University of Cambridge
Institute of Astronomy

A dissertation submitted to the University of Cambridge for the degree of Doctor of Philosophy

∽

Constraining Fundamental Physics From Cosmology

∽

Simeon Paul Bird
Pembroke College

Submitted to the Board of Graduate Studies
July 13, 2011

Under the Supervision of
Dr. Hiranya V. Peiris
Dr. Martin Haehnelt

DECLARATION OF ORIGINALITY

I, Simeon Paul Bird, declare that this thesis titled, 'Constraining Fundamental Physics From Cosmology' and the work presented in it are my own. I confirm that:

- This work was done wholly while in candidature for a research degree at this University.

- This thesis has not previously been submitted for a degree or any other qualification at this University or any other institution.

- Where I have consulted the published work of others, this is always clearly attributed, especially in Chapter 1 which serves as an introduction and largely draws on work in the literature.

- Where I have quoted from the work of others, the source is always given. With the exception of quotations or figures who are credited to others, this thesis is entirely my own work.

- I have acknowledged all main sources of help, in the acknowledgements and first chapter.

- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself. All writing has been put together by myself and I have asked others (credited in the introduction) for their aid in editing and revision.

- The length of this thesis does not exceed the stated limit of the Degree Committee of Physics and Chemistry of 60,000 words.

_____

Signed

_____

Date

Simeon Bird: Constraining Fundamental Physics From Cosmology

Summary

I use mathematical models and numerical simulations to constrain cosmological inflation, the seeds of structure, and the mass of the neutrino. I revisit arguments that simple models of inflation with a small red tilt in the scalar power spectrum generically yield an observable tensor spectrum. I show that criteria for fine-tuning based upon the algebraic simplicity of the potential depend strongly upon the assumptions they incorporate about the potential. Furthermore, several models with algebraically simple potentials require carefully tuned initial field configurations. I demonstrate the existence of potentials with vanishingly small tensor amplitudes which are natural in terms of both their algebraic form and initial conditions. I thus argue that proposed experiments which make highly sensitive measurements of the tensor amplitude cannot definitively rule out the inflationary paradigm.

The overshoot problem is the need to make the initial kinetic energy of the inflaton small enough to ensure slow-roll. I investigate claims that brane inflation solves the overshoot problem through microphysical restrictions on the phase space of initial conditions. By carrying out a comprehensive analysis of the parameter space allowed by the latest advances in brane inflation model-building, I find that the vast majority of the phase space of initial conditions is still dominated by overshoot trajectories.

Current results from the Lyman-α forest assume that the primordial power spectrum of density perturbations follows a simple power-law form with running. I perform a large suite of numerical simulations, using them to calibrate a minimally parametric framework for describing the power spectrum. Combined with cross-validation this framework allows me to directly reconstruct the power spectrum shape, a consistency check on the standard model. I find no evidence for deviation from scale-invariance, but current Lyman-α data do not have sufficient statistical power to robustly probe the shape of the power spectrum at these scales. In contrast, the ongoing Baryon Oscillation Sky Survey will be able to do so with high precision.

I perform an extensive suite of N-body simulations of the matter power spectrum, probing deep into the non-linear regime while incorporating massive neutrinos. I compare my results to the widely used HALOFIT approximation, and find that in the strongly non-linear regime it significantly over-predicts the suppression due to the free-streaming of neutrinos. Most published constraints are not affected, as they have used HALOFIT only in the linear or mildly non-linear regime. However, my results are important for future galaxy and weak lensing surveys.

Acknowledgements

Contents

"There are more things in heaven and earth, Horatio,
than are dreamt of in your philosophy."

– Hamlet, by William Shakespeare

1

Introduction

1.1 The Universe

Cosmology is the study of the Universe. It seeks to understand what the Universe is made of, and how it has evolved. Over the long history of science, few subjects have excited such wide speculation. However, in recent decades mankind is for the first time in the fortunate position of having large quantities of reliable data with which to turn speculation into testable hypothesis.

1.1.1 Composition

Dark Energy

The Universe is flat, with an energy density near the critical density (Melchiorri et al., 2000; Spergel et al., 2003). However, only thirty percent of this is matter; the other seventy percent of the Universe is dark energy. Early hints of dark energy were provided by simulations of structure formation (Efstathiou et al., 1990), which suggested that the matter density was only a third of the critical density. Later measurements of the supernovae distance-luminosity relation showed that the Universe is undergoing a period of accelerated expansion (Riess et al., 1998), implying the existence of either matter with negative pressure or a cosmological constant. Recently Blake et al. (2011) used measurements of the acoustic scale of oscillations in the primordial plasma as a standard ruler, and obtained independent confirmation of the acceleration of the Universe. Current data show no significant deviation from a cosmological constant, but are not sufficient to rule out alternatives such as exotic matter species.

Dark Matter

Dark matter describes the non-baryonic weakly interacting matter which makes up eighty percent of the matter in the Universe. It was first proposed by Zwicky (1933), who

obtained tentative evidence for it from estimates of the mass of the Coma cluster, and then rediscovered by Rubin and Ford (1970), whose observations of the rotation curve of the Andromeda galaxy implied the presence of a non-luminous galactic halo. A purely gravitational explanation was largely ruled out by Markevitch et al. (2004), who found the centre of mass for luminous matter in the Bullet cluster was offset from the total centre of mass. Further evidence appears from many other experiments, for example, the cosmic microwave background (Spergel et al., 2003) and galaxy surveys (Percival et al., 2002).

Massive active neutrino species were an early candidate for dark matter. However, because they retain significant thermal velocities from the early Universe, when they were relativistic, they cannot make up the majority of the dark matter; these residual thermal velocities would severely damp structure formation on small scales, to an extent which is incompatible with observations (Bond et al., 1980). Since neutrinos are known to be massive (Becker-Szendy et al., 1992), they do make up some small fraction of the dark matter, and I discuss using this fact to constrain their mass further in Section 1.6.

Baryons

The Universe manifestly also contains baryonic matter. Examination of absorption in high-redshift quasar spectra reveals that most of the baryons in the Universe are in the diffuse gas of hydrogen and helium making up the Intergalactic Medium (see, e.g. Rauch (1998) for a review). Absorption from residual hydrogen at $z \sim 3$ leads to the complex structure of overlapping hydrogen absorption called the Lyman-α forest, discussed further in Section 1.7. Detecting the Intergalactic Medium at low redshift is more challenging; here both hydrogen and helium are fully ionised, and thus do not cause significant absorption signals. Finally, trace amounts of radiation are present, mostly in the form of microwave background photons, discussed in Section 1.4.

1.1.2   History

Although accelerated expansion was only discovered recently, the expansion of the Universe has been known for far longer. Friedman (1922) and Lemaître (1927) proposed a cosmology with a uniform expansion of space, which was confirmed when Hubble (1929) observed a linear relation between distance and redshift for nearby galaxies. I discuss the Friedman-Lemaître model further in Section 1.2; one consequence is that in the distant past the Universe was much hotter and denser (Lemaître, 1931) and governed by progressively higher energy physics.

The concordance cosmology begins with inflation (Starobinsky, 1979; Guth, 1981), a period of accelerated exponential expansion in the very early Universe. The motivation for inflation, described further in Section 1.3, is to explain dynamically the initial conditions of the Universe. It occurs when the Universe is dominated by a form of matter with negative pressure (Linde, 1983), called the inflaton, whose properties I discuss further in Section 2.1. Chapter 3 presents work regarding some observational consequences of inflation.

Once inflation finishes, the inflaton is presumed to have decayed into radiation and matter, marking the start of the radiation era. The Universe is at this time a hot plasma,

gradually cooling and expanding. Once it cools to a temperature of 10^9 K, the neutron decay reaction ceases to be in equilibrium, leading to the decoupling of the neutrino background (Section 1.6). Free neutrons start to decay and combine with protons to form helium and lithium nuclei (Alpher et al., 1948). Radiation density redshifts faster than matter, and so matter soon starts to dominate, when $T \sim 10^4$ K.

When $T \sim 3000$ K, free electrons combine with protons, marking the end of the primordial plasma. The majority of photons emitted at this time are not scattered again, but are redshifted by the expansion of the Universe, forming the cosmic microwave background (CMB), which I discuss further in Section 1.4.

The CMB was emitted about 3.8×10^5 years after the end of inflation, and the Universe is now almost 1.4×10^{10} years old (see, e.g. Komatsu et al. (2011) for a compilation of the latest CMB and other results). The dominant physical process during the intervening time has been the collapse of matter under gravity to form structure. Section 1.5 describes an analytical approach to modelling this process, while Section 2.3 discusses the corresponding numerical methods. In Chapter 5, I present work using the Lyman-α forest, a system of absorption features in quasar spectra explained further in Section 1.7, to constrain the seeds of structure.

