to ordinary matter but are only gravitationally coupled to the inflaton. Such flat-directions have been classified and tabulated in the Minimal Supersymmetric Standard Model (MSSM) [112]. During inflation, the ψ fields undergo symmetry-breaking phase transitions, causing a sudden change in the effective mass of the inflaton (which is assumed to be a field in a hidden sector). A single flat-direction field would produce a "step" in the spectrum of the curvature perturbation [113], while if more than one flat-direction field is present, they can couple to the inflaton with opposite signs, and produce a "bump" feature in the power spectrum [89]. (A toy model that produces a similar "step" feature in the power spectrum has also been proposed [114] and has been studied with respect to fitting the WMAP 1-year [115] and 3-year [116] data, albeit in a different context to that considered here. Other signatures of multiple inflation, in particular the generation of associated non-Gaussianities, have also been studied [117].)

In this "bump" model, the potential for the inflaton ϕ and the flat-direction fields ψ_i is given by

$$\begin{aligned} V(\phi,\psi_1,\psi_2) = V_0 - \frac{1}{2}m^2 H_{\rm inf}^2\phi^2 &+ \frac{1}{2}\lambda_1 H_{\rm inf}^2\phi^2\psi_1^2 - \frac{1}{2}\mu_1^2 H_{\rm inf}^2\psi_1^2 + \gamma_1\psi_1^{n_1} \\ &+ \frac{1}{2}\lambda_2 H_{\rm inf}^2\phi^2\psi_2^2 - \frac{1}{2}\mu_2^2 H_{\rm inf}^2\psi_2^2 + \gamma_2\psi_2^{n_2}\,, \end{aligned} \tag{3.17}$$

where $mH_{\rm inf}$ and $\mu_i H_{\rm inf}$ are the masses of the ϕ and ψ_i fields respectively, $\lambda_i H_{\rm inf}^2$ is the coupling of the ψ_i field to the inflaton, γ_i is the coefficient of the non-renormalizable operator of order n_i which lifts the potential of the ψ_i field and $H_{\rm inf}$ is the Hubble scale during inflation (units of Planck mass $M_P = 1$ are used throughout). "Natural" values for the parameters of this model are $|\lambda_i| \sim 1$, $\mu_i^2 \sim 3$, and $\gamma_i \sim 1$, and it is assumed that the inflaton mass m is protected from SUSY-breaking corrections (the "η problem") so that sufficient e-folds of inflation occur. According to the list of flat-direction fields in [112], the ones with $n_i = 12$ and 16 will be the most relevant.

For a detailed discussion of the properties of the model, including the initial conditions, we refer the reader to the papers [111, 113, 89]. For our purposes it is enough to note that the two flat-direction fields ψ_i remain trapped at the origin until the phase transitions take place at times t_i. These times—or, equivalently, the scales at which the flat-direction fields introduce features in $\mathcal{P}_{\mathcal{R}}(k)$—can be chosen freely.

The primordial power spectrum produced by this 'bump' model can be calculated exactly by numerically solving the equation of motion of the ϕ field according to the potential in Eq. (3.17). This was performed in [89], where it was shown that a primordial power spectrum with such a "bump" would allow even an EdS universe with $\Lambda = 0$ to fit the WMAP 3-year data. This required a low global value of $h \simeq 0.44$, which is similar to that expected in a void model.

However, we do not wish to follow this prescription here for two reasons. Firstly, the full numerical solution is computationally expensive and slows an MCMC analysis considerably. Secondly, we are interested in the constraints on void models given the assumption of such a feature, and not in constraining the parameters of the theory of multiple inflation underlying Eq. (3.17). Eq. (3.17) demonstrates that such features in $\mathcal{P}_{\mathcal{R}}(k)$ are possible in an inflationary theory, but similar features with steps and oscillations may also be produced, e.g., in DBI inflation [118], and so there is no need to restrict the discussion to a single inflationary model. We need not even assume that the feature arises from inflation at all.

Therefore rather than reproduce the calculation of [89] exactly, we introduce a parameterization of $\mathcal{P}_{\mathcal{R}}(k)$ to capture the essential features of the reported best-fit bump model:

$$\mathcal{P}_{\mathcal{R}}(k) = \mathcal{P}_0 \left(1 - a \tanh(bx) + c \exp[-(bx)^2]\right) , \tag{3.18}$$

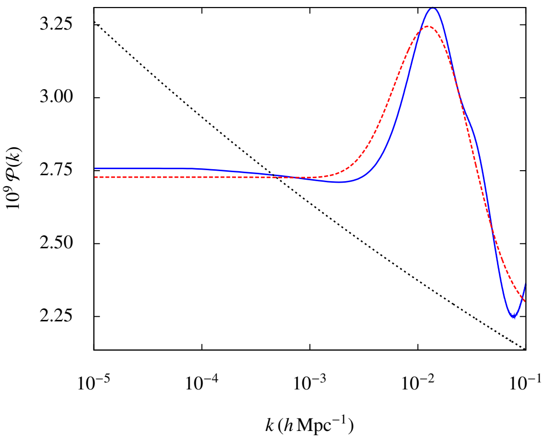

Figure 3.1: The solid (blue) curve shows the primordial power spectrum obtained from a full calculation of the multiple inflation model, with parameters as described in the text. The dashed (red) curve shows the simple parameterization (3.18), with parameter values $\mathcal{P}_0 = 2.48 \times 10^{-9}$, $a = 0.1$, $b = 2.0$, $c = 3.0$ and $k_0 = 0.015$. For comparison, the dotted (black) curve shows a standard power-law spectrum with slope $n_\mathrm{s} = 0.954$, amplitude $A_\mathrm{S} = 2.3 \times 10^{-9}$, which are the best-fit values for the ΛCDM model with no "running" to the CMB+SN$_\mathrm{CfA}$ datasets (Section 3.6).

where $x \equiv \log_{10}(k/k_0)$. For practical purposes, Eq. (3.18) is the form of the primordial power we are interested in, albeit motivated by an underlying physical model Eq. (3.17).

In Fig. 3.1 we plot $\mathcal{P}_\mathcal{R}(k)$ obtained from a full calculation in the "bump" model, with parameter values $n_1 = 16$, $n_2 = 12$, $\mu_1^2 = \mu_2^2 = 3$, $\gamma_1 = \gamma_2 = 1$, $\lambda_1 = -0.2$, $\lambda_2 = 1$ and $m^2 = 0.05$ (all set at their "natural" values, save for λ_1). H_inf, k_1 and k_2 are free parameters and are chosen so as to broadly reproduce the best-fit primordial spectrum from [89]. In the same figure we also plot the corresponding form of Eq. (3.18). We do not consider any tensor modes.

This form of $\mathcal{P}_\mathcal{R}(k)$ is thus determined by 5 free parameters: $\mathcal{P}_0$, a, b, c and k_0. This may appear to be a step backwards from the standard power-law form, which depends on only 3 (A_S, n_s and n_run — the amplitude, slope and "running" of the spectrum). However it is important to bear in mind that we are merely using an empirical parameterization for a $\mathcal{P}_\mathcal{R}(k)$ that is ultimately the result of some underlying theory, in which all of these parameters are in fact supposed to be determined by fundamental physics.

On the other hand, it can be argued that the standard parameterization of a power-law hides other parameters which are artificially set to zero through the overriding assumption that only a single scalar field is involved. In the absence of an accepted physical mechanism for inflation, we feel that all plausible alternatives ought to be considered, rather than judging purely on the basis of naive parameter-counting. It is with this rationale that we use this form of the primordial power in constructing this model.

3.4 Fitting the model to observations

In this section we will discuss how to fit the void model to the best available cosmological data: the WMAP 7-year release [15], SNe Ia data from the "Constitution" [12] and "Union2" compilations [13], the local Hubble rate measurement by the Hubble Key Project [8], constraints from big bang nucleosynthesis [119], and BAO measurements from SDSS Data Release 7 [40].

As cosmology in an LTB metric differs from the standard FRW approach, we need to develop some formalism to allow us to confront these datasets in a consistent manner. This is outlined in the following.

3.4.1 SNe Ia magnitudes

The distance modulus of a supernova is defined as the residual between its apparent and absolute magnitudes m and M, and is related to its luminosity distance as

$$\mu = m - M = 5\log_{10}\left[\frac{d_L}{\mathrm{Mpc}}\right] + 25\,, \tag{3.19}$$

so that, knowing μ for each object in a dataset, we can use Eq.(3.15) to compare a void model to the data.

We use two datasets: the Constitution compilation of 397 SNe in the redshift range $z = 0.015 - 1.55$ [12] and the Union2 compilation of 557 SNe in the range $z = 0.015 - 1.4$ [13]. The Constitution sample uses the SALT lightcurve fitter while Union2 uses the newer SALT2 fitter. Neither fitter, however, directly provides the value of μ for each object — for instance, the SALT fitter provides as output values of $m_B^{\max}$ (the rest-frame peak magnitude in the B-band), a time stretch factor s, and a colour parameter c for each supernova, from which the distance modulus is calculated as:

$$\mu_i = m_{B,i}^{\max} - M + \alpha \cdot (s_i - 1) - \beta \cdot c_i, \tag{3.20}$$

where M, α and β are empirical coefficients whose values are determined by marginalizing over the fit to a particular fiducial cosmology. The SALT2 fitter follows the same principle, though with a "stretch factor", x_1, that is analogous to but not the same as s.

Both the Constitution and Union2 compilations provide tabulated values of distance modulus μ and its estimated error σ_μ for a choice of coefficients M, α and β calculated for the best-fit flat ΛCDM model. In addition, σ_μ includes an important contribution from a systematic uncertainty whose value is *chosen* such that the reduced χ^2 value for the flat ΛCDM model is of order unity. Hence when comparing any alternative model to ΛCDM on the basis of their respective fits to the SNe Ia data, we ought to in principle perform the entire fitting analysis from scratch, choosing the values of M, α and β that produce the best fit for the particular model under consideration, and adjusting the systematic error inserted by hand in an equivalent manner. However, to do so we would require not only the actual lightcurve fitter outputs for each supernova (which the publicly downloadable data for the Constitution set does contain, but Union2 does not), but also the full covariance matrix of the lightcurve fits, which neither compilation provides.[2] Therefore, we are forced to adopt the common (but incorrect) procedure of fitting the void model by marginalizing over the unknown absolute magnitude M alone, while bearing in mind the caveat that this may bias our results and our best-fit void parameters in an unknown manner.

3.4.2 CMB power spectrum

It has been shown in several previous studies that void models are only very loosely constrained by the position of the first peak of the CMB [98, 102]. However, the WMAP satellite has measured the angular power spectrum C_ls over a wide range of multipoles, l, and it is preferable to use as much of this data as possible. This entails the calculation of the full predicted C_l spectrum for the void model. This has been done using various methods [106, 104, 105, 107], but always with the assumption of a nearly scale-invariant primordial spectrum.

In order to be able to calculate the predicted angular power spectrum for an LTB model using one of the publicly available Boltzmann codes (we use a version of CAMB

---

[2] This is provided in the SNLS 3-year data release [5]; however it contains only 231 objects, so we do not consider it here.

[29], modified to accept the primordial power spectra we consider), we use a version of the "effective EdS approach" [104, 105]. In brief, this consists of constructing an effective EdS model which has the same physics at recombination as the given void model, and the same angular diameter distance to the LSS. Then, given the same primordial power spectrum, the C_ls of the effective model and the void model will be the same, at intermediate and small angular scales. Thus the effective model can be used as a calculational tool to obtain the power spectrum for the void model under consideration. Note however that the effective EdS model will in general have a central temperature, $T_0^{\rm EdS}$, and a central Hubble rate, $H_0^{\rm EdS}$, that are *different* from the actual T_0 and H_0 that we observe today. This is because the physics has been matched at early times; at late times the models must then necessarily differ.

To generate the effective model, we adopt the following procedure. First we specify the void profile and local Hubble rate by choosing the parameters $\Omega_{\rm in}$, r_0 and H. Then, using Eqs. (3.13a) and (3.13b) we numerically integrate out from the centre of the void along the past light cone to obtain the coordinates (r_m, t_m) at an intermediate redshift z_m. As in [104, 105] we choose this redshift to be $z_m = 100$, where the spatial curvature of the void is negligible but the radiation density is still small and can justifiably be ignored in the calculations.

At these coordinates we now calculate the Hubble rate, $H_m = H_{\perp/\parallel}(r_m, t_m)$ and the angular diameter distance $A_m = A(r_m, t_m)$. Let us denote by $r^{\rm EdS}(t)$ the co-moving radial coordinate of a radial light ray in the EdS universe and by $z^{\rm EdS}$ the redshifts seen by an observer at $r^{\rm EdS} = 0$. To ensure the matching of the distances to the LSS, we impose the condition $A_m = a(r_m^{\rm EdS}) r_m^{\rm EdS}$ for EdS scale factor a. This provides us with the relation

$$A_m = \frac{2c}{\left(1 + z_m^{\rm EdS}\right) H_0^{\rm EdS}} \left(1 - \frac{1}{\sqrt{1 + z_m^{\rm EdS}}}\right), \tag{3.21}$$

where $z_m^{\rm EdS}$ is the redshift in the EdS universe at coordinates $\left(r_m^{\rm EdS}, t_m\right)$. (Note that our procedure implies $z_m^{\rm EdS} \neq z_m$ and thus differs slightly from the equivalent method outlined in [104, 105] where instead $z_0^{\rm EdS} \neq 0$.) To ensure the same physics at early times in the two models we also match the Hubble rates, $H_m = H^{\rm EdS}\left(z_m^{\rm EdS}\right)$, and this combined with Eq.(3.21) provides us with an expression for $z_m^{\rm EdS}$

$$z_m^{\rm EdS} = \frac{A_m^2 H_m^2 + 4cA_m H_m}{4c^2}. \tag{3.22}$$

Given the value of $z_m^{\rm EdS}$, we can then calculate the effective EdS mean temperature and Hubble rate via

$$T_0^{\rm EdS} = T_0 \left(\frac{1+z_m}{1+z_m^{\rm EdS}} \right), \tag{3.23a}$$

$$H_0^{\rm EdS} = \frac{H_m}{(1+z_m^{\rm EdS})^{3/2}}, \tag{3.23b}$$

where $T_0 = 2.726 \pm 0.001$ K is the CMB temperature observed today [44].

The parameters $T_0^{\rm EdS}, H_0^{\rm EdS}, \Omega_m = 1, \Omega_\Lambda = \Omega_k = 0$ and the baryon fraction Ω_b can then be fed into CAMB, together with the choice of primordial power spectrum, to find the spectrum of C_ls for the effective EdS model. This will be identical to the spectrum actually observed at the centre of the void, at large enough values of l.

At small values of l however, the void model will create an integrated Sachs-Wolfe (ISW) signal which is not captured by the effective EdS model. However, no rigorous calculation has yet been made of this expected ISW signal. In addition, the void may also in principle have a different reionization history than that of the effective EdS model. These two effects mean that at small l the C_l values calculated using the EdS model will differ, in both the TT and TE power spectra, from the actual spectrum due to the void.

In order to account for this, we choose to apply a cutoff at $l = 32$ in both the WMAP-7 TT and TE power spectra. The value $l = 32$ is chosen also to coincide with the switch-over point for the TT power spectrum at which the WMAP likelihood routine switches between the low-l, gibbs-sampling likelihood estimation technique and the master code for high-l (for the TE spectrum the switch-over point between pixel-based analysis and the master code is close by at $l = 24$). This also allows a more direct interpretation of the likelihood $\mathcal{L}$ in terms of a χ^2 value. We have checked that increasing the cutoff point does not materially affect our results.

Given that in [89] it is found that an EdS model can fit the WMAP data with a primordial power spectrum similar to the "bump" model we use here, we expect to find a good fit to the CMB with $h_0^{\rm EdS} \equiv H_0^{\rm EdS}/(100 \text{ km s}^{-1} \text{ Mpc}^{-1}) \simeq 0.44$ and $T_0^{\rm EdS} \simeq T_0$. This should be contrasted with the values of $h_0^{\rm EdS} \simeq 0.51$ and $T_0^{\rm EdS} \simeq 3.4\,\text{K}$ required with a power-law primordial spectrum [104, 105].

3.4.3 Local Hubble rates

It has been claimed in previous studies [104, 105, 107] that the *local* Hubble rate of void models that fit the CMB and SNe Ia data simultaneously must be very low (as low as $h_0 \simeq 0.45$ in [104, 105]) and that this argues against void models, since the measured local values (at $z < 0.1$) are significantly higher (see [10] for a review). This can provide an important discriminant against void models. Therefore in performing the MCMC analysis, we also fit the local Hubble rate for the void model $h_0^{\mathrm{LTB}} \equiv H/100$ to the Hubble Key Project (HKP) value $h_0 = 0.72 \pm 0.08$ [8]. We note that there is some variation in the value of h_0 obtained by different groups, ranging from $h_0 = 0.623 \pm 0.06$ [120] to $h_0 = 0.742 \pm 0.036$ [121], and the HKP value lies in between these two.