In the rest of this Chapter, I describe in more detail some of the physics alluded to above. Section 1.2 discusses the expanding Universe, Section 1.3 the motivation for inflation, and Section 1.4 the CMB. In Section 1.5 I describe the growth of structure, in Section 1.6 massive neutrinos, and in Section 1.7 the Lyman-α forest. I return to the technical background of inflation and the simulation of cosmic structure in Chapter 2, as these are the main topics of this thesis, and the subjects of Chapters 3-6.

1.2 Friedmann-Robertson-Walker Cosmology

In this Section I elaborate on the models of an expanding Universe discussed above. These models assume the cosmological principle, which embodies the idea that there should be nothing special about our position in the Universe, nor any preferred directions in which to look. Observational evidence is provided by the CMB, which is homogeneous and isotropic on large scales (Mather et al., 1990).

Solving the Einstein equations under the assumption of homogeneity and isotropy leads to the Friedmann-Robertson-Walker (FRW) metrics (Friedman, 1922), with the equation

$$ds^2 = -dt^2 + a^2(t) \left(dr^2 + f^2(r)(d\theta^2 + \sin^2\theta d\phi^2) \right) . \tag{1.1}$$

Here I use polar coordinates, units such that $c = 1$, and define a dimensionless expansion factor $a(t)$, which is by convention normalised to unity at the present day. $f(r)$ is a function dependent on curvature. Measurements of the microwave background acoustic peak combined with data on the Hubble constant imply that our Universe is spatially flat (Spergel et al., 2003), so $f(r) = r$.

The Friedmann equations can be derived from the Einstein equations for this metric

and are

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho + \frac{\Lambda}{3} - \frac{k}{a^2} \tag{1.2}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(3p + \rho) + \frac{\Lambda}{3}\,. \tag{1.3}$$

G is Newton's gravitational constant, Λ is a cosmological constant and k is spatial curvature. A dot denotes a derivative with respect to cosmic time. I assume a perfect fluid with density ρ and pressure p. Conservation of mass energy implies that

$$\dot{\rho} = -3\frac{\dot{a}}{a}(p + \rho)\,. \tag{1.4}$$

Thus, the assumption of an equation of state for the matter species fixes the phenomenological evolution of the Universe. Dark matter is pressureless, with $p = 0$, while relativistic species, such as radiation, have $p = \frac{\rho}{3}$. I can also simplify the above equations by defining the energy density in curvature and the cosmological constant (or dark energy) as

$$\rho_k = \frac{3k}{8\pi G a^2}\,, \qquad\qquad p_k = -\frac{\rho_k}{3}\,, \tag{1.5}$$

$$\rho_\Lambda = \frac{\Lambda}{8\pi G}\,, \qquad\qquad p_\Lambda = -\rho_\Lambda\,. \tag{1.6}$$

The Hubble parameter is defined as $H = \dot{a}/a$, and the critical density by

$$\rho_{\rm c} = \frac{3H^2}{8\pi G}\,. \tag{1.7}$$

The critical density can be used to define $\Omega_{\rm s} = \rho_{\rm s}/\rho_{\rm c}$, which has the advantage that the sum over all species $\Omega_{\rm total} = 1$.

The above equations show that matter density, $\rho_{\rm m} \sim a^{-3}$, dark energy density $\rho_\Lambda \sim a^0$, and radiation density $\rho_{\rm r} \sim a^{-4}$. In our Universe all three species are present, and there is a radiation-dominated phase, a matter-dominated phase, and finally a period of accelerated expansion.

1.3 Inflation

The CMB is observed to be a nearly perfect black body (Mather et al., 1990); thus at the time of last scattering, the entire sky was in thermal equilibrium. This seemingly innocuous fact demands further explanation, due to the scales involved. If we assume a matter-radiation dominated Universe like that described in the last Section, the average density of the Universe at early times ultimately reaches the Planck scale. This is the scale on which quantum corrections to general relativity are presumed to become important, requiring a theory of quantum gravity. The cosmological horizon is defined by the distance a signal at the speed of light can travel since this phase; essentially the maximal radius over which matter in the Universe can be in causal contact. If the Universe had always contained only matter and radiation, the horizon at last scattering would subtend less than a degree on the present-day sky. However, the CMB proves that at last scattering

the entire sky is in thermal equilibrium, and thus it seems it was in causal contact. This puzzle is the horizon problem.

There are a few potential solutions; perhaps some as-yet undiscovered physical principle demands that the initial conditions of the Universe are at a certain temperature, possibly due to unknown physics at the Planck scale. Alternatively, the expansion history of the Universe may be modified so as to dynamically make the effective horizon at large scattering larger. Inflation realises the second idea through a period of exponential expansion, which mimics domination by a cosmological constant (Starobinsky, 1979; Guth, 1981; Sato, 1981; Mukhanov and Chibisov, 1981; Linde, 1982b; Albrecht and Steinhardt, 1982).

A Universe dominated by a cosmological constant would have no means of decaying into the observed matter dominated Universe, nor would it be able to explain the origins of structure. Inflation instead postulates domination by a scalar field, the inflaton. The vacuum energy of the inflaton mimics a cosmological constant, but gradually decreases. Ultimately the energy of the inflaton is unable to sustain inflation, and the Universe transitions into a radiation dominated phase, while the inflaton may decay into more familiar particle species. The equations describing the dynamics of the inflaton will be presented in Section 2.1. The result is that the expansion factor during inflation is

$$ a(t) = a_0 e^{Ht} \,, \tag{1.8} $$

with a slowly-decreasing Hubble parameter. This exponential expansion allows a near Planck-scale homogeneous and isotropic patch, small enough to be in causal contact, to be stretched to cover the entire observable Universe. In effect, inflation solves the horizon problem by dynamically stretching the horizon.

Inflation can also provide a dynamical mechanism to explain the observed flatness of the Universe. Matter density redshifts as $\Omega_{\mathrm{M}} \sim a^{-3}$, while curvature redshifts as $\Omega_k \sim a^{-2}$, so during the matter era the ratio of Ω_k to Ω_{M} increases. There is a similar result for the radiation era, the end result of which is that satisfying the present day constraints on Ω_k requires an exceedingly small amount of initial curvature.

It is possible that curvature could be forbidden by an as-yet undiscovered symmetry principle. However, inflation solves the flatness problem dynamically; an exponential increase in the scale factor will exponentially suppress curvature. An adequate solution to the horizon problem requires ~ 60 e-folds of inflation ($Ht \sim 60$), so by the end of inflation curvature will be suppressed to at most $\Omega_k \sim e^{-120}$, which is far too small to be observed. The Universe may still be globally closed or open, but on scales vastly greater than the horizon.

The final, and possibly most compelling, motivation for inflation is that it provides an explanation for the origin of structure, while simultaneously solving the above problems. The earliest tracers of inhomogeneous structure so far observed are tiny temperature fluctuations in the CMB black body spectrum (Smoot et al., 1992; Bennett et al., 2003b), whose amplitude and distribution underpin our understanding of much of the later Universe. Inflation provides a framework for producing these fluctuations. According to inflation, structure originates as quantum fluctuations in the value of the inflaton field. Exponential expansion boosts these quantum effects to large scales, freezing their amplitude as fluctuation modes leave the horizon and allowing them to serve as the seeds of

structure. The details of the quantum field theory calculation which yields these results are beyond the scope of this thesis, but are described in, e.g., Mukhanov et al. (1992).

While inflation solves the flatness and horizon problems, as well as generating the seeds of structure, there is considerable uncertainty about its observational predictions. This is due to several factors, mainly related to our limited knowledge of particle physics at these energies. To make predictions for the amplitude of perturbations, and for their power spectrum, it is necessary to know the potential for the inflaton, discussed further in Section 2.1. Our incomplete understanding of particle physics at these scales also comes into play at the end of inflation. I mentioned above that the inflaton should decay into matter and radiation, but it is not known exactly how this might take place. These issues could be resolved if it were possible to embed inflation within a complete high-energy theory of particle physics and quantum gravity. The leading contender for this is string theory (Green et al., 1988). Although it is far from being experimentally proven, string theory allows us to examine fully realised case studies of inflationary models, and thus at least gain understanding of the challenges faced by them. In Chapter 4, I discuss the issue of overshoot in the context of brane inflation, the string theory inflation model which has received most attention.

1.4 The Cosmic Microwave Background

As discussed in Section 1.1, the cosmic microwave background (CMB) is composed of relic radiation, emitted when the temperature of the Universe was around 3000 K. At this time, free protons and electrons combined, causing the mean free path of a photon to change from extremely short to greater than the Hubble length. These photons were then largely decoupled from the rest of the Universe, and can now be observed in the microwave, allowing us a window into this early era.

The CMB was first observed accidentally by Penzias and Wilson (1965) as an unexpected microwave source, and initially suspected to be caused by pigeon droppings in the antenna. It quickly became clear that the new source of radiation was extragalactic, and the observational history of the CMB began.

Early work focussed on measuring the CMB temperature spectrum, which the COBE satellite (Mather et al., 1990) showed to be a nearly perfect black body at 2.7 K, showing that the Universe was homogeneous and isotropic on the largest scales, and detected small fluctuations in the temperature from a perfect black body (Smoot et al., 1992), at an amplitude of about 10^{-6} the total emission. Later experiments, such as the WMAP satellite (Bennett et al., 2003a), have mapped these fluctuations, finding their power spectrum and using it to determine the parameters of the concordance cosmology. We currently await results from the Planck satellite (Planck Collaboration, 2005), which aims to extract essentially all available information from the primary CMB temperature.

Future experiments will hopefully also make precision measurements of primordial polarization in the CMB photons (Bond and Efstathiou, 1984; Kamionkowski et al., 1997). This is more challenging; while the energy of the CMB is only weakly affected by intervening matter, the polarization is severely masked by foreground emission of polarized Galactic photons at all frequencies. However, CMB polarization is the current best hope for detecting the theorised primordial gravitational wave background, which

would give a second complementary probe into inflation (Starobinsky, 1979; Baumann et al., 2009b). In Chapter 3 I discuss the predictions simple inflationary models make for the amplitude of this background, an important piece of information for people building polarization experiments.

Great efforts have been expended in pursuit of ever more accurate data on the CMB, efforts justified by the rich harvest of theoretical insight it produces. Part of the power of the CMB lies in the fact that it provides a clear window into a period when deviations from homogeneity can be described by linear perturbation theory (Bond et al., 1994; Seljak and Zaldarriaga, 1996).

1.5 The Growth of Structure

As mentioned in the previous Section, observations of the CMB imply that the early Universe was homogeneous and isotropic on large scales, with small perturbations probably laid down during an early phase of inflation. As time passed, structure developed out of the gravitational collapse of these perturbations. In this Section I discuss some of the techniques used to describe this process.

Linear perturbation theory (Peebles, 1980) can be used when perturbations are small; if there is a species with a homogeneous background density $\rho(t)$ a dimensionless inhomogeneous perturbation from the background is defined such that $\delta(\mathbf{x}, t) \ll 1$. Two species are important at this time in the real Universe; cold dark matter and radiation, so I define δ_c for the matter and δ_r for the radiation. With this, I can approximate the Einstein equations as coupled second order linear differential equations, accurate to first order in the perturbations. I do not give the equations or their derivation here; instead I talk about the rate at which the solutions grow with time.

Structure growth is sensitive to two factors. First, the rate at which the background expands, meaning structure depends on which species dominates the Universe. Second, because perturbations on larger scales than the cosmological horizon cannot self-interact, they may grow at a different rate from those on sub-horizon scales.

The cold dark matter perturbation δ_c grows on all scales during the matter era like

$$\delta_c(k, t) \sim t^{2/3} \sim a(t) \,. \tag{1.9}$$

This is independent of the horizon scale because δ_c is a perturbation in pressure-less matter.

The radiation perturbation δ_r behaves the same in all eras, but is affected by the scale considered, since it possesses relativistic pressure and self-interaction. On super-horizon scales δ_r is always proportional to the matter perturbation; a consequence of entropy conservation in the Einstein equations. On sub-horizon scales, the radiation pressure becomes important and the perturbation amplitude oscillates, with no lasting structure growth.

Now let me turn to the behaviour of the matter perturbation during the radiation era. Since radiation is the dominant component in the Universe at the time, the coupling of matter to radiation becomes important on sub-horizon scales. The oscillations in radiation suppress the growth of the matter perturbation, so that $\delta_c \sim \ln t$. On super-horizon scales, however, this coupling is not effective. Because the expansion of the Universe is

slower during the radiation era, the matter perturbation grows somewhat faster, $\delta_c \sim t$. Finally, during late-time exponential expansion the growth of all perturbations is suppressed, and $\delta_c \sim$ constant.

δ_c grows efficiently during the matter era, implying that the linear description, which required $\delta_c \ll 1$, will eventually break down. To probe this regime, descriptions of non-linear structure growth are needed. The most accurate way to do this is to use numerical simulations, with methods described in Section 2.3. However, approximate analytic solutions can be very useful tools for some calculations, particularly when simulations are too time-consuming to run. One such approximate solution, used in Chapter 6, is HALOFIT (Smith et al., 2003).

This rests on the idea that once a perturbation grows to a certain level, it collapses into a virialized object, a halo. The halo has an internal structure independent of the internal distribution of the pre-collapse material. This model effectively reduces the non-linear problem of structure formation to the simpler one of finding approximate functional solutions for the internal structure of a generic halo, and, on larger scales, for correlations between halos. HALOFIT uses simulations to calibrate these functions, unlike older methods based on spherical collapse or stable clustering (Peebles, 1974; Hamilton et al., 1991; Peacock and Dodds, 1996).

1.6 Massive Neutrinos

Neutrinos are a family of neutral leptons, commonly produced in β-decay. Less well-known is the existence of a ubiquitous cosmic neutrino background (Kolb and Turner, 1990). The cosmic neutrino background was produced at a temperature of ~ 1 MeV, when weak interactions decoupled. Its current number density and temperature can be calculated with a straightforward application of relativistic statistical physics, for neutrinos light enough to be relativistic at decoupling. The only complication is the annihilation of positrons and electrons after neutrino decoupling, which boosts the photon temperature relative to the neutrinos. The result is that the temperature T_ν and number density for each neutrino species n_ν are

$$T_\nu = \left(\frac{4}{11}\right)^{1/3} T_\gamma \qquad\qquad n_\gamma = \left(\frac{3}{11}\right) n_\nu . \tag{1.10}$$

The effective number of active neutrino species, N_ν^{eff}, affects the expansion rate during radiation domination, and through this the production of helium during Big Bang Nucleosynthesis. Current constraints from BBN are $N_\nu^{\mathrm{eff}} = 2.4 \pm 0.4$ (Simha and Steigman, 2008), consistent with the standard model value of $N_\nu^{\mathrm{eff}} = 3.04$.

All three of these flavours of neutrino have now been confirmed in accelerator experiments. The measurement of oscillations between the three neutrino flavours has shown the existence of a mass splitting between the three known neutrino species, implying a lower limit on the sum of the neutrino masses, $M_\nu > 0.05$ eV (Becker-Szendy et al., 1992; Fukuda et al., 1998; Ahmed et al., 2004). Precise measurement of the mass of tritium β-decay products provides an upper limit of ~ 2.2 eV (Lobashev, 2003) on the mass of the electron neutrino, with future experiments hoping to reach constraints of 0.2 eV (Wolf and KATRIN collaboration, 2010). However, the mass of the electron neutrino (< 1 eV)

is so much smaller than the mass of other fundamental particles that obtaining better constraints from decay reactions is extremely challenging.

Fortunately, the large scale structure of the matter distribution places constraints on neutrino mass, as discussed in e.g. Lesgourgues and Pastor (2006). Once massive neutrinos have cooled sufficiently to become non-relativistic, they behave as a constituent of the dark matter. Furthermore, while most dark matter is cold, that is, lacking significant thermal dispersion, neutrinos are hot dark matter, which retain thermal velocity from the early Universe. Free streaming of hot dark matter erases a proportion of fluctuations on small scales, depending on what proportion of the total dark matter it makes up. This allows us to weigh the neutrino background. I describe this effect more quantitatively in Chapter 6.

Future galaxy surveys (Schlegel et al., 2009), CMB lensing (Planck Collaboration, 2005; Niemack et al., 2010) and weak lensing probes may measure the power spectrum sufficiently accurately to constrain the neutrino mass to ~ 0.05 eV (Abazajian et al., 2011), guaranteeing a positive measurement and potentially distinguishing the neutrino mass hierarchy, provided systematic errors can be kept under control. This would help inform the construction of particle physics theories beyond the standard model.

1.7 The Lyman-α Forest

As mentioned in Section 1.1, the Lyman-α forest is a feature of quasar spectra. Quasars are thought to be super-massive black holes with accretion discs which radiate, making them extremely luminous objects visible at high-redshift. These properties let them act as beacons, each illuminating a narrow beam of the Universe. Baryons present along the beam will leave their mark on quasar spectra as absorption features, and the varying redshifts to the absorber means a single transition line may appear many times. The Lyman-α forest is the overlapping absorption from the Lyman-α transition, when neutral hydrogen goes from its first excited state to its ground state.