The SNe Ia compilations, WMAP-7 observations and the HKP value for h_0 form the primary datasets that we use to constrain the void model. We also discuss the fit to some other cosmological data below.

3.4.4 Big bang nucleosynthesis

Our theoretical understanding of the physics of the epoch of big bang nucleosynthesis allows us to use observations of the abundances of various elements to constrain the baryon-to-photon ratio, $\eta = n_b/n_\gamma$, at that time. Although η is constant with time in FRW spacetime, this is not the case in LTB spacetime. However, assuming that η is spatially constant in the LTB model (which need not necessarily be true, see [122, 123]), we can use the effective EdS model model to calculate η for the void model. Since both models share the same early universe physics by construction, we have [105]

$$\eta_{10} \equiv 10^{10}\eta = 273.9 \left(\frac{T_0}{T_0^{\mathrm{EdS}}}\right)^3 \omega_b \,, \tag{3.24}$$

where $\omega_b \equiv \Omega_b \left(h_0^{\mathrm{EdS}}\right)^2$. The inferred primordial abundance of deuterium, together with that of helium, provides the constraint $5.1 \leq \eta_{10} \leq 6.5$ at 95% C.L. [119].

3.4.5 BAO scale

The BAO data provided by the SDSS collaboration [40] are essentially measurements of a feature in the correlation function of the observed galaxy distribution which is related to the physical sound horizon at the CMB scale, evolved down to the redshift at which the measurement is made. The full theory for how these perturbations should evolve in LTB spacetimes is unknown and a difficult problem, as the background curvature enters the Bardeen equation for the gravitational potential. As this may vary significantly over scales of ~ 150 Mpc, it can potentially add a significant and as yet unknown distortion to BAO scales [124].

Nevertheless, some efforts have been made to compare void models to BAO data under the assumption that the evolution of perturbations does *not* depend on scale (as is the case in FRW spacetime) [104, 105, 107]. Under this assumption the difference between LTB and FRW spacetimes is simply that as $H_{\|} \neq H_{\perp}$ for $r > 0$, physical length scales in an LTB spacetime evolve differently in radial and transverse directions at late times, whereas in FRW models the evolution is isotropic. It is not clear that this is necessarily a valid assumption to make. However, in order to provide a comparison with the previous studies we follow the same prescription as in [105], while bearing in mind that a better calculation may lead to significant changes in the results.

The first step is to construct another effective EdS model, referred to as the "BAO model", which shares the same early universe physics as the void model under consideration, but unlike in the previous case, need not match the angular diameter distance to the LSS. We choose in this case to match the central temperatures, $T_0^{\mathrm{BAO}} = T_0$, which then provides us with the relation:

$$H_0^{\mathrm{BAO}} = \frac{H_m}{(1+z_m)^{3/2}}, \tag{3.25}$$

as unlike in the previous Section, we will now have $z_m^{\mathrm{BAO}} = z_m$ (where, as before, we choose $z_m = 100$). We also match the values of Ω_b and $\Omega_m = 1$ at this redshift to ensure the same early universe physics.

As the BAO model and the LTB void model have, by definition, the same physics at coordinates $(r(z_m), t(z_m)) \equiv (r_m, t_m)$, and since at this redshift the LTB spacetime is sufficiently close to FRW, we may conclude that the physical sound horizon in the void model at time t_m is essentially isotropic and homogeneous, and equal to the sound horizon $s_{\mathrm{p}}(z_m)$ in the BAO model, which may be calculated according to [125]:

$$s_{\mathrm{p}}(z) = \frac{44.5 \ln\left[9.83/\left(\Omega_m h_0^2\right)\right]}{(1+z)\sqrt{1+10\left(\Omega_b h_0^2\right)^{3/4}}}\,\mathrm{Mpc}. \tag{3.26}$$

In order to evaluate the BAO scales observed at redshift z by an observer at the centre of the void, we need to evolve this physical sound horizon scale down to coordinates $(r(z), t(z))$ in the LTB background. Thus the radial and transverse physical BAO scales at redshift z will be, respectively,

$$l_{\parallel}^{\mathrm{BAO}}(z) = s_{\mathrm{p}}(z_m) \frac{A'(r(z), t(z))}{A'(r(z), t_m)}, \tag{3.27}$$

$$l_{\perp}^{\mathrm{BAO}}(z) = s_{\mathrm{p}}(z_m) \frac{A(r(z), t(z))}{A(r(z), t_m)}, \tag{3.28}$$

and these can in turn be rewritten in terms of the corresponding redshift and angular intervals:

$$\Delta z(z) = (1+z)\, l_{\parallel}^{\mathrm{BAO}}(z)\, H_{\parallel}(r(z), t(z)), \tag{3.29}$$

$$\Delta\theta(z) = \frac{l_{\perp}^{\mathrm{BAO}}}{A(r(z), t(z))}. \tag{3.30}$$

It is these values of $\Delta z(z)$ and $\Delta\theta(z)$ that are directly measurable. As mentioned in [105], the redshift scale $\Delta z(z)$ is expected to be a stronger discriminator of void models than $\Delta\theta(z)$. However, while estimates of the radial BAO scale have been made [126], the statistical significance of these claims has been questioned [127]. Given this ambiguity, we choose not to use this data; instead we use constraints on the ratio,

$$\Theta(z) \equiv \frac{s_{\mathrm{p}}(z_{\mathrm{rec}})}{D_V(z)}, \tag{3.31}$$

at redshifts $z = 0.2$ and $z = 0.35$ from [40], where z_{rec} is the redshift at recombination and $D_V(z)$ is an isotropized distance measure. This is given in a FRW model by [16]

$$D_V^{\mathrm{FRW}}(z) \equiv \left[\frac{z d_A^2(z)}{H(z)}\right]^{1/3}, \tag{3.32}$$

where $d_A(z)$ and $H(z)$ represent, respectively, the angular diameter distance and the Hubble rate in the FRW model.

As the BAO measurements are quoted in terms of the ratio $\Theta(z)$, it is necessary to convert them into a form that can be related to $\Delta z(z)$ and $\Delta\theta(z)$. It can be shown [107] that

$$Q(z) \equiv \left(\Delta\theta^2 \Delta z\right)^{1/3} = z^{1/3}\Theta(z), \tag{3.33}$$

and this is the measure we use to compare the void model to the data.

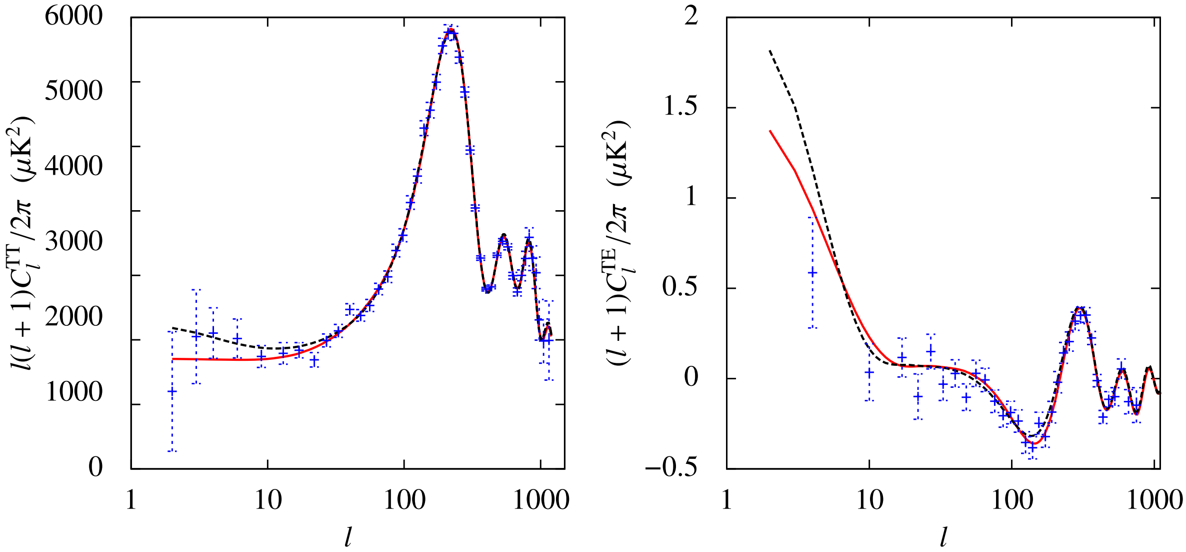

Figure 3.2: The TT and TE power spectra of the models that best fit the CMB and Constitution SNe Ia data. The solid (red) curve is for the void model, calculated using the effective EdS approach described in Section 3.4.2, and the dashed (black) curve is for ΛCDM. Binned WMAP-7 data is also shown.

3.5 Method

We perform a likelihood analysis using COSMOMC [128] to generate Markov-Chain-Monte-Carlo (MCMC) chains to estimate confidence limits on the parameters in fitting the model to the data. For each void model, specified by Ω_{in}, r_0 and H, we first calculate the effective parameters H_0^{EdS} and T_0^{EdS} as described in Section 3.4.2. We fix the optical depth to the LSS to the WMAP-7 value, although this choice is immaterial as our cutoff at $l = 32$ means that the data do not constrain reionization in any case. Therefore the full set of parameters for the void model is: local matter density Ω_{in}, void radius r_0, local Hubble value H, baryon fraction Ω_b, and the power spectrum parameters $\mathcal{P}_0$, a, b, c and k_0 (the cold dark matter density Ω_{c} is set equal to $1 - \Omega_b$). These are fed as inputs to CAMB [29] to generate the CMB spectrum for the model.

Once the output C_l values have been obtained from CAMB, we can fit the model to the WMAP-7 data [15], the SNe Ia data (the Constitution [12] and Union2 [13] datasets are fitted separately), the HKP value for the local Hubble rate [8], the BAO data [40] and the BBN constraint [119].

In order to compare the goodness of fit, we perform exactly the same fitting

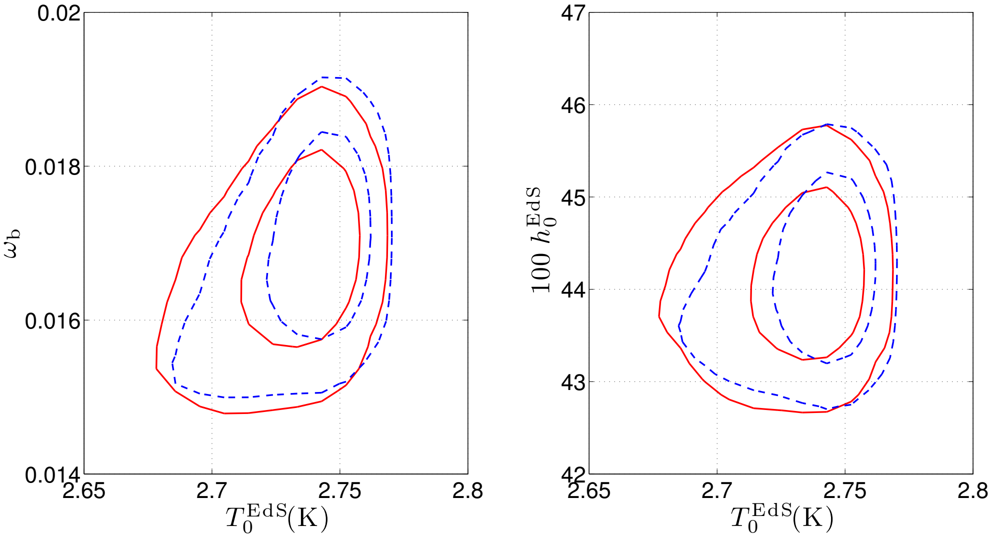

Figure 3.3: Constraints on the effective EdS parameters for the void model with a bump, from CMB and SNe Ia data. The solid (red) contours show the 1 and 2σ likelihood confidence intervals for the Constitution SNe Ia data set and the dashed (blue) contours are for Union2.

procedure for a vanilla ΛCDM model whose parameters are: baryon density $\Omega_b h_0^2$, cold dark matter density $\Omega_c h_0^2$, local Hubble rate $100h_0$,[3] and the power spectrum amplitude, A_S, and slope, n_s. We characterize the best-fit likelihood of the void model by the value $\Delta\chi^2 = -2\ln(\mathcal{L}_{\rm void}/\mathcal{L}_{\Lambda\rm CDM})$, which means that negative values of $\Delta\chi^2$ favour the void. In Fig. 3.2 the TT and TE power spectra are shown for two sample best-fit models. The quantitative results of the analysis are discussed in the next section.

3.6 Results

In this section we shall use $\rm SN_{CfA}$ to denote the Constitution data set and $\rm SN_{U2}$ to denote Union2; CMB denotes the WMAP-7 data and HKP the Hubble Key Project value for the Hubble rate.

In Figure 3.3 we plot the 2D likelihoods for the effective EdS model parameters $T_0^{\rm EdS}$, $H_0^{\rm EdS}$ and Ω_b obtained from the MCMC chains for both CMB+$\rm SN_{CfA}$ and CMB+$\rm SN_{U2}$ data. Clearly, $T_0^{\rm EdS} \simeq T_0$ and $h_0^{\rm EdS} \simeq 0.45$, which is exactly as we

---

[3] The COSMOMC code actually uses the ratio of the sound horizon to the angular diameter distance as the input parameter, but this is equivalent to using h_0.

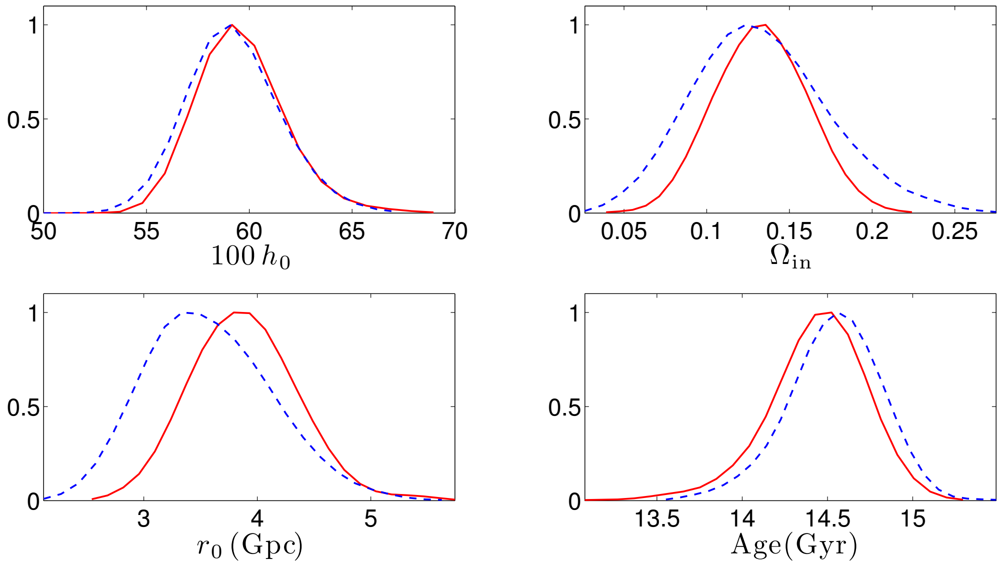

Figure 3.4: Marginalized 1D likelihoods for the void model with a bump, for the fit to CMB and SNe Ia data. The solid (red) curves are for the Constitution SNe Ia data set and the dashed (blue) curves are for Union2.

expect, given the results of [89]. Figure 3.4 shows the marginalized likelihoods of the void parameters and the resultant age of the universe, for the same datasets.

For both CMB+SN$_{\rm CfA}$ and CMB+SN$_{\rm U2}$ we obtain a local Hubble value $h_0 \sim 0.60 \pm 0.02$, which, although slightly low, is still consistent with the HKP value of $h_0 = 0.72 \pm 0.08$ to within 2σ. It is also much higher than the value of h_0 obtained in previous studies [104, 105]. This is because in order to fit the CMB with a power-law primordial spectrum a value of $T_0^{\rm EdS} \simeq 3.4$ K is required and in order to generate this, the void must be surrounded by a large over-dense shell. The presence of this shell means the asymptotic value $H_0(r \gg r_0) \sim H_0^{\rm EdS} \sim H_0(0)$ and hence in order to fit the CMB, $H_0(0)$ must be small. In contrast, our void profile only has a very small over-dense shell, and as shown in Figure 3.3, $T_0^{\rm EdS} \simeq T_0$. This allows us to generate the difference between $h_0 \simeq 0.60$ and $h_0^{\rm EdS} \simeq 0.45$.