The first cosmological use of the Lyman-α forest was that proposed by Gunn and Peterson (1965). They pointed out that in a completely neutral Universe, quasar spectra would be completely absorbed at shorter wavelengths than 1216Å, the rest frame wavelength of the Lyman-α transition. The lack of what became known as the Gunn-Peterson trough at $z = 3$ suggested that hydrogen is mostly ionised in the late Universe. Croft et al. (1998) later suggested that, by measuring the statistics of absorption in quasar spectra, it would be possible to constrain the clustering of matter. Advances in technology meant that it was possible to use numerical simulations to predict the Lyman-α absorption from cosmic structure (Hernquist et al., 1996), and thus quantify the relation between statistical properties of the Lyman-α forest and cosmological parameters.

Unlike most other cosmological probes, the Lyman-α forest is sensitive to gas and ionisation physics, which must be constrained in concert with the cosmology. The biggest systematic uncertainty comes from the requirement to simultaneously constrain the mean absorption, which depends on the neutral fraction of hydrogen. This affects the amplitude of the absorption correlations and so is partially degenerate with the amplitude of the density perturbations (McDonald et al., 2005a), which weakens the cosmological constraints. There is also uncertainty in calculating the temperature and temperature-

density relation of the gas. Both these uncertainties are a reflection of a lack of knowledge about the details of the ionising photon background due to stars and quasars.

Using the Lyman-α forest as a probe of structure has in recent years been refined and applied to increasingly large datasets by, e.g. Viel et al. (2004); McDonald et al. (2006); Seljak and Slosar (2006). In Chapter 5, I present work which builds on these analyses to test the assumption of a smooth and nearly scale-invariant primordial power spectrum on the smallest scales currently possible. I both analyse current data and make predictions for future surveys. These future surveys will increase the number of spectra further and measure not only the correlations along the line of sight but also those between quasar spectra, producing a full three dimensional map of the IGM (Slosar et al., 2009, 2011).

The unique feature of the Lyman-α forest as a probe of structure formation is that it is a direct probe of baryons. Thus it has allowed us to refine understanding of baryonic physics, doing for gas simulations what studies of large scale structure have done for numerical methods of describing gravitational collapse.

1.8 Organisational Notes

The structure of this thesis is as follows. Chapter 2 describes technical information and methods that underlie the work presented in the later chapters. Chapters 3-6 present the main work of this thesis, which has two themes: inflation, Chapters 3 and 4, and the growth of structure, Chapters 5 and 6.

Chapter 3 discusses fine-tuning criteria for a variety of inflaton potentials and how the choices one makes when defining a criterion affects predictions for the tensor-scalar ratio. In Chapter 4, I move to a particular inflation model, brane inflation, and explore the status of the overshoot problem.

Chapter 5 uses new simulations of the Lyman-α forest to reconstruct the shape of the power spectrum on the smallest scales, testing the assumption of near scale-invariance, motivated by inflation. In Chapter 6, I perform a suite of simulations of massive neutrinos to test the accuracy of current theoretical prescriptions for the effect they have on the matter power spectrum.

The papers on which this thesis is based are listed below. In order to present each chapter as an internally consistent piece of work, I have chosen not to alter the introductions of each paper from the published version, although I have in some cases made changes to the ordering of Sections, and I have changed formatting to fit the style of this thesis.

Chapter 3: *Fine-tuning criteria for inflation and the search*
for primordial gravitational waves
This work was published as S. Bird, H. Peiris and R. Easther, 2008, Phys. Rev. D, volume 78, article 083518 and was carried out in collaboration with the named co-authors.

Chapter 4: *Brane inflation and the overshoot problem*
This work was published as S. Bird, H. Peiris and D. Baumann, 2009, Phys. Rev. D, volume 80, article 023534 and was carried out in collaboration with the named co-authors.

Chapter 5: *Minimally parametric power spectrum reconstruction*
from the Lyman-α forest
This work was published as S. Bird, H. Peiris, M. Viel and L. Verde, 2011, MNRAS, volume 413, issue 3, pages 1717-1728 and was carried out in collaboration with the named co-authors.

Chapter 6: *Massive neutrinos and the non-linear matter power spectrum*
This work was carried out in collaboration with M. Viel and M. Haehnelt, and is in preparation for submission to MNRAS.

'Your years of toil,'
Said Ryle to Hoyle,
'Are wasted years, believe me.
The steady state
Is out of date.
Unless my eyes deceive me,

My telescope
Has dashed your hope;
Your tenets are refuted.
Let me be terse:
Our universe
Grows daily more diluted!'

         – Barbara Gamow

2

Methods

In this Chapter, I discuss in more detail two topics which are important for later Chapters: inflation and numerical simulation of structure. Section 2.1 explains how a scalar field can produce inflationary expansion, and Section 2.2 discusses a model of inflation derived from string theory. Section 2.3 is dedicated to the techniques and algorithms used when simulating structure, and finally Section 2.4 derives the formulae for Lyman-α forest absorption.

2.1 The Slow-Roll Formalism for Inflation

This Section discusses the mechanisms of inflation, and how a scalar field can produce inflationary expansion. As discussed in Section 1.3, inflation requires matter with negative pressure. A scalar field ϕ, with potential $V(\phi)$, can have this property. The Einstein equations imply it has equation of state

$$\rho = \frac{\dot{\phi}^2}{2} + V(\phi) \qquad\qquad p = \frac{\dot{\phi}^2}{2} - V(\phi) \, , \qquad\qquad (2.1)$$

which has negative pressure when moving sufficiently slowly that $V > \dot{\phi}^2/2$. This is known as slow-roll inflation. The Friedmann equations are

$$H^2 = \frac{1}{3M_{\mathrm{P}}^2} \left(\frac{\dot{\phi}^2}{2} + V(\phi) \right) \qquad\qquad \frac{\ddot{a}}{a} = \frac{1}{3M_{\mathrm{P}}^2} \left(V(\phi) - \dot{\phi}^2 \right) \, , \qquad (2.2)$$

defining the reduced Planck mass as $M_{\mathrm{P}}^{-2} = 8\pi G$ in units such that $c = \bar{h} = 1$.

    Energy conservation provides the final equation needed to close the system and implies that the scalar field behaves as a damped harmonic oscillator in a potential

$$\ddot{\phi} + 3H\dot{\phi} + V'(\phi) = 0 \, , \qquad\qquad (2.3)$$

where a prime denotes a derivative with respect to ϕ and the dot a derivative with respect to t, cosmic time.

Realising inflation requires a negative pressure and a slowly varying Hubble parameter. The first condition can be achieved by imposing a small kinetic energy $\dot{\phi}^2/2 \ll V(\phi)$, and the second with a small acceleration $\ddot{\phi} \ll V'(\phi)$ [1]. In this case, the equations of motion become

$$3H\dot{\phi} + V'(\phi) \approx 0 \qquad\qquad H^2 \approx \frac{V}{3M_{\mathrm{P}}^2}. \qquad (2.4)$$

To check this leads to accelerated expansion ($\ddot{a} > 0$), substitute Eq. 2.4 into Eq. 2.2 to find

$$\frac{\ddot{a}}{a} = H^2 \left(1 - \frac{\dot{\phi}^2}{2M_{\mathrm{P}}^2 H^2} \right) \qquad (2.5)$$

$$\approx H^2 \left[1 - \frac{M_{\mathrm{P}}^2}{2} \left(\frac{V'}{V} \right)^2 \right]. \qquad (2.6)$$

So, defining the first slow-roll parameter ϵ_V, accelerated expansion occurs when

$$\epsilon_V = \frac{M_{\mathrm{P}}^2}{2} \left(\frac{V'}{V} \right)^2 < 1. \qquad (2.7)$$

This condition for accelerated expansion implies the first slow-roll condition $\dot{\phi}^2/2 \ll V(\phi)$. Eq. 2.4 can be differentiated to find the second condition, $\ddot{\phi} \ll V'(\phi)$, in terms of the inflaton potential

$$\ddot{\phi} \approx \frac{\dot{\phi}}{3H} \left(V'' - 3H'\dot{\phi} \right) \qquad (2.8)$$

$$\approx \frac{V'}{3} \left(M_{\mathrm{P}}^2 \frac{V''}{V} - \epsilon_V \right). \qquad (2.9)$$

Defining the second slow-roll parameter $\eta_V = M_{\mathrm{P}}^2 V''/V$ shows that slow-roll holds when $\eta_V \ll 1$.

The slow-roll formalism can be used to derive the power spectrum of perturbations from inflation. However, this calculation is rather involved, requiring careful attention to issues of gauge, and somewhat tangential to the main direction of this work. I therefore simply quote the results and explain their physical significance. The power spectra of scalar perturbations $P_{\mathrm{s}}(k)$ and tensor perturbations $P_{\mathrm{t}}(k)$ are given by

$$P_{\mathrm{s}} = \left. \frac{H^2}{\pi \epsilon_V M_{\mathrm{P}}^2} \right|_{k=aH}, \qquad\qquad P_{\mathrm{t}} = \left. \frac{16H^2}{\pi M_{\mathrm{P}}^2} \right|_{k=aH}. \qquad (2.10)$$

These quantities are evaluated at horizon-crossing ($k = aH$) because this is the time at which the quantum scalar field perturbations are frozen in. The factor of 16 in P_{t} arises because the graviton is spin two, while the scalar field is spin zero.

---

[1] I can choose the direction of ϕ such that $V' > 0$.

Derivatives of the Hubble parameter can be used instead of an expansion in potential slow-roll parameters. While this makes the potential parameters less obviously related to slow-roll, it allows us to make use of the Hamilton-Jacobi formalism and construct a potential that adheres to given observational parameters.

The master equation for the Hubble expansion is obtained by differentiating Eq. 2.2 with respect to time and substituting $\ddot{\phi}$ using Eq. 2.3, to obtain the exact equation

$$H' = -\frac{\dot{\phi}}{2M_\mathrm{P}^2} \,.$$

(2.11)

The Hubble slow-roll parameters can be defined as

$$\epsilon_H = 2M_\mathrm{P}^2 \left(\frac{H'}{H}\right)^2 \qquad\qquad \eta_H = 2M_\mathrm{P}^2 \left(\frac{H''}{H}\right) \,.$$

(2.12)

With this definition, Eq. 2.6 becomes the exact equation

$$\frac{\ddot{a}}{a} = H^2 \left(1 - \epsilon_H\right) \,.$$

(2.13)

The second slow-roll condition becomes $\eta_H \ll 1$, and, by interpreting H as $H(\phi)$, the equations of motion can be recast with ϕ as the dependent variable.

Finally, the Hubble slow-roll parameters are related to the potential slow-roll parameters, to first order in slow-roll, by

$$\epsilon_H \approx \epsilon_V \,, \qquad\qquad \eta_H \approx \eta_V - \epsilon_V \,,$$

(2.14)

so in practice either formalism can be used as the need arises. In Chapter 3, I use the potential slow-roll formalism, while in Chapter 4, I use Hubble slow-roll parameters. In Chapter 4, I discuss the effect of a non-canonical kinetic term in the inflaton action and the resulting speed limit. This is realised through a Lorentz factor $\gamma(\phi)$, which prevents the inflaton from accelerating quickly, even where the potential is non-flat.

2.2 D-brane Inflation and String Theory

In this Section I give a brief overview of the physical motivation and reasoning behind the model of inflation investigated in Chapter 4 which is grounded in string theory.

String theory is currently the best candidate for a theory of quantum gravity, which also describes particle physics at very high energies. As mentioned in Section 1.3, one of the major unsolved problems in inflation is the lack of an inflaton potential derived from first principles. Since string theory is a theory of high energy particle physics, asking whether it can produce the inflaton is a very natural question.

String theory attempts to explain relativistic particles as one dimensional strings, fundamentally fuzzy on the string scale. Because the string scale is lower than the Planck scale, this fuzziness can resolve the ultraviolet catastrophes of quantum field theory. Furthermore, it tames the high energy quantum behaviour of gravity with higher order corrections corresponding to loop terms for the quantised string. String theories have an internal conformal gauge symmetry, related to the ability to re-parametrize the

coordinate system on the string. To preserve this symmetry, string theory requires several extra internal degrees of freedom, the famous extra dimensions. Because these dimensions are not observed, they are presumed to be compactified.

As well as one-dimensional string objects, string theory admits the existence of higher-dimensional membranes (or manifolds), known as D-branes [2]. D-branes are fully dynamical objects and may possess charge and mass. Furthermore, any number of the dimensions of the surface may be observably large. Thus the observable Universe can be viewed as a four dimensional D-brane moving in some higher dimensional space. The higher dimensional space is unobservable from our point of view, either because the higher dimensions are collapsed, or because we (and the rest of the Universe) are charged with some higher-dimensional field and thus confined to the brane.

The original formulation of brane inflation by Dvali and Tye (1999) postulated a D3-brane and a $\overline{\text{D3}}$-brane, which possess opposite sign charges of a higher-dimensional generalization of electromagnetism. Because they are charged, these two branes attract each other, and as they fall together their mutual potential energy is partly converted into expansion. From the point of view of an observer confined to the brane, it would appear as if the Universe is four dimensional and contains a scalar field moving in a potential. This is interpreted as the inflaton.

Unfortunately, it soon became apparent that this inflaton potential was not sufficiently flat to agree with observations. In order to resolve this, Kachru et al. (2003) proposed a stack of $\overline{\text{D3}}$-branes, which would use their charges to create a warped region of space, flattening the inflaton potential. It was initially believed that this potential would be ruined by the fields necessary for moduli stabilization, but Baumann and McAllister (2007) computed further corrections to the inflaton potential, due to interaction terms with other fields present in the wider, unwarped, space, which restored the flat potential necessary for inflation. This model is the basic picture in Chapter 4.

Unlike the original models of inflation, such as those discussed in Section 2.1, brane inflation does not end when $\epsilon_H \sim 1$. Instead, branes ultimately collide, annihilating their mutual charges and releasing their expanded energy into a low-energy brane corresponding to the present-day Universe [3]. There are also a few problems with the approach which have yet to be resolved. First, it is not clear that a space warped by $\overline{\text{D3}}$-branes is generic. This problem is somewhat analogous to the well-known problem of whether a sufficiently flat potential is generic in scalar field inflation.

It is important to remember that the model of D-brane inflation discussed above is not the only possible model of inflation from string theory; there is also, for example, the axion monodromy model of Silverstein and Westphal (2008). It is also likely that there are several further ways of realising inflation in string theory that have yet to be discovered.

For these reasons, it is probably best to regard D-brane inflation as a detailed case study, or toy model, and use it as a laboratory to determine the sorts of challenges that a particle physics model of inflation might face. This is the approach I take in Chapter 4,

---

[2]The name D-brane is used because these objects were originally introduced as surfaces upon which open strings obeying Dirichlet boundary conditions could end. Open strings have two ends; for massless strings, ends can either move at the speed of light (Neumann boundary conditions) or be held stationary, confined to some manifold by a charge (Dirichlet boundary conditions).

[3]The exact mechanism for this is not yet well understood.

where I examine whether D-brane inflation solves a particular problem of classical scalar field inflation, called the overshoot problem. I hope thereby to gain a hint into whether this problem generically exists in string theory models.

2.3 Simulating Structure

In this Section, I give an overview of some of the simulation techniques used in Chapters 5 and 6, and motivate my choice of codes.

2.3.1 Gravity

In this Subsection, I discuss some of the methods used for simulating the formation of structures under Newtonian gravity. Newtonian gravity is expected to be a good approximation in most circumstances because the Universe is dominated by cold dark matter and baryons, whose velocities and matter densities are non-relativistic. One circumstance where this is not always true is when simulating massive neutrinos, as in Section 2.3.4.

Since Birkhoff's theorem implies that spherically symmetric gravitating structures can be represented by point masses for external observers, it is natural to try to simulate gravity using particles. The basic problem is that of N-body dynamics; we must find some method of summing the force on each particle from a vast number of other particles. The first attempt at simulating structure formation numerically was probably that of Holmberg (1941) who, lacking modern computers, resorted to an ingenious method of modelling gravitational interactions. 37 light-bulbs were carefully positioned and lit, representing individual mass elements. Their gravitational evolution was computed by carefully moving each bulb by hand according to the intensity of the light falling upon it: the light being a proxy for gravitational attraction, since both radiation intensity and the gravitational force are proportional to $1/r^2$.

This is an early example of the direct summation method of N-body calculations. See, for example, Aarseth (2003) for more details of this approach. To simulate this process with a finite number of particles, the gravitational force on each particle must be calculated for each small time-step Δt. A direct sum would look something like

$$F_i = G \sum_j M_j (x_i - x_j)^{-2} \,, \tag{2.15}$$

where x_i are positions and j is summed over all particles. Since this sum must, in principle, be evaluated for every particle in turn, the cost of force calculations scales like N^2, where N is the number of particles, making direct summation impractically slow on cosmological scales. The problem of modelling gravitational collapse can therefore be restated as that of finding an approximation to the above summation which is both fast and accurate.

It is important to realise that these N-body particles are not physical point particle objects. When performing N-body simulations, Birkhoff's theorem is implicitly being used to approximate a large complex object as a point particle at the centre of mass of the original object. This neglects internal structure on scales smaller than the mass scale of the N-body particle. To check that structure on smaller scales than those resolved does

not have a strong impact on results, an identical simulation is run with slightly lower mass resolution. Boundary effects can be minimised by assuming periodic boundary conditions and surrounding a simulation box with an identical copy of itself; physically this is useful precisely because the Universe is homogeneous on large scales.