Our void model also does not suffer from the "old age" problem referred to in [105]: we find the age of the void universe to be 14.4 ± 0.3 Gyr for the CMB+SN$_{\rm CfA}$ chains and 14.5 ± 0.3 Gyr for CMB+SN$_{\rm U2}$, which are significantly less than the values of $\sim$ 18.8 Gyr quoted in [105]. This is unsurprising as the age t_0 is related to the value of h_0 by Eq. 3.9, and reasonable values for one ensure reasonable values for the

Datasets	# d.o.f.	$-2\ln(\mathcal{L}_{\Lambda\mathrm{CDM}})$	$\Delta\chi^2$
CMB	1936	5785.0	+0.9
CMB + $\mathrm{SN}_{\mathrm{CfA}}$	2333	6250.4	−3.3
CMB + $\mathrm{SN}_{\mathrm{U2}}$	2493	6315.7	+1.4
CMB + $\mathrm{SN}_{\mathrm{CfA}}$ + HKP	2334	6250.5	−0.6
CMB + $\mathrm{SN}_{\mathrm{U2}}$ + HKP	2494	6315.9	+3.4

Table 3.1: Best-fit likelihood values for ΛCDM and the relative $\Delta\chi^2$ values for the void model with a spectral "bump" for different choices of the fundamental constraining data sets. Here $\mathrm{SN}_{\mathrm{CfA}}$ refers to the Constitution sample and $\mathrm{SN}_{\mathrm{U2}}$ to Union2. CMB refers to WMAP-7 data and HKP to the Hubble Key Project value for h_0. Constitution data favour the void slightly and Union2 data favour ΛCDM slightly, but neither difference is significant given the number of degrees of freedom.

other.

To enable a comparison of the quality of the fits, we show in Table 3.1 the $\Delta\chi^2$ values for the best-fit model with a void and a spectral "bump", relative to the standard ΛCDM model, for different combinations of the primary constraining data sets. It is clear that the $\Delta\chi^2$ values in all cases are small given the number of degrees of freedom involved, which shows that the void model is perfectly compatible with these datasets. A more quantitative comparison cannot be made because as noted in Section 3.4.1 the SNe Ia data are adjusted to fit a fiducial ΛCDM model and the error bars are tuned to give a χ^2 per degree of freedom of order unity for this model.

It should be noted that the addition of the HKP constraint increases the $\Delta\chi^2$ value by ~ 2 for Union2 and ~ 2.7 for Constitution, which reflects the fact that CMB and SNe Ia data favour a value of h_0 that is lower than, but consistent with, the HKP value. Even with this increase however, the $\Delta\chi^2$ values are still small (and in fact the void provides a marginally better fit to the CMB+$\mathrm{SN}_{\mathrm{CfA}}$ data than does ΛCDM). It is possible that a thorough exploration of possible void profiles or use of unbiased SNe Ia data will make this fit better, but as this does not detract from the main results presented here, we do not investigate it further.

An interesting constraint comes from Big Bang nucleosynthesis. From our MCMC chains for CMB+$\mathrm{SN}_{\mathrm{CfA}}$ and CMB+$\mathrm{SN}_{\mathrm{U2}}$ we find in both cases $\eta_{10} = 4.6 \pm 0.2$, which is $\sim 2\sigma$ below the best-fit value. This arises primarily because ω_b is on the low side, as seen in Figure 3.3. When the BBN constraint is added to the MCMC chains, we find an increase in $\Delta\chi^2$ of ~ 2.5 relative to the values for CMB+$\mathrm{SN}_{\mathrm{CfA}}$+HKP and an increase of ~ 3.1 for CMB+SN$_{\rm U2}$+HKP. It should be noted however that the inferred primordial abundance of lithium indicates a significantly lower value of η than does deuterium [119], hence a better understanding of the chemical evolution of these fragile elements is required before the significance of the marginal discrepancy above can be assessed.

Finally, we find that adding the BAO data results in an increase in $\Delta\chi^2$ of 4.9 relative to the value for CMB+SN$_{\rm U2}$+HKP+BBN and 5.5 relative to the value for CMB+SN$_{\rm CfA}$+HKP+BBN, for an additional two degrees of freedom. The main reason for this increase is that the BAO data require a shallower void with $\Omega_{\rm in} \sim 0.17$, whereas the CMB+SN data favour a slightly deeper void with $\Omega_{\rm in} \sim 0.13$. Most of the increase in χ^2 arising from including the BAO data comes from the poorer fit to SNe Ia data that results. Given the uncertainty in interpreting $\chi^2_{\rm SN}$ that we have already mentioned, as well as the assumptions that have gone into our analysis in Section 3.4.5, we cannot draw any definite conclusions from these results. However it would appear that the BAO data are certainly not *in*consistent with a void model.

3.7 Discussion

We have presented here a local void model which fits SNe Ia and CMB data, local H_0 values, nucleosynthesis constraints and BAO *without* requiring dark energy and thus provides a counterexample to the claim that dark energy is necessary to fit these observations.

One objection that has subsequently been raised to such a model is the relatively low value of H_0 at the centre of the void, which is several standard deviations away from a more recent determination $H_0 = 73.8 \pm 2.4$ km s^{-1} Mpc^{-1} [9]. This has been claimed to rule out void models as an alternative to dark energy. As can be seen from Fig. 3.4, the mean value of h_0 obtained for the void model considered here is ~ 0.60, which is indeed low.[4] It is worth considering why this is so.

We find that if only the WMAP data are used in the fitting process, h_0 is effectively unconstrained: the data do not significantly distinguish between values in the range $0.45 < h_0 < 0.8$. This is understandable, because WMAP data provide a constraint

---

[4] However, measurements of h that do not rely on the distance ladder approach of [9] give somewhat lower values, e.g. $h = 0.67 \pm 0.03$ from the 6dF galaxy survey [129].

on h_0^{EdS} but not on Ω_{in} or r_0, and hence not on h_0. It is only when the SNe Ia are added that Ω_{in} and r_0 are constrained and we obtain the marginalized likelihood shown in Fig. 3.4. We have already argued that the SNe Ia datasets used here do not allow for a truly fair comparison, because the reported error bars include an arbitrary systematic uncertainty chosen specifically for ΛCDM, and correct marginalization over the nuisance parameters α and β is not possible. If the fit to the supernovae data alone were to be performed in some truly model-independent way, although it is unlikely that χ^2_{SNe} would be reduced significantly (the χ^2 per degree of freedom is already significantly less than 1 for Union2), we believe it is possible that the best-fit values of Ω_{in} and r_0 could change enough to allow higher h_0 values when fitting to WMAP and SNe Ia data simultaneously.[5] This is a loophole in the argument against void models which should be investigated further.

Ultimately however, the best way of ruling out void models may come from non-geometrical observations. One such observable is the measurement of the time-variation of $H(z)$ using redshift drift [130]. This can provide direct confirmation of acceleration, and such observations may be possible in the near future [131].

It is also possible in principle to constrain a violation of the Copernican principle via the y-distortion of the CMB spectrum that is produced by the Compton scattering of photons by reionized gas in regions of the void that see a highly anisotropic LSS [132]. In the single-scattering and linear approximations and under the assumption that the dipole anisotropy dominates the distortion, this can be written as [105]

$$y = \frac{7}{10} \int_0^{z_{\mathrm{re}}} dz \frac{d\tau}{dz} \beta\left(z\right)^2 \,, \tag{3.34}$$

where τ is the optical depth, $\beta\left(z\right)$ is the dipole temperature anisotropy in the CMB observed at redshift z, and the integral is taken up to the redshift of reionization z_{re}. The FIRAS instrument on COBE provides an upper bound $y < 1.5 \times 10^{-5}$ (at 2σ) [43]. While this can rule out some void models, it was shown in [105] that a suitable choice of the void profile can evade this constraint.

---

[5] When fitting SNe Ia data to flat ΛCDM, it is argued that the change in the best-fit parameters obtained by using a truly model-independent fitting method is small. However flat ΛCDM models form a one-parameter family of d_L-z curves, whereas LTB models have at least two parameters, as in the case of the model studied here—and, more generally, an entire functional degree of freedom—so it is not immediately clear that this argument should hold.

Epilogue

Another direct test of a local void is via the kinetic Sunyaev-Zel'dovich (kSZ) effect [133]. This has been applied to cluster kSZ observations (e.g., [103, 134]). A recent paper [135] finds that cluster kSZ observations place severe constraints on simple void models with a homogeneous bang time. Allowing the bang-time to vary significantly increases the freedom of the models, but it is found that it is not possible simultaneously to obtain a small kSZ effect and a large local Hubble value, putting void models in conflict with observation. However, this analysis does not allow for the primordial power spectrum to differ from near-scale-invariance, and allowing for this additional freedom, or for separately inhomogeneous matter and radiation fluids [123], may weaken these constraints.

It is also possible to use upper bounds on the kSZ effect from the Atacama Cosmology Telescope [136] to constrain void models. In [137] this was done for simplistic void models treated as two homogeneous regions joined by a singular mass shell and the authors concluded that void models were ruled out. However, more recently a careful analysis using LTB voids [138] somewhat weakened these constraints and highlighted the various assumptions that must be made in obtaining them. Technically it is possible that by changing these assumptions it may still be possible to evade the kSZ constraints, but it appears unlikely. In fact the main source of uncertainty would appear to come from the use of the biased SNe Ia fitting method, which is common to all of the papers above.

Ultimately, one does not expect the perfectly spherically symmetric LTB void model to provide an accurate description of the real universe, for which a more sophisticated treatment of inhomogeneities must surely be required. It seems inevitable that in the future these simple models will eventually be ruled out once all loopholes are closed. Nevertheless, as we have discussed in this chapter, they demonstrate the difficulty in drawing any unambiguous conclusions from the cosmological data free from the influence of prior assumptions of homogeneity or the nature of inflation implicit in ΛCDM. In the next Chapter we examine whether these assumptions themselves are consistent with the observed structure of the universe.

Chapter 4

The ISW imprint of cosmic superstructures

4.1 Introduction

Throughout this thesis we have highlighted the cosmological observations for which the ΛCDM model provides a consistent explanation, including CMB anisotropies, SNe Ia, local Hubble value, galaxy clustering, baryon acoustic oscillations, etc. However, as we stressed in Chapter 3, this evidence is mostly *geometrical*, as it is based on interpreting measurements of distances—made using 'standard rulers' (the sound horizon at last scattering) and 'standard candles' (SNe Ia)—as due to accelerated expansion. We showed in Chapter 3 how the same data can be equally well fitted without dark energy in a local void model if other assumptions such as a power-law spectrum for the primordial density perturbations are relaxed. It is therefore important to check if there are other observational tests of ΛCDM, which could provide evidence for the *dynamical* effects of dark energy.

An example of such an observation is the late-time integrated Sachs-Wolfe (ISW) effect [46]. The decay of gravitational potentials after dark energy begins to dominate (at redshift $z \lesssim 1$) should lead to secondary CMB anisotropies as the CMB photons traverse regions of over- or under-density. If the universe is spatially flat, then detection of the linear ISW effect through cross-correlation of the CMB with large-scale structure would provide direct evidence of dark energy's effect on the growth of structure [139] to complement the geometrical evidence for the ΛCDM model.[1]

---

[1] The ISW effect should also boost low multipoles in the CMB angular power spectrum, whereas these are in fact anomalously low on the observed sky. However given the large cosmic variance on

Unfortunately the ISW signal in full-sky CMB-galaxy cross-correlation studies is hard to detect at high significance as it requires z measurements for over 10 million galaxies [140], and such datasets are not yet available. However, several authors (e.g., [141, 142, 143]) have examined smaller source catalogues and reported marginal detections with $< 3\sigma$ significance. On the other hand, other authors [144] were unable to reject the null hypothesis (no ISW effect) and [145] even found a slight *anti*-correlation thus rejecting ΛCDM at $2 - 3\sigma$ significance. Some groups have combined several different data sets and report up to 4σ detection of ΛCDM [146], but this interpretation is complicated by the difficulty of estimating errors for correlated galaxy data sets, and it has been claimed that errors have been significantly underestimated [147].

Much of the uncertainty in full-sky studies arises from the difficulty in reconstructing the underlying density field from galaxy survey data, given Poisson noise in the distribution of galaxies. A different approach to this problem is followed in [148, 149], who study the Sloan Digital Sky Survey (SDSS) Data Release 6 (DR6) luminous red galaxies (LRGs). They use 3D galaxy information rather than the projected 2D density, and select only the most extreme density perturbations, which are unambiguously identified despite Poisson noise. Along the lines of sight corresponding to these 'superstructures' they report a 4.4σ detection of the ISW effect. Aside from being the most significant detection of the ISW signal to date, this approach naturally provides information about the sizes and distribution of extreme structures in the universe, so can be used to check the consistency of the standard ΛCDM model of structure formation, and can constrain, e.g., primordial non-Gaussianity.

However, the magnitude of the temperature signal reported in 2008 by Granett et al. [148] (hereafter G08a) is surprisingly large and has been argued [90, 150] to be quite inconsistent with ΛCDM. The original authors have responded by noting that the assumed profile of the superstructures has a large effect on the theoretical prediction for their temperature signal [151], such that with a different assumption than the 'compensated top-hat' profile adopted by [90] and [150], and when using a template fit approach, the discrepancy is only at the 2σ level.

---

these scales, the discrepancy with ΛCDM is not too significant.

In this work we aim to clarify this important issue. We calculate the expected temperature signal from these superstructures making no *a priori* assumptions about their nature except that they arose in a ΛCDM cosmology with gaussian primordial density perturbations. We find that while [151] are right in that the expected signal does depend somewhat on the assumed density profile, its value calculated using an exact treatment of (initially) gaussian perturbations [152] is still discrepant at $> 3\sigma$ with the observations reported by G08a.

In Section 4.2 we briefly review the ISW effect and in Section 4.3 we calculate the expected temperature signal of superstructures in the standard ΛCDM model. Section 4.4.1 describes the key features of the observation of G08a which must be accounted for in making a comparison with the theoretical calculation. In Section 4.4.2 we show that, even if G08a had selectively picked out the regions in the survey region with the biggest ISW signal, there is still a significant discrepancy. Finally in Section 4.5 we discuss possible reasons for this discrepancy and future observational tests.

4.2 The ISW effect of structures

The ISW effect is a late-time secondary CMB anisotropy caused when the CMB photons traverse a region with a time-varying gravitational potential Φ. If the potential decays with time then for a photon passing through an overdense region the energy it gains while falling into the potential well is not cancelled by the energy it later loses in climbing out of the evolved, shallower, potential well. Overdense regions therefore appear as hot spots in the CMB. Conversely underdense regions, or voids, will appear as cold spots as the photon loses more energy climbing the potential hill than it subsequently gains while descending.

The necessary ingredient for the existence of a late-time ISW effect is that the gravitational potential Φ should decay. In a universe with $\Omega_m = 1$ and no dark energy, density perturbations δ grow at exactly the same rate as the scale factor of the universe a ($\equiv (1+z)^{-1}$), so at the linear level there is no evolution of $\Phi \propto -\delta/a$. However, in a ΛCDM universe a grows at a faster rate than (linear) density perturbations, so perturbations in the gravitational potential decay with time, leading to ISW temperature fluctuations as described above.

The temperature fluctuation $\Delta T(\hat{n})$ induced along a direction $\hat{n}$ is [46]

$$\Delta T(\hat{n}) = \frac{2}{c^3}\bar{T}_0 \int_0^{r_{\mathrm{L}}} \dot{\Phi}(r, z, \hat{n})\ a\ dr\ , \tag{4.1}$$

where $\bar{T}_0$ is the mean CMB temperature, r_{L} is the radial comoving distance to the last scattering surface (LSS), $\dot{\Phi}(r, z, \hat{n})$ is the time derivative of the gravitational potential along the photon geodesic and c is the speed of light.