The gravitational dynamics can be described by the Poisson equation,

$$\nabla^2 \phi(x) = 4\pi G\rho\,, \tag{2.16}$$

where ϕ is the gravitational potential. Once the density field, ρ, is known on a grid, this equation can be solved using Fourier methods, which are much faster than direct summation. Therefore, one important approximation is the particle-mesh approach, where particles are smoothed to the width of a grid cell and interpolated onto a grid, which is then Fourier transformed, and the forces computed.

This can resolve long-range forces very accurately for little cost, but it has finite resolution, corresponding to the spacing of the grid used. Collapse under gravity is a problem where the dynamics cover an extremely wide range of scales, so this loss of accuracy is problematic. Early codes, such as that of Efstathiou and Eastwood (1981), used a particle mesh grid on large scales and resolved small scales with direct summation. This results in the so-called particle-particle-particle-mesh, or P³M, algorithm. However, the reliance on direct summation on small scales is still prohibitively slow for many problems.

The final development in gravity solvers is the Tree-PM algorithm. Like P³M, this uses a particle mesh to calculate long-range forces. The innovation is to use a tree algorithm for short-range forces. This is essentially a multipole method, which groups particles into a hierarchy of cells. For sufficiently distant force interactions, the effect of multiple particles can be approximated into a single multipole force, substantially reducing the number of computations and allowing essentially arbitrary resolution. TreePM is the algorithm used in GADGET (Springel, 2005), the code used to run my simulations.

2.3.2 Hydrodynamics

In Chapter 5, I investigate the Lyman-α forest. The Lyman-α forest measures the distribution of neutral hydrogen, so to simulate it I must describe not only dark matter, but baryons. The first step to doing this is a description of adiabatic gas physics and pressure forces. In this Section I give an overview of the commonly used models.

The natural way to describe an adiabatic gas is with a fluid approximation, unlike the collisionless gravitational dynamics described in 2.3.1, which was modelled using particles. Computational fluid dynamics is a well-studied subject, and modern approaches often use an Eulerian grid approach. Here space is split into fixed grid cells, and the evolution of the fluid is tracked with the flow across cell boundaries. This approach works very well for computational fluid dynamics. However, in cosmology, hydrodynamics must almost always be coupled to gravity, which introduces several complications. First, Eulerian grid codes with a fixed grid for cosmology suffers from the same drawback as the bare particle mesh method discussed in the last Section for gravity; limited spatial resolution. Many pure fluid dynamical problems have a fixed length scale, so this is not necessarily a problem. However, cosmological problems require a high dynamic range.

To overcome these limitations, some Eulerian codes use Adaptive Mesh Refinement (AMR). In AMR, grid cells are dynamically refined to a finer sub grid as the need arises. This allows unbounded dynamic range, while not wasting computation time or memory. However, AMR brings its own tradeoffs. The first problem is a practical one; the dynamic increase in the grid size means that it is very hard to predict the memory requirements of a fully adaptive code in advance. In addition, to refine a cell, a description of the fluid distribution on scales of the refined mesh is needed, while the code can only store information on the unrefined scales. In practice the fact that objects are constantly collapsing under gravity ameliorates this problem. There are several cosmological AMR codes in existence, for example RAMSES (Teyssier, 2002) and ENZO (O'Shea et al., 2004), which couples hydrodynamics using AMR to a particle mesh gravity solver. The simplistic gravity solver used in ENZO points to another problem with using grid codes for cosmology; modern gravity solvers are inherently particle based and fit badly into a grid-based fluid description. Having two entirely separate representations of matter adds to programming complexity and computational overhead. Modelling cosmological hydrodynamics with a particle-like description can ameliorate these difficulties.

This is the motivation for smoothed particle hydrodynamics (SPH). In SPH the fluid is sampled by a number of tracer particles, each of which corresponds to a Lagrangian fluid element of constant density. SPH discretises the fluid in mass rather than in spatial coordinates, and so it fits well with the particle based description of gravity. In particular, it automatically has a high dynamic range, because of the concentration of particles in areas of high matter density. A fluid description is then obtained by interpolating smoothly between particles. My simulations in Chapter 5 use GADGET, whose formulation of SPH explicitly conserves both entropy and energy (Springel and Hernquist, 2002). Entropy conservation avoids spurious mixing, but means that an artificial viscosity force is required to allow particle collisions to heat the gas, which in turn makes it hard for SPH to model shocks correctly. Recently, Springel (2010) has presented AREPO, which aims to combine the advantages of SPH and grid-based codes by discretizing the fluid using a Lagrangian moving mesh.

One potential problem with SPH is that because it concentrates particles in regions of high-density, voids are comparatively under-resolved, and problems sensitive to low-density regions may suffer. In Chapter 5, I use GADGET to simulate the Lyman-α forest. Most of the Lyman-α forest comes from the weakly non-linear regime, in regions of comparatively small overdensity. Since there will be few shocks on these scales, of small dynamical importance, an SPH code should be appropriate. Lack of resolution in the voids, in principle, can cause concern. However, when Regan et al. (2007) compared GADGET to ENZO, they found that the theoretically higher resolution in voids did not in practice have a strong effect on the Lyman-α forest. In addition, the computational overheads of AMR and the ability to switch off hydrodynamics in overdense regions meant that GADGET was significantly faster when optimised for the Lyman-α forest.

2.3.3   Ionisation and Thermal Physics

Understanding the neutral hydrogen which makes up the Lyman-α forest also requires an understanding of the ionisation and thermal physics of the gas. In my simulations, the gas is supposed to be kept in ionisation equilibrium by an ultraviolet photon background,

assumed to be spatially uniform and supplied as an external input to the simulation. This assumption of a uniform UV background is justified because cosmic hydrogen is mostly ionised at redshifts of $z = 2 - 4$, at least in the near mean density gas which gives rise to the Lyman-α forest, and hence the gas is optically thin to hydrogen-ionising photons. Optically thin gas also justifies my assumption that the gas is in ionisation equilibrium. For optically thin gas, the neutral fraction will be changed by alterations in the amplitude of the UV background, or by the growth of structure, both much slower processes than ionisation. On the other hand, if the gas were optically thick the neutral fraction would change on the much faster timescale of an ionisation front.

While the Lyman-α gas is optically thin to hydrogen-ionising photons, the situation is different with respect to helium. Helium becomes doubly ionised at $z \sim 3$; before the completion of this process the gas is not optically thin to helium-ionising photons and several of the above assumptions break down. Fortunately the Lyman-α forest is only directly sensitive to neutral hydrogen, and so the main effect of helium reionisation is to increase the temperature of the gas (Abel and Haehnelt, 1999). In order to quantify this effect, I would need to include a robust model of radiative transfer into my simulations, which is not currently feasible. Instead, I account for the uncertainty caused in the thermal history by considering simulations with a variety of temperatures compatible with observations, each generated by using a different effective helium photo-heating rate.

2.3.4 Neutrinos

In this Section I examine some of the methods for simulating the effect of neutrinos on the growth of structure. As discussed in Section 1.6, neutrinos are weakly interacting light particles, which retain significant thermal dispersion and thus a long free-streaming length in the late Universe. Because they are weakly interacting particles, they can be treated in our simulations as an extra variety of dark matter, but, unlike the dominant cold dark matter, with significant thermal velocity. This approach allows neutrinos to be simulated fully self-consistently, including non-linear effects. Furthermore, since neutrinos do not cluster significantly on scales smaller than their free-streaming length, computation of the tree component of the gravitational forces required for cold dark matter can be omitted. This significantly speeds up computation.

One complication with this approach is the impact of shot noise. For low-mass neutrinos the thermal velocity may be much higher than the gravitational flow velocity. The neutrino distribution is only sampled at finite accuracy, which can lead to thermal shot noise dominating the power spectrum. This is most acute at high redshifts, where the neutrino temperature is greater and structure growth has not yet progressed to the same extent. Shot noise can be avoided by simply increasing the number of neutrino particles. While this is rather a brute force approach, and quickly reaches practical limits, it can be extremely effective.

An alternative method, which avoids a neutrino particle species altogether, is to treat the neutrinos as a perturbation to a single dark matter species. Here the cold dark matter is treated normally, but additional power, representing the massive neutrinos, is added to the particle-mesh grid (Brandbyge and Hannestad, 2009). The neutrino contribution is calculated analytically, using linear perturbation theory, which gives it arbitrarily high

resolution but neglects non-linear structure in the neutrinos. This approach is clearly preferred at very high redshift, when non-linear clustering in the neutrinos is minimal, and shot noise is of most concern. However, for $z \leq 2$ the situation reverses, and particle-based simulations incorporate more of the relevant effects.

With this in mind, Brandbyge and Hannestad (2010) proposed a hybrid scheme. Massive neutrinos are initially included using the grid-based method. Once the neutrino thermal velocity has dropped below a critical level, the lowest momentum grid cells are converted into particles, with initial properties given by those of the grid cell before conversion. The mass corresponding to these neutrino particles is then removed from the grid. With this method, they claim to combine the advantages of particle and grid-based simulations and are able to examine small-scale structures forming in the neutrinos.

I do not use this hybrid approach in Chapter 6, preferring purely particle or grid based methods. I am interested only in the effect of neutrinos on the non-linear matter power spectrum, and not on the small-scale neutrino structure. Furthermore, the criteria for converting grid cells into particles adds complexity and potential fragility to the numerical implementation, making it more difficult to prove convergence. Therefore in Chapter 6, I have taken the approach of performing both grid and particle-based simulations that are fully converged and checking them for consistency with each other. Since I find explicitly that shot noise is negligible, I prefer the particle-based simulations when there is disagreement.

In practice, the most important limitation of my simulations in Chapter 6 is that in the early Universe, neutrinos are somewhat relativistic, reducing their effective mass. Because GADGET is a Newtonian code, which incorporates constant mass neutrino particles, it cannot adequately account for this and over-estimates the neutrino suppression, even in the linear regime. I avoid this by starting my simulations at a sufficiently low redshift, but a code which models the changing effective mass of the neutrino would be preferred.

2.3.5   Initial Conditions

In this Section I give a brief description of the methods used to generate the initial conditions in my simulations. The basic problem is to turn the power spectrum of density fluctuations, calculated using perturbation theory, into a three dimensional distribution of particles. First the density field must be realised from the power spectrum. It is easiest to do this in Fourier space, as the power spectrum gives the amplitude of Fourier modes directly. The phases of the Fourier modes are generated randomly; this is equivalent to assuming that the statistics of the initial field are Gaussian. In practice this is not entirely accurate; neglecting the possibility of small primordial non-Gaussianities, gravitational evolution causes some information leakage into the phases, and a deviation from pure non-Gaussianity. The error from this can be easily made negligible by starting simulations at high redshift.

The initial positions of the particles, assumed to be on a regular grid in my simulations, need to be perturbed by the density field. To do this I use the first-order Lagrangian approximation of Zel'dovich (1970). The Zel'dovich approximation assumes that a particle with initial position vector q_i has been moved inertially to its new location x_i by a

displacement vector $\Phi_i(q_i)$, generated from the initial density field, such that

$$x_i = a(t)q_i + D(t)\Phi_i(q_i) \tag{2.17}$$

$$\frac{\partial \Phi_i}{\partial q_i} = -\delta(q_i) \, , \tag{2.18}$$

where $a(t)$ is the scale factor, $D(t)$ the linear growth rate and $\delta(q_i)$ the density perturbation field. The initial velocity can be found easily by differentiating with respect to time

$$v_i = -\frac{dD(t)}{dt}\Phi_i(q_i) \, . \tag{2.19}$$

The Zel'dovich approximation is accurate for small displacements, but it breaks down once particle displacements are sufficiently large for particle trajectories to intersect. It is therefore important to start simulations at a sufficiently early redshift, so that the initial particle displacements are on average smaller than the mean particle spacing. An alternative is to use second-order perturbation theory, which adds a correction term to the above equation (e.g. Scoccimarro (1998).

2.4 Simulating Lyman-α Spectra

In this Section I detail the procedure for simulating a Lyman-α forest spectrum, given a snapshot of a structure simulation, performed using the methods described in Section 2.3. The first step is to find the velocity of each particle, including both peculiar velocities and the Hubble flow. In my simulations, peculiar velocities have an effect around the 10% level.

Next, I calculate the optical depth, τ_λ, at a wavelength λ, as defined by the line integral

$$\tau_\lambda = \int \sigma_\lambda n_H dl \, , \tag{2.20}$$

where σ_λ is the cross-section for the transition and n_H is the number density of the neutral hydrogen. σ_λ is given by the rest cross-section, σ_α, multiplied by a broadening function, Φ,

$$a_\lambda = \sigma_\alpha \times \Phi \, . \tag{2.21}$$

The oscillator strength of the Lyman-α transition, f_α, is defined by the ratio between the cross-section of the Lyman-α transition and the cross section of the transition involving a free electron

$$f_\alpha = \frac{\sigma_\alpha}{\pi r_0 \lambda} \, , \tag{2.22}$$

where λ is the rest wavelength of the Lyman-α transition. The classical radius of the electron, r_0 is related to the Thomson cross-section, σ_T by

$$r_0 = \sqrt{\frac{3\sigma_T}{8\pi}} \, . \tag{2.23}$$

Hence

$$\sigma_\alpha = \sqrt{\pi \frac{3\sigma_T}{8}} f_\alpha \lambda \, . \tag{2.24}$$

To compute the broadening function, I neglect natural broadening from the intrinsic uncertainty in the energy levels of the hydrogen atom. Natural broadening is only important in the densest absorbers (damped Lyman-α systems), which my simulations lack the resolution to adequately resolve. The effect of DLAs can be included by a correction applied when calculating the likelihood.

Hence the only form of broadening present is Doppler broadening. In an absorber of temperature T_H, mass m_H, and velocity v, the probability of a particle having zero velocity relative to an incoming photon is

$$\Phi = \frac{c}{\sqrt{\pi}b} \exp\left[-\frac{v^2}{b}\right] , \tag{2.25}$$

$$\text{where } b = \sqrt{\frac{2kT_H}{m_H}} . \tag{2.26}$$

Hence, a wavelength bin at position i will suffer absorption from an HI absorber in bin j according to:

$$\tau_{ij} = \sigma_\alpha \Phi n_H a \Delta \tag{2.27}$$

$$= \sigma_\alpha \frac{c}{\sqrt{\pi}b} n_H a \Delta \exp\left[-\left(\frac{v_i - v_j}{b}\right)^2\right] . \tag{2.28}$$

Here Δ is the bin width, and a is the expansion factor.

The flux in each bin is then simply $\mathcal{F} = e^{-\tau}$. Each spectrum is smoothed with a simple 3 point boxcar average, following Viel et al. (2004), and the flux power spectrum from the simulation box is defined to be the average over a number of simulated spectra.

"Inflation is as violent as a mugger, as frightening as an armed robber and as deadly as a hit man."

– Ronald Reagan

3

Fine-tuning Criteria for Inflation and the Search for Primordial Gravitational Waves

Abstract

We revisit arguments that simple models of inflation with a small red tilt in the scalar power spectrum generically yield an observable tensor spectrum. We show that criteria for fine-tuning based upon the algebraic simplicity of the potential depend strongly upon the explicit assumptions they incorporate, particularly regarding the end of inflation. In addition, some models with algebraically simple potentials require carefully tuned initial field configurations, and not all types of fine-tuning are identifiable via the algebraic simplicity of the potential. Conversely, in the absence of a strong prior on the mechanism that ends inflation, we demonstrate the existence of potentials with vanishingly small tensor amplitudes, which are natural in terms of both their algebraic form and initial conditions. We thus argue that proposed experiments (CMBPol or BBO) which make highly sensitive measurements of the tensor amplitude cannot definitively rule out the inflationary paradigm.

3.1 Introduction

In the simplest scenarios, inflation is driven by a single scalar field, the "inflaton", rolling in a slowly varying potential (Guth, 1981; Sato, 1981; Mukhanov and Chibisov, 1981; Linde, 1982b; Albrecht and Steinhardt, 1982). Sufficient inflation (≈ 60 e-folds) smooths the Universe and solves the flatness, horizon, and monopole problems. Initial inhomogeneities originate as quantum fluctuations in the inflaton field and the metric, sourcing scalar (Hawking, 1982; Guth and Pi, 1982; Starobinsky, 1982; Bardeen et al., 1983) and tensor (Grishchuk, 1975; Starobinsky, 1979) perturbations which are almost Gaussian and scale-invariant (Bardeen et al., 1983). These broad-brush predictions are consistent

with observations (Komatsu et al., 2009; Peiris and Easther, 2006a,b, 2008).

The shape of the inflaton potential is not predicted by theory; it is only required to satisfy the slow-roll conditions. However, the shape of the potential is constrained by the scalar spectral index, n_s, which measures the deviation of the scalar power spectrum from scale invariance, and the tensor-scalar ratio, r, which measures the amplitude of the tensor modes relative to that of the scalar modes. Data from the *Wilkinson Microwave Anisotropy Probe* (WMAP) combined with a compilation of supernovae and baryon acoustic oscillations data gives $n_s = 0.960^{+0.014}_{-0.013}$ at the 68% confidence level (CL) if the tensor-scalar ratio $r = 0$, and $n_s = 0.968 \pm 0.015$ when we marginalize over r (Dunkley et al., 2009; Komatsu et al., 2009). There is currently no evidence for the existence of a primordial tensor background: the 95% CL upper limit from the same data compilation is $r < 0.20$.