The Poisson equation relates Φ to the density contrast $\delta \equiv (\rho - \bar{\rho})/\bar{\rho}$ (where $\bar{\rho}(t)$ is the mean density) through:

$$\nabla^2\Phi(\mathbf{x}, t) = 4\pi G\bar{\rho}(t)a^2\delta(\mathbf{x}, t). \tag{4.2}$$

This can be written in Fourier space as

$$\Phi(\mathbf{k}, t) = -\frac{3}{2}\left(\frac{H_0}{k}\right)^2 \Omega_m \frac{\delta(\mathbf{k}, t)}{a}\ . \tag{4.3}$$

Taking the time derivative of this equation yields

$$\dot{\Phi}(\mathbf{k}, t) = \frac{3}{2}\left(\frac{H_0}{k}\right)^2 \Omega_m \left[\frac{\dot{a}}{a^2}\delta(\mathbf{k}, t) - \frac{\dot{\delta}(\mathbf{k}, t)}{a}\right]. \tag{4.4}$$

At this point we make the assumption that linear theory holds on the large scales of interest, and hence that the perturbations grow as $\delta(\mathbf{k}, t) = D(t)\delta(\mathbf{k}, z = 0)$, where $D(t)$ is the linear growth factor, Eq. (1.45). This is an approximate treatment that only captures linear effects, but it has been shown that non-linear corrections, although important relative to the linear effects at early times, represent only a 10% correction at the low redshifts we are interested in [153].[2] In this approximation,

$$\dot{\Phi}(\mathbf{k}, z) = \frac{3}{2}\left(\frac{H_0}{k}\right)^2 \Omega_m \frac{H(z)}{a}\left(1 - \beta(z)\right)\delta(\mathbf{k}, z)\ , \tag{4.5}$$

where $\beta(z) \equiv d\ln D/d\ln a$ is the linear growth rate. Hence under the linear approximation, the time evolution of $\dot{\Phi}$ is described by the ISW linear growth factor, $G(z) = H(z)\left(1 - \beta(z)\right)D(z)/a$. It is only for an $\Omega_m = 1$ universe with no dark energy that $\beta(z) = 1$ for all z and so there is no ISW effect in the linear approximation.

---

[2] Both the linear and non-linear effects grow with time, however at late times and large scales ($\sim$ 100 Mpc/h), the linear effect dominates while at early times (when $\Omega_\Lambda \simeq 0$), both effects are smaller but the non-linear effect dominates.

Given the density profile δ of any isolated superstructure, Eqs. (4.1) and (4.5) can be used to calculate the temperature fluctuation it induces in the CMB. Assuming spherical symmetry of the density profile, Eq. (4.5) becomes

$$\dot{\Phi}(r, z) = \frac{3}{2}\Omega_m H_0^2 G(z) F(r) \tag{4.6}$$

in real-space, where the function $F(r)$ is defined as

$$F(r) = \int_0^r \frac{r'^2}{r}\delta(r')\, dr' + \int_r^\infty r'\delta(r')\, dr'\,, \tag{4.7}$$

with $\delta(r')$ being evaluated at redshift $z = 0$.[3] Thus $F(r)$ contains all information about the structure in question, while the assumed cosmology enters through the prefactor and the ISW growth term $G(z)$ in Eq. (4.6).

4.3 Expected signal from superstructures in the ΛCDM model

Here we use the term 'superstructures' to refer to density perturbations extending over $\gtrsim 100\ h^{-1}$Mpc. These should not be thought of as non-linear collapsed structures in the usual sense, rather as smooth hills and valleys in the density distribution. Collapsed structures form only where the density perturbation $\delta(\mathbf{r})$ exceeds unity, which happens on much smaller scales than those of interest here.

The question of what size of ISW effect to expect from superstructures in the ΛCDM model and whether this is consistent with the signal observed in G08a has been addressed by several authors, but with differing conclusions. G08a stated that the most extreme structures in the $(500\ h^{-1}\text{Mpc})^3$ box of the Millennium N-body simulation [154], when placed at $z = 0$, would produce a signal of $\Delta T \sim 4.2\ \mu$K. However, when Hunt and Sarkar calculated the signal distribution for supervoids with the densities and sizes reported by G08a, they obtained only $\langle \Delta T \rangle = -0.42\ \mu$K [90]. They assumed a compensated top-hat profile for the gravitational potential motivated by the asymptotic final state of a void [155]. A similar profile was assumed in [150], which reported a similar average signal $\langle \Delta T \rangle = -0.51\ \mu$K for the 50 most extreme density perturbations of fixed radius $r = 130\ h^{-1}$Mpc expected in a ΛCDM cosmology.

---

[3] This could be evaluated at any redshift, provided an appropriate modification were made to $G(z)$. In fact we choose to evaluate it at $z = 0.52$.

Subsequently it was argued [151] that this profile is *not* the appropriate choice for density perturbations on $\gtrsim 100\ h^{-1}$Mpc scales, and that an *uncompensated* gaussian density profile will give larger values of ΔT. Using such a profile, [156] performed a template fit to the data and claimed the discrepancy with theory was only $\sim 2\sigma$. However, to arrive at this result they have used a template fitting approach that implicitly includes underdense regions with a physical density contrast $\delta < -1$ in the template, which is physically impossible.[4]

It is thus necessary to revisit this issue. Using the statistics of a homogeneous, isotropic, gaussian density field, we now derive the expected mean density profiles of superstructures of all density contrasts and all sizes, as well as the expected number density of such superstructures.

4.3.1 The number density of structures on different scales

We identify superstructures of different sizes with points of extrema of the linear density perturbation field $\delta(\mathbf{r})$ when smoothed over different scales. Overdensities correspond to peaks of the smoothed field and underdensities to points of minima. In the ΛCDM model, $\delta(\mathbf{r})$ is a homogeneous and isotropic, gaussian-distributed random field and the statistical properties of the maxima and minima have been calculated by [152] (hereafter BBKS). We briefly review below their key results and introduce necessary notation. Some lengthy expressions are however relegated to the Appendix.

Let $P(k,t)$ denote the matter power spectrum, defined as the Fourier transform of the two-point correlation function $\xi(r,t)$ of the density field at time t. We define a set of spectral moments weighted by powers of k,

$$\sigma_j^2(t) = \int \frac{k^2\,dk}{2\pi^2} W^2(kR_f) P(k,t) k^{2j}\,, \tag{4.8}$$

where $W(kR_f)$ is the window function appropriate to the filter used to smooth the density field, and R_f is the comoving smoothing scale. Thus σ_0 is just the standard rms fluctuation of the smoothed density field. We use a Gaussian filter to smooth the density field, corresponding to $W(kR_f) = \exp(-k^2R_f^2/2)$ to define the spectral

---

[4] While the linearly extrapolated density contrast can be taken to arbitrarily low values for calculating abundances, the ISW effect is sensitive to the true density contrast which by definition must always be > -1.

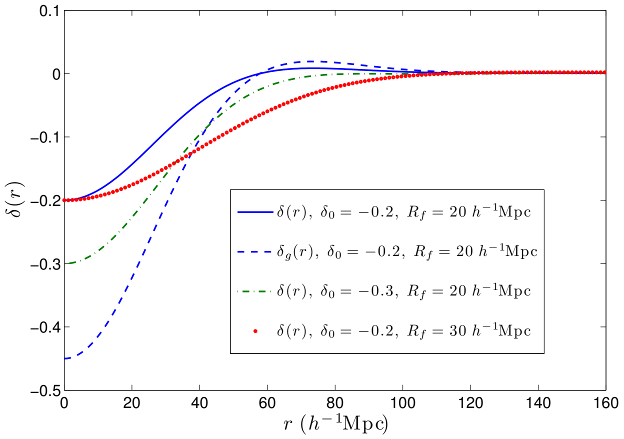

Figure 4.1: Mean radial profiles for voids obtained from Eq. (4.13) for different values of central underdensity δ_0 and the smoothing scale R_f. The blue (solid) line and the green (dash-dot) line are for the same R_f (= 20 h^{-1}Mpc) but different values of δ_0, whereas the red (dotted) profile has the same δ_0 (= −0.2) as the blue (solid) profile, but a larger smoothing scale. The blue (dashed) curve is the biased galaxy density contrast corresponding to matter density contrast given by the blue (solid) line, with bias factor $b = 2.25$ as is appropriate for LRGs.

parameters

$$\gamma \equiv \frac{\sigma_1^2}{\sigma_2\sigma_0}\,, \quad R_* \equiv \sqrt{3}\frac{\sigma_1}{\sigma_2}\,. \tag{4.9}$$

In terms of these quantities, the (comoving) differential number density $\mathcal{N}_{\rm max}(\nu)$ of maxima of height $\delta_0 = \nu\sigma_0$ has been shown to be [152]

$$\mathcal{N}_{\rm max}(\nu)d\nu = \frac{1}{(2\pi)^2R_*^3}e^{-\nu^2/2}G(\gamma,\gamma\nu)d\nu\,. \tag{4.10}$$

The explicit form of the function $G(\gamma,\gamma\nu)$ is given in Eq. (A.3) in the Appendix. We use a fitting form, accurate to better than 1%, which is provided in Eqs. (A.5-A.7). The density of minima is related to that of maxima through $\mathcal{N}_{\rm min}(\nu) = \mathcal{N}_{\rm max}(-\nu)$.

From Eq. (4.10) it is clear that the number density of maxima falls off exponentially with the relative height of these maxima, as intuitively predicted. Equally intuitive, though less obvious from the equation as written, is the fact that the number density also drops off sharply with increasing smoothing scale R_f.

4.3.2 Mean radial profiles

Having identified superstructures with the sites of maxima or minima in the smoothed density field, we wish to determine the mean radial variation of the density field in the neighbourhood of these extrema. BBKS show that, given that a maximum $\delta = \delta_0$ exists at the point $\mathbf{r} = 0$, the mean shape in the vicinity of this point, after averaging over all possible orientations of the principal axes as well as all values of the curvature at $\mathbf{r} = 0$, is

$$\bar{\delta}(r) = \frac{\delta_0}{(1-\gamma^2)}\left(\psi + \frac{R_*^2}{3}\nabla^2\psi\right) - \frac{\langle x|\delta_0\rangle\sigma_0}{\gamma(1-\gamma^2)}\left(\gamma^2\psi + \frac{R_*^2}{3}\nabla^2\psi\right) , \tag{4.11}$$

where $\psi(r) \equiv \xi(r)/\xi(0)$ is the normalised density-density correlation function and $x = -\nabla^2\delta/\sigma_2$. The expectation value of x given the knowledge that a peak of height δ_0 exists at $\mathbf{r} = 0$, denoted by $\langle x|\delta_0\rangle$, is approximately

$$\langle x|\delta_0\rangle = \gamma\nu + \Theta(\gamma, \gamma\nu) , \tag{4.12}$$

where $\nu = \delta_0/\sigma_0$ and $\Theta(\gamma, \gamma\nu)$ given by the fitting function in Eq. (A.10) in the Appendix. It follows from the definitions of the various quantities that Eq. (4.11) can be rewritten as [157]

$$\bar{\delta}(r,t) = \frac{1}{\sigma_0}\int_0^\infty \frac{k^2}{2\pi^2}\frac{\sin(kr)}{kr}W^2(kR_f)P(k,t)\left[\frac{\nu - \gamma^2\nu - \gamma\Theta}{1-\gamma^2} + \frac{\Theta R_*^2 k^2}{3\gamma(1-\gamma^2)}\right] dk . \tag{4.13}$$

We use this form for numerical evaluation of profiles. Note that γ and R_* depend on the smoothing scale R_f.

For the purposes of calculating the temperature signal, we make the simplifying approximation that all superstructures in the ensemble have the mean profile of Eqs. (4.11) or (4.13), dependent on the height of the peak after smoothing and on the smoothing scale (the actual distribution of the size of structures depends on the distribution of x values, which can be obtained from Eq. (A.8)). If the ensemble contains a sufficiently large number of superstructures this amounts to assuming that the average temperature signal for the ensemble of different profiles is the same as the temperature signal due to a superstructure with the average profile. Such an assumption is justified because of the linear relationship between ΔT, Φ and δ.

In Fig. 4.1 we plot some example underdense profiles for selected values of δ_0 and R_f. These examples are chosen for clarity and are not representative of the most likely actual underdensities. It can be seen that the size of the structures is rather larger than the smoothing scale R_f. Profiles calculated in this manner are somewhat narrower than those obtained from the simpler Gaussian form used in [151] and [156], which is the appropriate form for the density profile around a point of δ_0, *without* the knowledge that the point is an extremum of the density field. Nevertheless, the numerical difference is small and is not the cause of the difference in our conclusions. Note also that all our profiles turn over (i.e. $\bar{\delta}(r)$ and $\bar{\delta}'(r)$ both change sign at large r), though this cannot be seen in all cases in Fig. 4.1 because of the vertical scale used.

To identify superstructures in galaxy surveys (the methodology of G08a is discussed in more detail in Section 4.4.2) the galaxy density contrast δ_g is assumed to be linearly biased with respect to the matter density: $\delta_g = b\delta$. Denoting by $\rho_{\rm sl}$ the value of the density field at turnover, and by ρ_0 the minimum density at the centre, a selection cut is made on $w \equiv \rho_{\rm sl}/\rho_0$, effectively amounting to a lower bound on the absolute value of δ_0. This avoids false detections of over- and under-dense regions that are just Poisson fluctuations in galaxy number counts and ensures that only extreme superstructures are included in the ensemble.

The definition of the radius R_v of any of the voids shown in Fig. 4.1 is slightly ambiguous. We choose it to be the radius of turnover in the density profile less the smoothing scale R_f, since smoothing necessarily increases the radius of structures somewhat. Note that this is a small correction since in general $R_v \gg R_f$.

4.3.3 Temperature signal

The ISW signal of any one superstructure will be too small compared to the primordial CMB anisotropies to be observed individually. Therefore, what is measured is the average CMB temperature anisotropy over all the patches along the lines of sight of selected sample of either over- or under-densities. The primordial anisotropies are not correlated with the large scale structure and average out. Therefore, if a large enough number of patches are taken, the correlated ISW signal will dominate. Our calculation of this averaged signal is calculated as follows.

At a given z, we use Eq. (4.13) to calculate the matter and galaxy density profiles about extrema of the density field, as functions of central density contrast δ_0 and smoothing scale R_f, and obtain $\dot{\Phi}$ along any line of sight as discussed in Section 4.2. This allows us to calculate the temperature shift induced in the CMB, $\Delta T(\theta; \delta_0, R_f)$, as a function of the angle θ in the sky, where $\theta = 0^\circ$ corresponds to the line of sight passing directly through the centre of the superstructure.

To compare with observations we first apply the selection criterion on δ_0 through the limit on w. Then, to calculate the expectation value $\langle \Delta T \rangle$ for the resulting ensemble of superstructures, we weight the results appropriately using the number density of extrema, Eq. (4.10). This quantity can then be compared to the signal that should be observed when averaging over a number of CMB patches. For an ensemble of voids, the expectation value is

$$\langle \Delta T \rangle = \frac{\iint W(\theta)\Delta T(\theta;\delta_0)\mathcal{N}_{\rm min}\sigma_0^{-1}\ d^2\theta d\delta_0}{\pi\theta_c^2 \int \mathcal{N}_{\rm min}\sigma_0^{-1}\ d\delta_0}, \tag{4.14}$$

where $0 \leq \theta \leq \theta_{out}$; $W(\theta)$ is a filter whose form is chosen in order to match that used in the actual observation; and $-1 \leq \delta_0 \leq \delta_0^{\rm c}$, where $\delta_0^{\rm c}$ is the (mildly R_f-dependent) cutoff on the minimum underdensity required for the void to pass the selection criterion.

The choice of the smoothing scale R_f determines the mean radial size of the voids included within the ensemble, and affects the expectation value $\langle \Delta T \rangle$. Although structure-finding algorithms may not have an explicit size dependence, in practice there is obviously a lower limit on the size of the over- or under-density that can be reliably found. As smaller structures are overwhelmingly more probable than larger ones and also cause smaller temperature shifts, it is important to capture this effect. We discuss this in more detail in the next Section. For a given ensemble of voids specified by $R_{\rm f}$ and $\delta_0^{\rm c}$, the observed temperature signals ΔT will be distributed about this mean value. By calculating $\langle \Delta T^2 \rangle$ in a similar fashion to Eq. (4.14) we can estimate the standard deviation of this distribution simply as $(\langle \Delta T^2 \rangle - \langle \Delta T \rangle^2)^{1/2}$.

The expression for the expected signal from an ensemble of overdensities follows in an exactly analogous manner to Eq. (4.14).

4.4 Comparing theory to observation

4.4.1 The measured ISW signal of superstructures

In order to compare our expectation for the ISW signal to the specific observation made in G08a, it is necessary to carefully follow exactly the same procedure. We therefore briefly recap the aspects of the observation that are relevant to our analysis.

G08a use a sample of 1.1 million LRGs in a redshift range $0.4 < z < 0.75$ (with a mean redshift of $z = 0.52$) from the SDSS DR6 [158], which covers 7500 square degrees of the sky, and a total volume of about 5 $h^{-3}\mathrm{Gpc}^3$. Within this sample they search for 'supervoids' and 'superclusters' using two publicly-available structure-finding algorithms: ZOBOV (ZOnes Bordering On Voidness, [159]) for supervoids, and VOBOZ (VOronoi BOund Zones, [160]) for superclusters.

It is necessary to mimic the precise way in which these algorithms find and select structures when choosing the ensemble for which to calculate $\langle \Delta T \rangle$ from Eq. (4.14). ZOBOV uses a parameter-free Voronoi tessellation to estimate the density at each galaxy in the sample, based on the distance to its nearest neighbours. Around each density minimum it then finds the region of the density depression, or void. Of course large voids can contain multiple smaller voids, or even isolated high-density regions. The 'significance' of the depression is estimated by comparing the density contrast w — defined as the ratio of the density at the lip of the void to the density at the the point of minimum — to a uniform Poisson point sample in order to obtain the likelihood that a void of density contrast w could arise from Poisson noise, i.e., that it is a false positive detection. A 3σ cutoff is then applied on the likelihood, which translates to a cutoff of $w > w_c = 2.0$ on the density contrast [159]. Using this cutoff, the authors find a total of 50 voids in the survey volume, the properties of which are tabulated in [149] (hereafter G08b).

The VOBOZ cluster finder uses the same algorithm but applied to the inverse of the density field, with density contrast defined as the ratio of the peak density to the density at the edge of the structure. However, in this case the 3σ cut on the likelihood that an overdensity of given w could have arisen due to Poisson noise is at $w > 6.8$ [160]. In fact G08a instead impose a slightly tighter cut, corresponding to $w > 8.35$, in order to obtain a sample of exactly 50 such superclusters: their properties are also tabulated in G08b.

Having thus identified the desired sample of superstructures in the SDSS survey, G08a then search for the ISW temperature signals of these structures using an inverse-variance weighted combination of the WMAP 5-year Q, V and W maps [161] with foreground subtracted and the KQ75 mask applied. They build stacked images by averaging the CMB temperature in the regions around the lines of sight passing

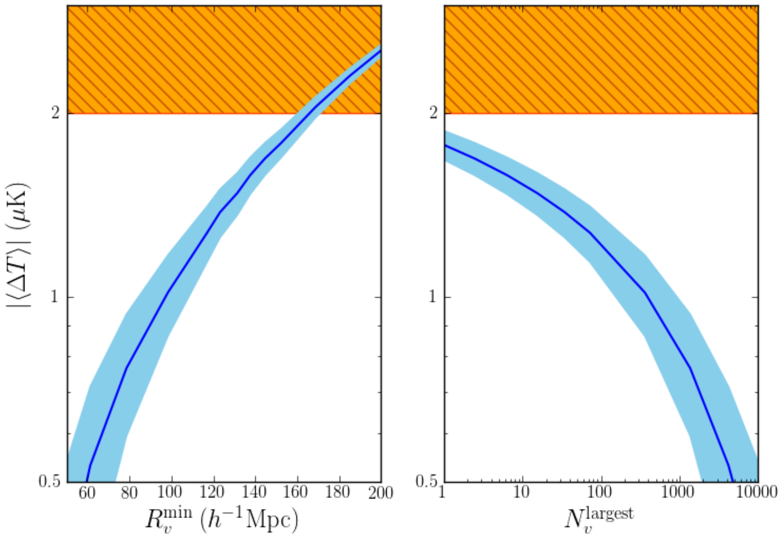

Figure 4.2: *Left panel*: The absolute value of $\langle\Delta T\rangle$ for an ensemble of supervoids which satisfy the ZOBOV selection condition on density (see text), as a function of the minimum radius of supervoids included. The solid (blue) curve shows the mean value and the shaded (lighter blue) contours the 1σ region. The (orange) cross-hatched area is the lower end of the 3σ range of the observed value $\langle\Delta T\rangle_{\rm obs} = -11.3 \pm 3.1$ μK. *Right panel*: As above, but showing $\langle\Delta T\rangle$ as a function of the number of supervoids in the ensemble from which the observed sample of 50 supervoids is to be drawn, when only the $N_{\rm v}$ *largest* supervoids also meeting the ZOBOV selection condition on density are included.

through the centres of the identified superstructures and use a compensated top-hat filter of width $\theta_{\rm c}$ in order to perform the averaging. This corresponds to making the choice

$$W(\theta) = \begin{cases} 1\,, & 0 \leq \theta \leq \theta_c; \\ -1\,, & \theta_c < \theta \leq \theta_{out}, \end{cases} \tag{4.15}$$

in Eq. (4.14), with $\theta_{out} = \sqrt{2}\theta_c$. For a filter radius $\theta_c = 4°$,[5] the sample of supervoids gives an average $\Delta T = -11.3 \pm 3.1$ μK, and the sample of superclusters an average $\Delta T = 7.9 \pm 3.1$ μK. When averaged together, with the negative of the supercluster image added to that of the supervoids, the value obtained is $\Delta T = -9.6 \pm 2.2$ μK, i.e., a 4.4σ detection.

---

[5] The choice of filter radius $\theta_c \sim 4°$ appears to be based on the theoretical prediction that the full CMB-galaxy cross-correlation should peak at about 4° [142]. G08a repeat the observation with a few other widths, $3° \leq \theta_c \leq 5°$ and obtain a maximum detection significance for $\theta_c = 4°$.

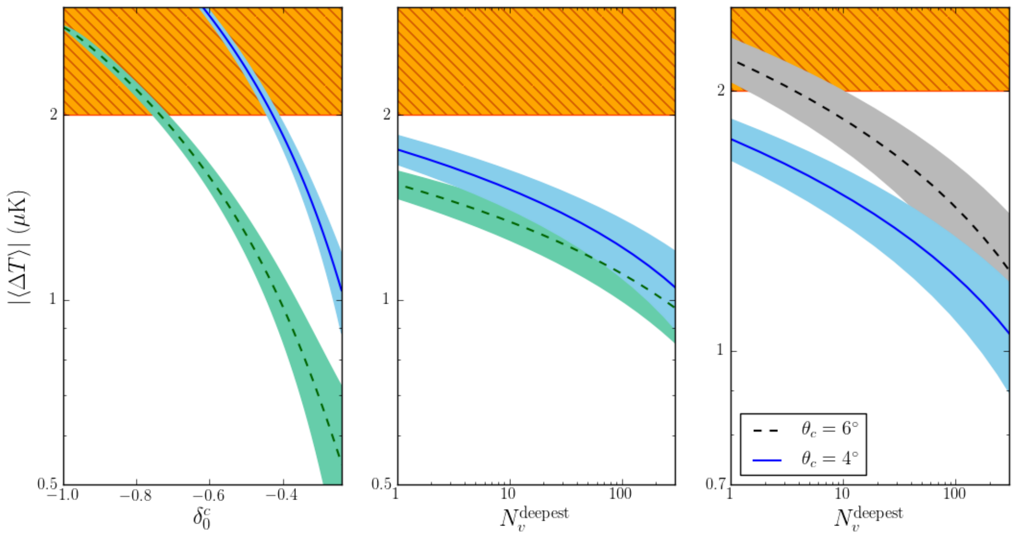

Figure 4.3: *Left panel*: The absolute value of $\langle\Delta T\rangle$ for an ensemble of supervoids which satisfy $\delta_0 < \delta_0^c$, as a function of δ_0^c. The dashed (green) curve shows the case when R_f is chosen such that voids with $R_{\rm v} \gtrsim 70\ h^{-1}$Mpc are included in the ensemble; the solid (blue) curve is for $R_{\rm v} \gtrsim 100\ h^{-1}$Mpc. Shaded contours show the 1σ region about the mean and the orange cross-hatched area is as in Fig. 4.2. *Middle panel*: As before, but showing $\langle\Delta T\rangle$ as a function of the number of supervoids in the ensemble from which the observed sample of 50 supervoids is to be drawn, when only the $N_{\rm v}$ *deepest* supervoids are included. *Right panel*: As before, but for two different choices of the radius $\theta_{\rm c}$ of the compensating top-hat filter used in eq. (4.14). The solid (blue) curve is the mean value for $\theta_{\rm c} = 4°$, and the broken (black) curve for $\theta_{\rm c} = 6°$. Shaded regions show the 1σ deviations from the mean.

4.4.2 Comparison

We now compare the observed result described above to the theoretical result described in Section 4.3. We assume for simplicity that the centres of all superstructures are located at the mean redshift $z = 0.52$. We estimate this simplification to have negligible effects on the end result, as discussed in Section 4.4.2. We use a standard ΛCDM cosmological model with $\Omega_m = 0.29$, $\Omega_\Lambda = 0.71$, $n_s = 0.96$ and $\sigma_8 = 0.83$ (with $h = 0.69$ where required) to obtain the matter power spectrum at redshift $z = 0.52$ using CAMB [29].[6] The bias factor for LRGs is taken to be $b = 2.25$.

Our first finding is that there are *no* overdense superstructures within the linear regime (i.e., with $\delta_0 < 1$) which meet the VOBOZ 3σ-significance selection criterion

---

[6] These are the mean parameter values obtained from a fit to WMAP 7-year [15] and SDSS DR7 [162] data using COSMOMC [128].

that $w > 6.8$, let alone the $w > 8.35$ cut applied by G08a. Such a ratio of densities between the lip of an overdensity and the centre can only be achieved in highly non-linear collapsed structures. We conclude that Table 5 of G08b does not list the most overdense large-scale linear perturbations, but the (mild) large-scale linear perturbations that happen to contain the most overdense small-scale collapsed structures. This means that the criteria used to select the overdensities tabulated in G08b are being affected by collapsed structures in a manner that our methodology is unable to capture, so we are unable to estimate the expected ISW effect for that sample. Note however that our calculations below for the maximum possible amplitude of ISW signal from superstructures in the SDSS region holds equally well for over- and under-densities in the linear regime. Any contamination of the regions selected by VOBOZ will only *reduce* the expected signal, although we are unable to estimate by how much.

Hence we concentrate only on the sample of underdense regions (supervoids), which does not suffer from this problem, and for which the observed average signal is $\langle \Delta T \rangle_{\rm obs} = -11.3 \pm 3.1\ \mu$K. To calculate the expected $\langle \Delta T \rangle$ using Eq. (4.14), we must first choose the smoothing scale R_f. This choice determines the distribution of void radii R_v in the ensemble according to Eq. (A.8). For each R_f we calculate $R_v^{\rm min}$ such that 95% of all voids in the ensemble have radius $R_v > R_v^{\rm min}$, this being a convenient way to characterise the ensemble. At $R_f = 20\ h^{-1}$Mpc, we find $R_v^{\rm min} \sim 70\ h^{-1}$Mpc, which is similar to the mean radius of the supervoids in Table 4 of G08b.[7] At this smoothing scale, the number of such voids within the SDSS survey volume that should satisfy the selection criterion on the density is $N_v \sim 10^4$. This is to be compared with the $N_v = 50$ voids that are actually tabulated in G08b. For the larger ensemble, we find an expectation value $\langle \Delta T \rangle = -0.3 \pm 0.2\ \mu$K, i.e., consistent with zero and indicative of enormous tension with the observation.

However, since G08a only see such a small fraction of the total number of supervoids expected in the SDSS survey volume, strong selection effects must be in operation. These will be due to the fact that we have neglected shot noise, i.e., our assumption that the galaxy distribution smoothly traces the total matter distribution.

---

[7] A precise comparison between our radii and theirs is not possible because due to effects of shot noise. In particular, at the radii reported in G08b, the density has not yet reached the background level and so their values must systematically somewhat underestimate the sizes of the voids compared to our definition.

While this assumption has no effect on our ISW predictions, because they relate to the total matter distribution, it will have an effect on the selection of superstructures from the LRG data. The selection effects that could enhance the expected signal are skews towards larger and deeper regions, i.e., if G08a did not see a random sample of 50 of the $\sim 10^4$ expected supervoids but chose some sample that is skewed towards regions with larger ΔT values. We show below that the expected signal from the 50 *most extreme* regions is indeed ~ 6 times larger than $-0.3\ \mu$K; nevertheless the discrepancy with observation is still $> 3\sigma$. Therefore, irrespective of how the supervoids were selected by G08a, tension remains with the expectation of the standard ΛCDM model.

Accounting for selection effects

We first consider the possibility that the void-finding algorithm ZOBOV is sensitive to only the largest (and least common) supervoids in the matter distribution. These voids will also produce larger ISW temperature signals. Even though ZOBOV does not explicitly take the void size into account, this is highly plausible. The average density of the LRGs in the SDSS DR6 sample is roughly 1 galaxy per $(15\ h^{-1}\text{Mpc})^3$. In underdense regions LRGs will be even more sparsely distributed, so ZOBOV will certainly be less able to identify smaller underdense structures, thus biasing the sample towards larger voids.

In order to model the effect of such a selection bias, we increase the value of R_f in Eq. (4.14); this is equivalent to only including the N_v *largest* supervoids, with radius $R_v \geq R_v^{\min}$, in the ensemble from which $\langle \Delta T \rangle$ is calculated. In the left panel of Fig. 4.2 we plot the expectation $\langle \Delta T \rangle$ as a function of $R_v^{\min}$. The orange cross-hatched area shows the region that is within 3σ of the observed value $\langle \Delta T \rangle_{\text{obs}} = -11.3 \pm 3.1\ \mu$K; only the lower portion of this region is shown for clarity. The theoretical value of $\langle \Delta T \rangle$ becomes marginally consistent ($\sim 3\sigma$) with the observed value when $R_v^{\min} \sim 170\ h^{-1}$Mpc. However the probability that within the SDSS survey volume there are 50 supervoids of radius $R_v \geq 170\ h^{-1}$Mpc that also meet the ZOBOV selection criterion is negligibly small.

In the right-hand panel of Fig. 4.2 we plot $\langle \Delta T \rangle$ as a function of the size N_v of the ensemble of largest supervoids that should exist within the SDSS volume. It is from this ensemble that the 50 observed supervoids should be regarded as having been drawn. It can be seen that even under the assumption that the ZOBOV algorithm selected exactly the 50 largest supervoids in the entire SDSS survey volume, the expected signal is only $\langle \Delta T \rangle = -1.33 \pm 0.13\ \mu\mathrm{K}$ which is still $> 3\sigma$ discrepant with the observed value. We conclude that the observed signal cannot be explained due to a simple bias towards selecting only the largest supervoids.

It is interesting to note that the 50 largest supervoids expected within the SDSS survey volume correspond to $R_v^{\mathrm{min}} \sim 120\ h^{-1}$Mpc. The largest void radius reported in G08a is $\sim 125\ h^{-1}$Mpc and the mean is $\sim 70\ h^{-1}$Mpc. We have argued that these values somewhat underestimate the size of the structures compared to our definition of R_v, yet it seems unlikely that the difference could be so large that *all* the voids tabulated in G08a should have $R_v \geq 120\ h^{-1}$Mpc.

We next consider whether the ZOBOV algorithm is more sensitive to deeper supervoids. This is also quite likely. In Table 4 of G08b, the edge of most of the voids is defined at a radius where the density contrast is still negative. This means ZOBOV will be systematically underestimating the value of w relative to the definition we have used (where $\delta_{\mathrm{edge}} \simeq 0$), so the G08a observation effectively used a more stringent cut on w than the one we have used. We can model this effect by varying δ_0^{c} from the value determined by the stated algorithm. In the left-hand panel of Fig. 4.3 we plot $\langle \Delta T \rangle$ as a function of δ_0^{c} for two radius cutoff values $R_v^{\mathrm{min}} \sim 70\ h^{-1}$Mpc (the mean radius of the supervoids in G08b) and $R_v^{\mathrm{min}} \sim 100\ h^{-1}$Mpc as examples. The middle panel shows $\langle \Delta T \rangle$ as a function of N_v, the number of voids included in the ensemble when δ_0^{c} is varied in each case. For the smaller radius cutoff, the ensemble is dominated by smaller voids with a small ISW effect, so increasing R_v^{min} increases $\langle \Delta T \rangle$ at any δ_0^{c}. With $R_v^{\mathrm{min}} \sim 100\ h^{-1}$Mpc, $\langle \Delta T \rangle$ becomes marginally consistent with observation for $\delta_0^{\mathrm{c}} \lesssim -0.5$. However, the probability of obtaining 50 supervoids with $R_v \geq 100\ h^{-1}$Mpc and $\delta_0 \leq -0.5$ is again negligibly small, as the middle panel clearly demonstrates.

Discussion of potential systematic errors

We now discuss the expected corrections due to the simplifying assumptions we made in calculating the signal. Our first approximation was to neglect non-linear effects and the time evolution of the void density profiles [155, 163]. On small scales voids evolve towards a compensated top-hat profile through non-linear evolution [155], but this produces a *smaller* net ISW effect due to the effect of the overdense ridge at the boundary [150, 151]. Thus by using the linear theory profiles we are somewhat *over*estimating the expected ΔT, although on the scales of interest, $\gtrsim 100\ h^{-1}$Mpc, the effects of the non-linear evolution will be small in any case.

A more subtle assumption is that the real supervoid profiles are adequately described by the profiles obtained following our smoothing prescription. The effect of the smoothing is to slightly broaden the $\delta(r)$ profile, which in turn leads to a small broadening of the $\Delta T(\theta)$ profile. This may become a problem if the $\Delta T(\theta)$ profile is significantly broadened so that the compensated top-hat filter of radius 4° then underestimates the real signal. If this were the case then it would be more appropriate to use a broader top-hat than 4° in Eq. (4.14). To check this effect, we repeated our analysis with a filter of radius 6°, which is a generous overestimate of the degree of broadening caused by the smoothing. In the right-hand panel of Fig. 4.3 we show the effect of this change for the largest and deepest supervoids. As expected, increasing the filter radius does increase $\langle\Delta T\rangle$, but the effect is small. Even with a 6° filter the expected signal remains $> 3\sigma$ discrepant with observation at $N_v = 50$; as mentioned above, the actual effect on $\langle\Delta T\rangle$ due to the smoothing will be less than this.

As a further test of the robustness of our calculation, we compare our results with Fig. 1 of [153], which shows the ISW map from a cosmological N-body simulation for a volume comparable to the SDSS DR6, but at $z = 0$ rather than $z = 0.52$. The *very* largest density perturbations in this map correspond to a maximum ISW temperature shift of $|\Delta T| \sim 4\ \mu$K before applying a filter analysis. Taking into account the fact that ΔT is more pronounced at smaller redshift, this appears to tally very well with our prediction from Fig. 4.3 that the most extreme supervoid in the SDSS volume should produce $\langle\Delta T\rangle \sim -2\ \mu$K. This indicates that any correction due to ignoring the redshift distribution must also be small.

As seen in the full sky maps of [153], there are lines of sight along which the chance alignment of several large structures between us and the LSS several Gpc away can lead to isolated ISW 'cold spots' with $\Delta T < -10\ \mu$K. However, this is a cumulative effect of several structures over Gpc scales and does not affect our conclusions about the correlation between these cold spots and the N isolated structures at $z \sim 0.5$ when N is large. Similarly, there will be several 4° circles on the sky for which the underlying CMB anistropy alone can give an average ΔT of similar magnitude—it is precisely these fluctuations that generate the observational uncertainty of $\pm 3.1\ \mu$K. However such fluctuations are not due to the ISW effect and there is no reason for them to be correlated with large structures in the SDSS catalogue. Therefore, assuming only that the lines of sight along which to perform the analysis were not chosen *a posteriori*, the observed signal can only result from rare ($> 3\sigma$) fluctuations or anomalously large density perturbations at $z \sim 0.5$.

Interestingly G08a report a slightly lower significance detection for both $N = 30$ and $N = 70$ compared to $N = 50$. Fluctuations due to the underlying CMB anisotropy dominate at small N and false positive identifications of structures increase at large N so the signal-to-noise ratio is expected to have a maximum at some intermediate N, but a more detailed study is needed to quantify where this should be.

G08a also study the variation of the signal-to-noise ratio with the width of the compensating top-hat filter, θ_c, and report a maximum at $\theta_c = 4°$ for the combined sample of over-and underdense structures. However, for supervoids alone we find that $|\langle \Delta T \rangle|$ increases slightly as θ_c is increased from 4° to 6° (Fig. 4.3). This is further evidence that the G08a sample of superstructures *cannot* be modelled by linear structures in ΛCDM.

4.5 Summary and prospectives

We have calculated the integrated Sachs-Wolfe effect expected in ΛCDM from superstructures of sizes $\sim 100\ h^{-1}$Mpc, using the density profiles predicted by the linear theory of Gaussian perturbations [152], rather than assuming their profiles *a priori.* We find that the few most extreme regions surveyed by SDSS will produce an ISW signal of $\sim 2\ \mu$K. This matches well with ISW maps generated from N-body simulations of the ΛCDM cosmology [153].

Our result is about 4 times larger than earlier calculations which assumed compensated top-hat density profiles [90, 150], which while well motivated for non-linear voids that have formed at small scales [155], should not apply on the large scale of the superstructures considered here. This is almost entirely because of the choice of density profile. Due to the fact that the gravitational potential leaks out of non-compensated profiles it was pointed out in [151] that the ISW effect from such structures should be larger, and our results confirm this.

Nevertheless we have demonstrated that the ISW signal claimed to have been detected by Granett et al. [148] is still $> 3\sigma$ larger than the signal expected in ΛCDM. This tension persists even after allowing for likely selection effects. In fact, even the most extreme underdensities in the SDSS volume would still produce a signal $> 3\sigma$ smaller than the observed signal. Therefore, no possible selection effect could be responsible for the observed signal. Our conclusion is that large, deep superstructures appear to be significantly more numerous than expected from a ΛCDM model with a Gaussian primordial power spectrum.

This differs from the conclusion of Papai et al. [156] who also used similar density profiles. This is not due to the small difference between their profiles and those assumed here. Instead it is because those authors incorrectly applied a method that was calibrated correctly at small density contrasts to density contrasts where it necessarily breaks down for voids (i.e., allowing $\delta < -1$ in their templates). Imposing the physical restriction $\delta > -1$ will decrease the amplitude of the template ISW signal calculated by these authors.

An interesting question is whether the expected signal of the most extreme superstructures in a ΛCDM universe is possible to detect in principle. For 50 4° circles of the CMB, [148] finds the statistical uncertainty is 3.1 μK. Hence the expected statistical uncertainty in the average for a sample of N such patches will be approximately $(3.1 \times \sqrt{50}/\sqrt{N})$ μK, so that for $N \simeq 3000$, the estimated error on a measurement of $\langle \Delta T \rangle$ is $\Delta T_{\mathrm{noise}} \simeq 0.4$ μK. Thus a detection of $|\langle \Delta T \rangle| = 1$ μK averaged over an ensemble of ~ 3000 superstructures can be made with roughly 2.5σ significance. From Fig. 4.2 it can be seen that for the 3000 largest supervoids in the SDSS DR6 survey volume, $|\langle \Delta T \rangle| \sim 1$ μK which is of the right order of magnitude but somewhat too small for detection at high significance. However, the SDSS window is not large enough to contain 3000 independent 4° patches on the sky so in any case a larger survey would be needed in order to measure a statistically significant signal, and such a survey would contain more supervoids. This order-of-magnitude estimate indicates that even if the G08a observation is a statistical anomaly, the ISW imprint of superstructures in a ΛCDM cosmology may be large enough to be detected in future surveys.

As noted earlier, the detection of the ISW imprint of individual superstructures provides an important complement to full-sky CMB-galaxy cross-correlation studies. It has the potential to provide information about the radii, density contrasts and density profiles of specific structures that lie in the extreme tail of the probability distribution function. Our calculation demonstrates that the predicted ISW signal from the most extreme superstructures is far too small to explain the temperature fluctuations seen by G08a, indicating a failing of the standard ΛCDM cosmology.

A likely explanation for this deviation is that the primordial perturbations are non-Gaussian. This would influence both the abundance of these extreme regions — e.g., see [164] — as well as their density profile, thus changing their expected ISW signal. As pointed out by [165], a primordial skewness, parameterised by $f_{\rm NL}$, would not be able to enhance the abundance of both over- and underdense regions simultaneously; however a primordial kurtosis, parameterised by a positive $g_{\rm NL}$, would indeed do so, and be less constrained by the CMB. Note that non-Gaussianity disproportionately affects the tail of the distribution of density perturbations, which is where most of the contribution to the ISW effect of individual superstructures comes from. Therefore, primordial non-Gaussianity may be able to explain this signal while preserving the success of ΛCDM on other fronts. However, this would then undermine the use of the ISW effect as an independent test for Λ.

Another possible explanation might lie in a modification of the growth rate of perturbations as can happen in, e.g., models based on scalar-tensor gravity [166]. The presence of large-scale inhomogeneities can themselves alter the growth rate and this too deserves further attention.

Chapter 5

Conclusions

In this thesis we have studied aspects of inflation, the effect of inhomogeneities on observations, the generation of structure in the universe and its signatures. In particular, we have used current observational data to test aspects of the currently accepted ΛCDM cosmological model and to constrain alternative scenarios.

The density perturbations observed today are generated at the time of inflation, so it is important to ask what models of inflation can generate a power spectrum of perturbations consistent with observation, and how these models can be accommodated within theories of particle physics beyond the Standard Model (SM). An obvious extension of the SM is supersymmetry, and in particular the minimal supersymmetric SM (MSSM). In Chapter 2 we briefly discussed some models of inflation based on MSSM flat direction fields, and specifically one particular such model with a renormalizable potential. We examined the constraints on the parameter space of this model arising simply from the properties of the primordial power spectrum of fluctuations inferred from the measurements of CMB anisotropies by the WMAP satellite in the standard cosmological framework. We quantified the fine-tuning problem that exists in this model and showed how the model might be extended in order to alleviate it.

However, the assumed power-law form of the primordial power, although consistent with the WMAP data when the standard cosmological assumptions are made, is not the only possible prediction that can arise from inflation, especially in models where flat direction fields are coupled to the inflaton. Such flat direction fields are generically present in supersymmetric theories, and if they were to undergo a phase transition during inflation, this might cause a localised feature in the power spectrum. In Chapter 3, we discussed how such a localised feature can even allow an alternative cosmological model, with a local void but no dark energy, simultaneously to fit data from WMAP, Type Ia supernovae, local Hubble measurements, baryon acoustic oscillations, and nucleosynthesis. This model, though simplistic, serves as an example that the standard ΛCDM model need not be the only one that can fit the cosmological data. Although ultimately simple void models may be found to be in conflict with some other observations, they highlight the degeneracies in cosmological datasets, and the need to find dynamical evidence of the effects of dark energy.

In Chapter 4 we examine the ISW effect, which is exactly such a dynamical test. We discuss a recent observation of the ISW temperature signal from giant superstructures identified from galaxy surveys, which has been commented on in the literature. We provide a theoretical calculation of the expected size of this signal in a ΛCDM universe with Gaussian primordial fluctuations, making as few *a priori* assumptions as possible, and find that the observed signal is discrepant with the prediction at high significance. If this signal is indeed due to the ISW effect, we conclude that the observed universe is significantly inhomogeneous on larger scales than predicted in ΛCDM, and explaining this discrepancy may lead us to new physical insights or new models, possibly including models of inflation that can generate significant primordial non-Gaussianity.

This thesis has of necessity touched on only a few of the aspects of modern cosmological enquiry. With the advent of the era of precision cosmological measurements, observations will continue to throw up many fascinating puzzles for theory to explain, and this will remain a fascinating field for many years to come.

Appendix A

Statistics of a Gaussian random field

In this appendix we list some of the functions used in Chapter 4. A full derivation of these expressions from a detailed treatment of the statistics of Gaussian random fields is beyond the scope of this thesis, but has been provided in the very thorough paper by Bardeen, Bond, Kaiser and Szalay [152]. We provide the details of the expressions merely to aid reference.

A.1 The number density of extrema

Let $\delta(\mathbf{r})$ be a *homogeneous and isotropic Gaussian random field*, i.e., for each point $\mathbf{r}$ in three-dimensional real space $\delta(\mathbf{r})$ is a random variable such that the m-point probability distribution functions

$$P\left[\delta(\mathbf{r}_1), \delta(\mathbf{r}_2), \ldots, \delta(\mathbf{r}_m)\right] d\delta(\mathbf{r}_1) d\delta(\mathbf{r}_2) \ldots \delta(\mathbf{r}_m) \ , \tag{A.1}$$

are multivariate Gaussians with zero mean, all finite dimensional distributions are invariant under translations of the points $\mathbf{r}_i$ by the same vector and the two-point correlation function $\xi(\mathbf{r}_1 - \mathbf{r}_2) = \langle \delta(\mathbf{r}_1)\delta(\mathbf{r}_2) \rangle = \xi(|\mathbf{r}_1 - \mathbf{r}_2|)$ is rotation-invariant.

We define the power spectrum $P(k,t)$ as the Fourier transform of $\xi(r,t)$ at time t, and define a set of spectral moments weighted by powers of k, $\sigma_j^2(t; R_f)$, as in Eq. (4.8). Given the spectral parameters γ and R_* from Eq. (4.9), the (comoving) differential number density of peaks of the random field δ is given by

$$\mathcal{N}_{\max}(\nu) d\nu = \frac{1}{(2\pi)^2 R_*^3} e^{-\nu^2/2} G(\gamma, \gamma\nu) d\nu \ . \tag{A.2}$$

The function $G(\gamma, \gamma\nu)$ is defined by the integral

$$G(\gamma, x_*) = \int_0^\infty dx\ f(x) \frac{\exp[-(x-x_*)^2/2(1-\gamma^2)]}{[2\pi(1-\gamma^2)]^{1/2}}\ , \tag{A.3}$$

where the function $f(x)$ can be expressed in closed form as:

$$\begin{aligned} f(x) &= \frac{(x^3-3x)}{2}\left\{\text{erf}\left[\left(\frac{5}{2}\right)^{1/2} x\right] + \text{erf}\left[\left(\frac{5}{2}\right)^{1/2}\frac{x}{2}\right]\right\} \\ &+ \left(\frac{2}{5\pi}\right)^{1/2}\left[\left(\frac{31x^2}{4}+\frac{8}{5}\right)e^{-5x^2/8} + \left(\frac{x^2}{2}-\frac{8}{5}\right)e^{-5x^2/2}\right]. \end{aligned} \tag{A.4}$$

A fitting formula that agrees with the above form for $G(\gamma, w)$ to better than 1% accuracy over the range $0.3 < \gamma < 0.7$ and $-1 < w < \infty$, and with accuracy better than 0.1% for $w > 1$, is

$$G(\gamma, w) = \frac{w^3 - 3\gamma^2 w + [B(\gamma)w^2 + C_1(\gamma)]\exp[-A(\gamma)w^2]}{1 + C_2(\gamma)\exp[-C_3(\gamma)w]}\ , \tag{A.5}$$

where the various quantities A, B, C_i are:

$$A = \frac{5/2}{(9-5\gamma^2)}\ ,\ B = \frac{432}{(10\pi)^{1/2}(9-5\gamma^2)^{5/2}}\ , \tag{A.6}$$

$$C_1 = 1.84 + 1.13(1-\gamma^2)^{5.72}\ ,\quad C_2 = 8.91 + 1.27e^{6.51\gamma^2}\ ,\quad C_3 = 2.58e^{1.05\gamma^2}\ . \tag{A.7}$$

We use this fitting formula for numerical evaluation of number densities in Chapter 4.

Note that $\mathcal{N}_{\rm max}(\nu)d\nu$ specifically refers to the differential number density of points of points of maxima, rather than just extrema. The number density of minima is related to that of maxima by $\mathcal{N}_{\rm min}(\nu) = \mathcal{N}_{\rm max}(-\nu)$.

A.2 Density profiles around extrema

In the vicinity of a maximum or minimum of the density field $\delta(\mathbf{r})$, the averaged density profile can be written in terms of the variables we used in Section 4.3.2, $\psi(r) \equiv \xi(r)/\sigma_0^2$ and $x = -\nabla^2\delta/\sigma_2$ is a measure of the curvature at the extremum point. The conditional probability of x given that a peak of height $\delta_0 = \nu\sigma_0$ exists, is

$$P_x(x|\nu)dx = \frac{\mathcal{N}_{\rm max}(\nu, x)dx}{\mathcal{N}_{\rm max}(\nu)} = \frac{\exp[-(x-x_*)^2/2(1-\gamma^2)]}{[2\pi(1-\gamma^2)]^{1/2}}\frac{f(x)dx}{G(\gamma,\gamma\nu)}\ . \tag{A.8}$$

Given this probability distribution, the expectation value of x, given δ_0, is

$$\langle x|\delta_0\rangle = \gamma\nu + \Theta(\gamma, \gamma\nu)\ , \tag{A.9}$$

where

$$\Theta(\gamma, \gamma\nu) = \frac{3(1-\gamma^2) + (1.216 - 0.9\gamma^4)\exp[-\gamma/2(\gamma\nu/2)^2]}{[3(1-\gamma^2) + 0.45 + (\gamma\nu/2)^2]^{1/2} + \gamma\nu/2} \tag{A.10}$$

is an approximate fitting form. Using this expectation value for x, the mean, spherically-averaged, radial profile of the inhomogeneity about the central point of extremum can be written as in Eq. (4.11).

Bibliography

[1] **The Supernova Cosmology Project** Collaboration, S. Perlmutter *et. al. Astrophys.J.* **517** (1999) 565–586, [`astro-ph/9812133`].

[2] A. Conley, J. Guy, M. Sullivan, N. Regnault, P. Astier, *et. al. Astrophys.J.Suppl.* **192** (2011) 1, [`arXiv:1104.1443`].

[3] D. Larson, J. Dunkley, G. Hinshaw, E. Komatsu, M. Nolta, *et. al. Astrophys.J.Suppl.* **192** (2011) 16, [`arXiv:1001.4635`].

[4] For a review of the physics of the early universe, see the following.

E. W. Kolb and M. S. Turner *Front.Phys.* **69** (1990) 1–547.

[5] J. Guy, M. Sullivan, A. Conley, N. Regnault, P. Astier, *et. al. Astron.Astrophys.* **523** (2010) A7, [`arXiv:1010.4743`].

[6] J. A. Holtzman, J. Marriner, R. Kessler, M. Sako, B. Dilday, *et. al. Astron.J.* **136** (2008) 2306–2320, [`arXiv:0908.4277`].

R. Kessler, A. Becker, D. Cinabro, J. Vanderplas, J. A. Frieman, *et. al. Astrophys.J.Suppl.* **185** (2009) 32–84, [`arXiv:0908.4274`].

[7] A. G. Riess, L.-G. Strolger, S. Casertano, H. C. Ferguson, B. Mobasher, *et. al. Astrophys.J.* **659** (2007) 98–121, [`astro-ph/0611572`].

[8] **HST** Collaboration, W. Freedman *et. al. Astrophys.J.* **553** (2001) 47–72, [`astro-ph/0012376`].

[9] A. G. Riess, L. Macri, S. Casertano, H. Lampeitl, H. C. Ferguson, *et. al. Astrophys.J.* **730** (2011) 119, [`arXiv:1103.2976`].

[10] W. L. Freedman and B. F. Madore *Annu. Rev. Astron. Astrophys.* **48** (2010) 673–710, [`arXiv:1004.1856`].

[11] **Supernova Search Team** Collaboration, A. G. Riess *et. al. Astron.J.* **116** (1998) 1009–1038, [`astro-ph/9805201`].

[12] M. Hicken, W. Wood-Vasey, S. Blondin, P. Challis, S. Jha, *et. al. Astrophys.J.* **700** (2009) 1097–1140, [`arXiv:0901.4804`].

[13] R. Amanullah, C. Lidman, D. Rubin, G. Aldering, P. Astier, *et. al. Astrophys.J.* **716** (2010) 712–738, [`arXiv:1004.1711`].

[14] For a review of supernovae observations, see the following.

B. Leibundgut *Gen.Rel.Grav.* **40** (2008) 221, [`arXiv:0802.4154`].

[15] **WMAP** Collaboration, E. Komatsu *et. al. Astrophys.J.Suppl.* **192** (2011) 18, [`arXiv:1001.4538`].

[16] **SDSS** Collaboration, D. J. Eisenstein *et. al. Astrophys.J.* **633** (2005) 560–574, [`astro-ph/0501171`].

[17] G. Efstathiou, W. Sutherland, and S. Maddox *Nature* **348** (1990) 705–707.

S. D. White, G. Efstathiou, and C. Frenk *Mon.Not.Roy.Astron.Soc.* **262** (1993) 1023–1028.

[18] **2dFGRS** Collaboration, G. Efstathiou *et. al. Mon.Not.Roy.Astron.Soc.* **330** (2002) L29, [`astro-ph/0109152`].

[19] S. Sarkar *Gen.Rel.Grav.* **40** (2008) 269–284, [`arXiv:0710.5307`].

[20] S. Weinberg *Rev.Mod.Phys.* **61** (1989) 1–23.

[21] The following paper is a review of the status of quintessence models.

E. V. Linder *Gen.Rel.Grav.* **40** (2008) 329–356, [`arXiv:0704.2064`].

[22] For reviews, see the following papers.

S. Capozziello and M. Francaviglia *Gen.Rel.Grav.* **40** (2008) 357–420, [`arXiv:0706.1146`].

K. Koyama *Gen.Rel.Grav.* **40** (2008) 421–450, [`arXiv:0706.1557`].

[23] R. Bousso *Gen.Rel.Grav.* **40** (2008) 607–637, [`arXiv:0708.4231`].

[24] E. Bianchi and C. Rovelli `arXiv:1002.3966`.

[25] C.-P. Ma and E. Bertschinger *Astrophys.J.* **455** (1995) 7–25, [`astro-ph/9506072`].

[26] D. H. Lyth and A. R. Liddle, *The primordial density perturbation: Cosmology, inflation and the origin of structure.* Cambridge University Press, 2009.

[27] S. Weinberg, *Cosmology.* Oxford University Press, 2008.

S. Dodelson, *Modern Cosmology.* Academic Press, 2003.

A. R. Liddle and D. Lyth, *Cosmological inflation and large scale structure.* Cambridge University Press, 2000.

[28] J. M. Bardeen *Phys. Rev.* **D22** (1980) 1882–1905.

[29] A. Lewis, A. Challinor, and A. Lasenby *Astrophys.J.* **538** (2000) 473–476, [`astro-ph/9911177`].

[30] U. Seljak and M. Zaldarriaga *Astrophys.J.* **469** (1996) 437–444, [`astro-ph/9603033`].

[31] J. M. Bardeen, P. J. Steinhardt, and M. S. Turner *Phys.Rev.* **D28** (1983) 679.

D. Wands, K. A. Malik, D. H. Lyth, and A. R. Liddle *Phys.Rev.* **D62** (2000) 043527, [`astro-ph/0003278`].

[32] D. H. Lyth and A. Riotto *Phys.Rept.* **314** (1999) 1–146, [`hep-ph/9807278`].

[33] H. Kodama and M. Sasaki *Prog.Theor.Phys.Suppl.* **78** (1984) 1–166.

[34] S. M. Carroll, W. H. Press, and E. L. Turner *Ann.Rev.Astron.Astrophys.* **30** (1992) 499–542.

[35] B. Harrison *Phys.Rev.* **D1** (1970) 2269–2271.

Y. Zel'dovich *Astron.Astrophys.* **5** (1970) 84–89.

[36] R. C. Nichol *Gen.Rel.Grav.* **40** (2008) 249–267, [`arXiv:0708.2824`].

[37] C. Blake, A. Collister, and O. Lahav *Mon.Not.Roy.Astron.Soc.* **385** (2008) 1257–1269, [`arXiv:0704.3377`].

[38] **SDSS** Collaboration, D. J. Eisenstein *et. al.* *Astron.J.* **122** (2001) 2267, [`astro-ph/0108153`].

D. A. Wake, R. K. Sheth, R. C. Nichol, C. M. Baugh, J. Bland-Hawthorn, *et. al.* *Mon.Not.Roy.Astron.Soc.* **387** (2008) 1045–1062, [`arXiv:0802.4288`].

[39] B. A. Reid, W. J. Percival, D. J. Eisenstein, L. Verde, D. N. Spergel, *et. al.* *Mon.Not.Roy.Astron.Soc.* **404** (2010) 60–85, [`arXiv:0907.1659`].

[40] **SDSS** Collaboration, B. A. Reid *et. al.* *Mon.Not.Roy.Astron.Soc.* **401** (2010) 2148–2168, [`arXiv:0907.1660`].

[41] F. S. Labini, N. L. Vasilyev, Y. V. Baryshev, and M. Lopez-Corredoira *Astron.Astrophys.* **505** (2009) 981–990, [`arXiv:0903.0950`].

[42] E. A. Kazin, M. R. Blanton, R. Scoccimarro, C. K. McBride, A. A. Berlind, *et. al.* *Astrophys.J.* **710** (2010) 1444–1461, [`arXiv:0908.2598`].

[43] D. Fixsen, E. Cheng, J. Gales, J. C. Mather, R. Shafer, *et. al.* *Astrophys.J.* **473** (1996) 576, [`astro-ph/9605054`].

[44] D. Fixsen *Astrophys.J.* **707** (2009) 916–920, [`arXiv:0911.1955`].

[45] W. Hu and S. Dodelson *Ann.Rev.Astron.Astrophys.* **40** (2002) 171–216, [`astro-ph/0110414`].

[46] R. Sachs and A. Wolfe *Astrophys.J.* **147** (1967) 73–90.

[47] **Particle Data Group** Collaboration, K. Nakamura *et. al.* *J.Phys.G* **G37** (2010) 075021.

[48] A. H. Guth *Phys. Rev.* **D23** (1981) 347–356.

[49] A. A. Starobinsky *JETP Lett.* **30** (1979) 682.

A. A. Starobinsky *Phys.Lett.* **B91** (1980) 99.

D. Kazanas *Astrophys.J.* **241** (1980) L59.

K. Sato *Mon.Not.Roy.Astron.Soc.* **195** (1981) 467.

[50] E. J. Copeland, E. W. Kolb, A. R. Liddle, and J. E. Lidsey *Phys. Rev.* **D48** (1993) 2529–2547, [`hep-ph/9303288`].

[51] A. R. Liddle and S. M. Leach *Phys.Rev.* **D68** (2003) 103503, [`astro-ph/0305263`].

[52] A. R. Liddle and D. H. Lyth *Phys.Rept.* **231** (1993) 1–105, [`astro-ph/9303019`]. Revised version.

[53] T. Bunch and P. Davies *Proc.Roy.Soc.Lond.* **A360** (1978) 117–134.

[54] J. R. Ellis, D. V. Nanopoulos, and S. Sarkar *Nucl.Phys.* **B259** (1985) 175.

M. Kawasaki, F. Takahashi, and T. Yanagida *Phys.Rev.* **D74** (2006) 043519, [`hep-ph/0605297`].

M. Kawasaki, F. Takahashi, and T. Yanagida *Phys.Lett.* **B638** (2006) 8–12, [`hep-ph/0603265`].

[55] A. D. Linde *Phys.Lett.* **B129** (1983) 177–181.

[56] M. Kawasaki, M. Yamaguchi, and T. Yanagida *Phys.Rev.Lett.* **85** (2000) 3572–3575, [`hep-ph/0004243`].

[57] A. Mazumdar and J. Rocher *Phys.Rept.* **497** (2011) 85–215, [`arXiv:1001.0993`].

[58] S. Hotchkiss, G. German, G. G. Ross, and S. Sarkar *JCAP* **0810** (2008) 015, [`arXiv:0804.2634`].

[59] D. S. Goldwirth and T. Piran *Phys.Rev.Lett.* **64** (1990) 2852–2855.

D. S. Goldwirth and T. Piran *Phys.Rept.* **214** (1992) 223–291.

E. Calzetta and M. Sakellariadou *Phys.Rev.* **D47** (1993) 3184–3193, [`gr-qc/9209007`].

T. Vachaspati and M. Trodden *Phys.Rev.* **D61** (1999) 023502, [`gr-qc/9811037`].

[60] G. Gibbons and N. Turok *Phys.Rev.* **D77** (2008) 063516, [`hep-th/0609095`].

N. Arkani-Hamed, S. Dubovsky, A. Nicolis, E. Trincherini, and G. Villadoro *JHEP* **0705** (2007) 055, [`arXiv:0704.1814`].

A. D. Linde *JCAP* **0706** (2007) 017, [`arXiv:0705.1160`].

A. Linde and M. Noorbala *JCAP* **1009** (2010) 008, [`arXiv:1006.2170`].

[61] S. Hotchkiss, A. Mazumdar, and S. Nadathur *JCAP* **1106** (2011) 002, [`arXiv:1101.6046`].

[62] K. Enqvist and A. Mazumdar *Phys.Rept.* **380** (2003) 99–234, [`hep-ph/0209244`].

[63] S. Nadathur and S. Sarkar *Phys.Rev.* **D83** (2011) 063506, [`arXiv:1012.3460`].

[64] S. Nadathur, S. Hotchkiss, and S. Sarkar `arXiv:1109.4126`.

[65] R. Allahverdi, K. Enqvist, J. Garcia-Bellido, and A. Mazumdar *Phys.Rev.Lett.* **97** (2006) 191304, [`hep-ph/0605035`].

[66] R. Allahverdi, K. Enqvist, J. Garcia-Bellido, A. Jokinen, and A. Mazumdar *JCAP* **0706** (2007) 019, [`hep-ph/0610134`].

[67] R. Allahverdi, A. Kusenko, and A. Mazumdar *JCAP* **0707** (2007) 018, [`hep-ph/0608138`].

[68] R. Allahverdi, B. Dutta, and A. Mazumdar *Phys.Rev.Lett.* **99** (2007) 261301, [`arXiv:0708.3983`].

[69] R. Allahverdi, B. Dutta, and A. Mazumdar *Phys.Rev.* **D75** (2007) 075018, [`hep-ph/0702112`].

[70] R. Allahverdi, B. Dutta, and Y. Santoso *Phys.Rev.* **D82** (2010) 035012, [`arXiv:1004.2741`].

[71] R. Allahverdi, A. Ferrantelli, J. Garcia-Bellido, and A. Mazumdar *Phys.Rev.* **D83** (2011) 123507, [`arXiv:1103.2123`].

[72] Z. Lalak and K. Turzynski *Phys.Lett.* **B659** (2008) 669–675, [`arXiv:0710.0613`].

K. Enqvist, L. Mether, and S. Nurmi *JCAP* **0711** (2007) 014, [`arXiv:0706.2355`].

[73] R. Allahverdi, A. R. Frey, and A. Mazumdar *Phys.Rev.* **D76** (2007) 026001, [`hep-th/0701233`].

[74] K. Enqvist, A. Mazumdar, and P. Stephens *JCAP* **1006** (2010) 020, [`arXiv:1004.3724`].

[75] A. D. Linde *Phys.Lett.* **B259** (1991) 38–47.

A. D. Linde *Phys.Rev.* **D49** (1994) 748–754, [`astro-ph/9307002`].

[76] E. J. Copeland, A. R. Liddle, D. H. Lyth, E. D. Stewart, and D. Wands *Phys.Rev.* **D49** (1994) 6410–6433, [`astro-ph/9401011`].

[77] C. Burgess, R. Easther, A. Mazumdar, D. F. Mota, and T. Multamaki *JHEP* **0505** (2005) 067, [`hep-th/0501125`].

[78] J. Bueno Sanchez, K. Dimopoulos, and D. H. Lyth *JCAP* **0701** (2007) 015, [`hep-ph/0608299`].

[79] R. Allahverdi and A. Mazumdar *Phys.Rev.* **D78** (2008) 043511, [`arXiv:0802.4430`].

[80] K. Enqvist, A. Mazumdar, and O. Taanila *JCAP* **1009** (2010) 030, [`arXiv:1007.0657`].

[81] R. Mohapatra, S. Antusch, K. Babu, G. Barenboim, M.-C. Chen, *et. al.* *Rept.Prog.Phys.* **70** (2007) 1757–1867, [`hep-ph/0510213`].

A. de Gouvea *Phys.Rev.* **D72** (2005) 033005, [`hep-ph/0501039`].

[82] R. Allahverdi, B. Dutta, and A. Mazumdar *Phys.Rev.* **D78** (2008) 063507, [`arXiv:0806.4557`].

K. Kamada and J. Yokoyama *Prog.Theor.Phys.* **122** (2010) 969–986, [`arXiv:0906.3402`].

[83] M. Maniatis *Int.J.Mod.Phys.* **A25** (2010) 3505–3602, [`arXiv:0906.0777`].

[84] H.-S. Lee, K. T. Matchev, and S. Nasri *Phys.Rev.* **D76** (2007) 041302, [`hep-ph/0702223`].

C. Arina and N. Fornengo *JHEP* **0711** (2007) 029, [`arXiv:0709.4477`].

T. Asaka, K. Ishiwata, and T. Moroi *Phys.Rev.* **D75** (2007) 065001, [`hep-ph/0612211`].

[85] R. Allahverdi and A. Mazumdar *JCAP* **0610** (2006) 008, [`hep-ph/0512227`].

[86] R. Allahverdi and A. Mazumdar *Phys.Rev.* **D76** (2007) 103526, [`hep-ph/0603244`].

[87] **WMAP** Collaboration, D. Spergel *et. al.* *Astrophys.J.Suppl.* **148** (2003) 175–194, [`astro-ph/0302209`].

[88] **WMAP** Collaboration, D. Spergel *et. al.* *Astrophys.J.Suppl.* **170** (2007) 377, [`astro-ph/0603449`].

[89] P. Hunt and S. Sarkar *Phys.Rev.* **D76** (2007) 123504, [`arXiv:0706.2443`].

[90] P. Hunt and S. Sarkar *Mon.Not.Roy.Astron.Soc.* **401** (2010) 547, [`arXiv:0807.4508`].

[91] A. Blanchard, M. Douspis, M. Rowan-Robinson, and S. Sarkar *Astron.Astrophys.* **412** (2003) 35–44, [`astro-ph/0304237`].

[92] S. Rasanen *Class.Quant.Grav.* **28** (2011) 164008, [`arXiv:1102.0408`].

[93] C. Clarkson, B. Bassett, and T. H.-C. Lu *Phys.Rev.Lett.* **101** (2008) 011301, [`arXiv:0712.3457`].

S. Rasanen *JCAP* **0902** (2009) 011, [`arXiv:0812.2872`].

T. Clifton and P. G. Ferreira *Phys.Rev.* **D80** (2009) 103503, [`arXiv:0907.4109`].

[94] V. Marra, E. W. Kolb, S. Matarrese, and A. Riotto *Phys.Rev.* **D76** (2007) 123004, [`arXiv:0708.3622`].

V. Marra, E. W. Kolb, and S. Matarrese *Phys.Rev.* **D77** (2008) 023003, [`arXiv:0710.5505`].

K. Kainulainen and V. Marra *Phys.Rev.* **D80** (2009) 127301, [`arXiv:0906.3871`].

[95] T. Biswas and A. Notari *JCAP* **0806** (2008) 021, [`astro-ph/0702555`].

R. Vanderveld, E. E. Flanagan, and I. Wasserman *Phys.Rev.* **D78** (2008) 083511, [`arXiv:0808.1080`].

W. Valkenburg *JCAP* **0906** (2009) 010, [`arXiv:0902.4698`].

E. E. Flanagan, N. Kumar, I. Wasserman, and R. Vanderveld `arXiv:1109.1873`.

[96] J. Moffat and D. Tatarski *Astrophys.J.* **452** (1995) 17, [`astro-ph/9407036`].

K. Tomita *Astrophys.J.* **529** (2000) 26, [`astro-ph/9905278`].

[97] M.-N. Celerier *Astron.Astrophys.* **353** (2000) 63–71, [`astro-ph/9907206`].

[98] H. Alnes, M. Amarzguioui, and O. Gron *Phys.Rev.* **D73** (2006) 083519, [`astro-ph/0512006`].

[99] T. Biswas, R. Mansouri, and A. Notari *JCAP* **0712** (2007) 017, [`astro-ph/0606703`].

H. Alnes and M. Amarzguioui *Phys.Rev.* **D75** (2007) 023506, [`astro-ph/0610331`].

S. Alexander, T. Biswas, A. Notari, and D. Vaid *JCAP* **0909** (2009) 025, [`arXiv:0712.0370`].

[100] T. Clifton, P. G. Ferreira, and K. Land *Phys.Rev.Lett.* **101** (2008) 131302, [`arXiv:0807.1443`].

[101] K. Bolejko and J. B. Wyithe *JCAP* **0902** (2009) 020, [`arXiv:0807.2891`].

[102] J. Garcia-Bellido and T. Haugboelle *JCAP* **0804** (2008) 003, [`arXiv:0802.1523`].

[103] J. Garcia-Bellido and T. Haugboelle *JCAP* **0809** (2008) 016, [`arXiv:0807.1326`].

[104] J. Zibin, A. Moss, and D. Scott *Phys.Rev.Lett.* **101** (2008) 251303, [`arXiv:0809.3761`].

[105] A. Moss, J. P. Zibin, and D. Scott *Phys.Rev.* **D83** (2011) 103515, [`arXiv:1007.3725`].

[106] T. Clifton, P. G. Ferreira, and J. Zuntz *JCAP* **0907** (2009) 029, [`arXiv:0902.1313`].

[107] T. Biswas, A. Notari, and W. Valkenburg *JCAP* **1011** (2010) 030, [`arXiv:1007.3065`].

[108] K. Enqvist and T. Mattsson *JCAP* **0702** (2007) 019, [`astro-ph/0609120`].

[109] J. P. Zibin *Phys.Rev.* **D78** (2008) 043504, [`arXiv:0804.1787`].

[110] J. Martin and C. Ringeval *Phys.Rev.* **D69** (2004) 083515, [`astro-ph/0310382`].

N. Kogo, M. Sasaki, and J. Yokoyama *Phys.Rev.* **D70** (2004) 103001, [`astro-ph/0409052`].

N. Kogo, M. Sasaki, and J. Yokoyama *Prog.Theor.Phys.* **114** (2005) 555–572, [`astro-ph/0504471`].

A. Shafieloo, T. Souradeep, P. Manimaran, P. K. Panigrahi, and R. Rangarajan *Phys.Rev.* **D75** (2007) 123502, [`astro-ph/0611352`].

A. Shafieloo and T. Souradeep *Phys.Rev.* **D70** (2004) 043523, [`astro-ph/0312174`].

D. Tocchini-Valentini, Y. Hoffman, and J. Silk *Mon.Not.Roy.Astron.Soc.* **367** (2006) 1095–1102, [`astro-ph/0509478`].

D. Tocchini-Valentini, M. Douspis, and J. Silk *Mon.Not.Roy.Astron.Soc.* **359** (2005) 31–35, [`astro-ph/0402583`].

G. Nicholson, C. R. Contaldi, and P. Paykari *JCAP* **1001** (2010) 016, [`arXiv:0909.5092`].

G. Nicholson and C. R. Contaldi *JCAP* **0907** (2009) 011, [`arXiv:0903.1106`].

K. Ichiki and R. Nagata *Phys.Rev.* **D80** (2009) 083002.

[111] J. A. Adams, G. G. Ross, and S. Sarkar *Nucl.Phys.* **B503** (1997) 405–425, [`hep-ph/9704286`].

[112] T. Gherghetta, C. F. Kolda, and S. P. Martin *Nucl.Phys.* **B468** (1996) 37–58, [`hep-ph/9510370`].

[113] P. Hunt and S. Sarkar *Phys.Rev.* **D70** (2004) 103518, [`astro-ph/0408138`].

[114] J. A. Adams, B. Cresswell, and R. Easther *Phys.Rev.* **D64** (2001) 123514, [`astro-ph/0102236`].

[115] **WMAP** Collaboration, H. Peiris *et. al.* *Astrophys.J.Suppl.* **148** (2003) 213, [`astro-ph/0302225`].

L. Covi, J. Hamann, A. Melchiorri, A. Slosar, and I. Sorbera *Phys.Rev.* **D74** (2006) 083509, [`astro-ph/0606452`].

[116] J. Hamann, L. Covi, A. Melchiorri, and A. Slosar *Phys.Rev.* **D76** (2007) 023503, [`astro-ph/0701380`].

[117] S. Hotchkiss and S. Sarkar *JCAP* **1005** (2010) 024, [`arXiv:0910.3373`].

[118] P. Brax and E. Cluzel *JCAP* **1104** (2011) 014, [`arXiv:1102.1917`].

[119] See review by B. Fields and S. Sarkar in [47].

[120] G. Tammann, A. Sandage, and B. Reindl *Astrophys.J.* **679** (2008) 52–70, [`arXiv:0712.2346`].

[121] A. G. Riess, L. Macri, S. Casertano, M. Sosey, H. Lampeitl, *et. al.* *Astrophys.J.* **699** (2009) 539–563, [`arXiv:0905.0695`].

[122] M. Regis and C. Clarkson `arXiv:1003.1043`.

[123] C. Clarkson and M. Regis *JCAP* **1102** (2011) 013, [`arXiv:1007.3443`].

[124] S. February, J. Larena, M. Smith, and C. Clarkson *Mon.Not.Roy.Astron.Soc.* **405** (2010) 2231, [`arXiv:0909.1479`].

[125] D. J. Eisenstein and W. Hu *Astrophys.J.* **496** (1998) 605, [`astro-ph/9709112`].

[126] E. Gaztanaga, A. Cabre, and L. Hui *Mon.Not.Roy.Astron.Soc.* **399** (2009) 1663–1680, [`arXiv:0807.3551`].

E. Gaztanaga, R. Miquel, and E. Sanchez *Phys.Rev.Lett.* **103** (2009) 091302, [`arXiv:0808.1921`].

[127] E. A. Kazin, M. R. Blanton, R. Scoccimarro, C. K. McBride, and A. A. Berlind *Astrophys.J.* **719** (2010) 1032–1044, [`arXiv:1004.2244`].

J. Miralda-Escude `arXiv:0901.1219`.

[128] A. Lewis and S. Bridle *Phys.Rev.* **D66** (2002) 103511, [`astro-ph/0205436`].

[129] F. Beutler, C. Blake, M. Colless, D. Jones, L. Staveley-Smith, *et. al.* `arXiv:1106.3366`.

[130] J.-P. Uzan, C. Clarkson, and G. F. Ellis *Phys.Rev.Lett.* **100** (2008) 191303, [`arXiv:0801.0068`].

[131] M. Quartin and L. Amendola *Phys.Rev.* **D81** (2010) 043522, [`arXiv:0909.4954`].

C.-M. Yoo, T. Kai, and K.-i. Nakao *Phys.Rev.* **D83** (2011) 043527, [`arXiv:1010.0091`].

[132] J. Goodman *Phys.Rev.* **D52** (1995) 1821–1827, [`astro-ph/9506068`].

R. Caldwell and A. Stebbins *Phys.Rev.Lett.* **100** (2008) 191302, [`arXiv:0711.3459`].

[133] R. Sunyaev and Y. Zeldovich *Comments Astrophys. Space Phys.* **4** (1972) 173–178.

[134] C.-M. Yoo, K.-i. Nakao, and M. Sasaki *JCAP* **1010** (2010) 011, [`arXiv:1008.0469`].

[135] P. Bull, T. Clifton, and P. G. Ferreira `arXiv:1108.2222`.

[136] S. Das, T. A. Marriage, P. A. Ade, P. Aguirre, M. Amir, *et. al.* *Astrophys.J.* **729** (2011) 62, [`arXiv:1009.0847`].

[137] P. Zhang and A. Stebbins *Phys.Rev.Lett.* **107** (2011) 041301, [`arXiv:1009.3967`].

[138] J. P. Zibin and A. Moss *Class.Quant.Grav.* **28** (2011) 164005, [`arXiv:1105.0909`].

[139] R. G. Crittenden and N. Turok *Phys.Rev.Lett.* **76** (1996) 575, [`astro-ph/9510072`].

[140] N. Afshordi *Phys.Rev.* **D70** (2004) 083536, [`astro-ph/0401166`].

M. Douspis, P. G. Castro, C. Caprini, and N. Aghanim *Astron.Astrophys.* **485** (2008) 395, [`arXiv:0802.0983`].

[141] P. Fosalba, E. Gaztanaga, and F. Castander *Astrophys.J.* **597** (2003) L89–92, [`astro-ph/0307249`].

N. Afshordi, Y.-S. Loh, and M. A. Strauss *Phys.Rev.* **D69** (2004) 083524, [`astro-ph/0308260`].

S. Boughn and R. Crittenden *Nature* **427** (2004) 45–47, [`astro-ph/0305001`].

WMAP Collaboration, M. R. Nolta *et. al.* *Astrophys.J.* **608** (2004) 10–15, [`astro-ph/0305097`].

[142] N. Padmanabhan, C. M. Hirata, U. Seljak, D. Schlegel, J. Brinkmann, *et. al.* *Phys.Rev.* **D72** (2005) 043525, [`astro-ph/0410360`].

[143] T. Giannantonio, R. G. Crittenden, R. C. Nichol, R. Scranton, G. T. Richards, *et. al.* *Phys.Rev.* **D74** (2006) 063520, [`astro-ph/0607572`].

A. Cabre, E. Gaztanaga, M. Manera, P. Fosalba, and F. Castander *Mon.Not.Roy.Astron.Soc.* **372** (2006) L23–L27, [`astro-ph/0603690`].

A. Raccanelli, A. Bonaldi, M. Negrello, S. Matarrese, G. Tormen, *et. al.* *Mon.Not.Roy.Astron.Soc.* **386** (2008) 2161, [`arXiv:0802.0084`].

[144] A. Rassat, K. Land, O. Lahav, and F. B. Abdalla *Mon.Not.Roy.Astron.Soc.* **377** (2007) 1085–1094, [`astro-ph/0610911`].

C. Francis and J. Peacock *Mon.Not.Roy.Astron.Soc.* **406** (2010) 2, [`arXiv:0909.2494`].

[145] U. Sawangwit, T. Shanks, R. Cannon, S. Croom, N. P. Ross, *et. al.* *Mon.Not.Roy.Astron.Soc.* **402** (2010) 2228, [`arXiv:0911.1352`].

[146] S. Ho, C. Hirata, N. Padmanabhan, U. Seljak, and N. Bahcall *Phys.Rev.* **D78** (2008) 043519, [`arXiv:0801.0642`].

T. Giannantonio, R. Scranton, R. G. Crittenden, R. C. Nichol, S. P. Boughn, *et. al.* *Phys.Rev.* **D77** (2008) 123520, [`arXiv:0801.4380`].

[147] M. Lopez-Corredoira, F. Labini, and J. Betancort-Rijo *Astron.Astrophys.* **513** (2010) A3, [`arXiv:1001.4000`].

[148] B. R. Granett, M. C. Neyrinck, and I. Szapudi *Astrophys.J.* **683** (2008) L99, [`arXiv:0805.3695`].

[149] B. R. Granett, M. C. Neyrinck, and I. Szapudi `arXiv:0805.2974`.

[150] K. T. Inoue, N. Sakai, and K. Tomita *Astrophys.J.* **724** (2010) 12–25, [`arXiv:1005.4250`].

[151] P. Papai and I. Szapudi *Astrophys.J.* **725** (2010) 2078–2086, [`arXiv:1009.0754`].

[152] J. M. Bardeen, J. Bond, N. Kaiser, and A. Szalay *Astrophys.J.* **304** (1986) 15–61.

[153] Y.-C. Cai, S. Cole, A. Jenkins, and C. S. Frenk *Mon.Not.Roy.Astron.Soc.* **407** (2010) 201, [`arXiv:1003.0974`].

[154] V. Springel, S. D. White, A. Jenkins, C. S. Frenk, N. Yoshida, *et. al.* *Nature* **435** (2005) 629–636, [`astro-ph/0504097`].

[155] R. K. Sheth and R. van de Weygaert *Mon.Not.Roy.Astron.Soc.* **350** (2004) 517, [`astro-ph/0311260`].

[156] P. Papai, I. Szapudi, and B. R. Granett *Astrophys.J.* **732** (2011) 27, [`arXiv:1012.3750`].

[157] O. Lahav and P. B. Lilje *Astrophys.J.* **374** (1991) 29–43.

[158] **SDSS** Collaboration, J. K. Adelman-McCarthy *et. al.* *Astrophys.J.Suppl.* **175** (2008) 297–313, [`arXiv:0707.3413`].

[159] M. C. Neyrinck *Mon.Not.Roy.Astron.Soc.* **386** (2008) 2101, [`arXiv:0712.3049`].

[160] M. C. Neyrinck, N. Y. Gnedin, and A. J. Hamilton *Mon.Not.Roy.Astron.Soc.* **356** (2005) 1222, [`astro-ph/0402346`].

[161] **WMAP** Collaboration, G. Hinshaw *et. al.* *Astrophys.J.Suppl.* **180** (2009) 225–245, [`arXiv:0803.0732`].

[162] **SDSS** Collaboration, K. N. Abazajian *et. al.* *Astrophys.J.Suppl.* **182** (2009) 543–558, [`arXiv:0812.0649`].

[163] J. M. Colberg, R. K. Sheth, A. Diaferio, L. Gao, and N. Yoshida *Mon.Not.Roy.Astron.Soc.* **360** (2005) 216–226, [`astro-ph/0409162`].

[164] S. Matarrese, L. Verde, and R. Jimenez *Astrophys.J.* **541** (2000) 10, [`astro-ph/0001366`].

M. Kamionkowski, L. Verde, and R. Jimenez *JCAP* **0901** (2009) 010, [`arXiv:0809.0506`].

[165] K. Enqvist, S. Hotchkiss, and O. Taanila *JCAP* **1104** (2011) 017, [`arXiv:1012.2732`].

[166] R. Nagata, T. Chiba, and N. Sugiyama *Phys.Rev.* **D69** (2004) 083512, [`astro-ph/0311274`].