A detection of primordial tensors would have tremendous implications for cosmology as their existence is a key prediction of inflation (Starobinsky, 1979). Moreover, measuring a non-zero value of r would eliminate the ekpyrotic/cyclic scenarios, which predict an unobservably small primordial tensor amplitude (Khoury et al., 2002; Boyle et al., 2004). Finally, the observed amplitude of the tensor spectrum fixes the energy scale of inflation, removing a major uncertainty in inflationary model building. On the other hand, an observable tensor spectrum would imply the need for super-Planckian field variations during inflation in simple single field models (Lyth, 1997), which raises specific theoretical challenges.

Cosmic variance, polarized foregrounds, and weak lensing of the E-mode may ensure that $r \sim 10^{-3}$ is the effective lower limit for a detection of tensor modes in the cosmic microwave background (CMB) (Knox and Song, 2002; Kesden et al., 2002; Verde et al., 2006; Amarie et al., 2005). Consequently, it is important to know whether inflation makes a *generic* prediction for the value of r and, if so, whether it falls within the detectable range. A plethora of explicit inflationary models with varying degrees of physical motivation have been proposed. Many models do have observable values of r (e.g. the chaotic inflation of Linde (1983)); conversely, a detection of tensors would rule out many (Baumann and McAllister, 2007), but not all (Dimopoulos et al., 2008; Easther and McAllister, 2006; Silverstein and Westphal, 2008), explicit string-theory constructions. Consequently, to show that inflation has a "natural" range of r values, we must establish a "weight" that will give preference to certain classes of models.

One weighting proposal is that of Boyle et al. (2006) (BST), who offer a fine-tuning statistic based on the algebraic simplicity of the potential, along with five conditions which a successful inflationary potential should satisfy. These criteria are then applied to quartic polynomial potentials, with the conclusion that a high degree of fine-tuning is required for $n_s > 0.98$, or for $r < 10^{-2}$ with $n_s > 0.95$. If future observations confirm that $0.95 < n_s < 0.98$, the BST criterion suggests that r is within the observable range.

In addition to tuning the *parameters* of the potential, some potentials need a very special initial field configuration if inflation is to begin. This tuning can involve both the suppression of the kinetic energy relative to the potential energy (to ensure that the inflaton does not "overshoot" the inflationary portion of the potential) or the absence of inhomogeneities on the initial inflationary patch. This topic has been investigated for many of the common models of inflation (Belinsky et al., 1985; Goldwirth, 1990; Mendes and Liddle, 2000; Underwood, 2008; Chiba, 2008).

In particle physics, models are regarded as fine-tuned if they require correlations between parameters beyond those imposed by the symmetries of the underlying theory ('t Hooft, 1980), or when certain parameters are arbitrarily set to zero; models which are not fine-tuned in this sense are known as "technically natural" (Hotchkiss et al., 2008).

There are other definitions of naturalness based on the sensitivity of observables to small changes in their parameters (Barbieri and Giudice, 1988; Anderson and Castano, 1995). Hotchkiss et al. (2008) examined technically natural potentials which did not obey all of the BST viability conditions, and found red-tilted examples with $r < 0.01$ which were not fine-tuned under their definition. Other studies came to similar conclusions (Efstathiou and Chongchitnan, 2006; Smith et al., 2008). Inflationary potentials are often written down without reference to an underlying theory, so we do not necessarily know the relevant symmetries, nor the couplings between the inflaton and other fields which contribute loop corrections to the inflaton potential. Consequently, while technical naturalness is a powerful tool, it is less useful when applied to generic inflationary potentials. Moreover, inflationary potentials typically require some tuning in order that inflation happens at all.

We tackle two questions. Firstly, we show that even a mild relaxation of the rules imposed by BST significantly weakens their conclusions for the likely range of r. In addition, we ask whether the inflaton's initial conditions must be tuned in order for viable inflation to occur. We consider only the ratio of the kinetic to potential energy, treating the Universe as a homogeneous system and ignoring the ability of primordial inhomogeneities to suppress the onset of inflation. However, this requirement places a strong constraint on the inflationary model space, which does not overlap with the BST proposal, demonstrating the interplay between the *definition* of tuning, and the resulting conclusions for the likely values of r and n_s. Our aim is not to assess the intrinsic merit of different naturalness criteria. Rather, our purpose is to investigate whether conclusions about the natural value of r are sensitive to assumptions encoded in the tuning criteria. This question is of considerable importance when assessing and justifying CMB B-mode or gravitational wave experiments sensitive to r. Likewise, any conclusions we reach regarding the "natural" value of r in inflation will determine the significance of a null result from these experiments.

3.2 Methodology

3.2.1 Notation

We use natural units, $c = \hbar = 1$, and M_P is the reduced Planck mass. The first two slow-roll parameters are

$$\epsilon = \frac{M_P^2}{2} \left[\frac{V'(\phi)}{V(\phi)} \right]^2 , \tag{3.1}$$

$$\eta = M_P^2 \left[\frac{V''(\phi)}{V(\phi)} \right] , \tag{3.2}$$

where $V(\phi)$ is the inflaton potential, ϕ is the inflaton and primes denote derivatives with respect to ϕ. To first order, the spectral index and the scalar-tensor ratio are

$$n_s = 1 + 2\eta - 6\epsilon, \tag{3.3}$$
$$r = 16\epsilon. \tag{3.4}$$

These quantities are evaluated 60 e-folds before the end of inflation, ϕ_e, when $\phi = \phi_{60}$, which is taken to correspond to the CMB scales[1]. A minimally coupled inflaton with a canonical kinetic term has the equation of motion

$$\ddot{\phi} + 3H\dot{\phi} + V'(\phi) = 0, \tag{3.5}$$

where H is the Hubble parameter. The scale factor $a(t)$ is described by

$$H^2 = \left(\frac{\dot{a}}{a}\right)^2 = \frac{1}{3M_P^2}\left[V(\phi) + \frac{1}{2}\dot{\phi}^2\right], \tag{3.6}$$

where the overdot denotes a derivative with respect to coordinate time t. It is convenient to use the number of e-folds, N, as the time variable, via

$$dN = H dt. \tag{3.7}$$

Hence both N and ϕ increase as t increases[2]. Defining $\dot{\phi} = \Pi$, the equations of motion are

$$\frac{d\phi}{dN} = \frac{\Pi}{H}, \tag{3.8}$$
$$\frac{d\Pi}{dN} = -\left(3\Pi + \frac{V'}{H}\right). \tag{3.9}$$

Provided $\eta - \epsilon \ll 1$, an attractor solution exists for this system (Goldwirth, 1990), satisfying

$$\frac{d\Pi}{dN} = 0 \implies \Pi = -\frac{V'}{3H}. \tag{3.10}$$

In every result quoted in this Chapter, we have solved the full equations of motion numerically; no slow-roll approximation has been employed.

3.2.2   Selection Conditions

Following BST (Boyle et al., 2006), we consider quartic polynomial potentials and apply five conditions which define a viable inflationary model. These conditions largely overlap with those of BST; since we are investigating the impact of *changing* the viability conditions we specify them in some detail:

---

[1] Our conclusions are not sensitive to the details of reheating.
[2] The opposite sign convention can be found in the literature; with consistent usage, there is no practical distinction.

1. The density perturbations have an amplitude of roughly 10^{-5} 60 e-folds before the end of inflation, in agreement with observations.

2. At least 60 e-folds of inflation occur, with the trajectory starting on the slow-roll inflationary attractor and ϕ evolves monotonically as a function of time.

3. Inflation terminates in a deterministic and smooth manner, via $\epsilon = 1$, rather than through a hybrid-type transition (Linde, 1994).

4. After inflation, the field evolves smoothly to a minimum of the potential located at ϕ_m, and $V(\phi_m) = 0$.

5. V is bounded below, and the minimum, if metastable, is long-lived.

Beyond specifying these rules, we must also write down a generic potential. We follow BST and write down the quartic polynomial [3]

$$V(\phi) = V_0(\phi^4 + B\phi^3 + A\phi^2 + C\phi + D) . \qquad (3.11)$$

To satisfy condition (4), we require a minimum at ϕ_m, with $V(\phi_m) = 0$. Without loss of generality, we can set $\phi_m = 0$, which implies that $C = D = 0$. To satisfy condition (5), the potential must be bounded below, which implies $V_0 \geq 0$. The magnitude of V_0 is dynamically irrelevant, so we can choose it such that condition (1) is satisfied. Typical values of V_0 were around 10^{-16}; in no case did one of our tested potentials exceed $V(\phi) = M_P^4$. This is necessary (Linde, 1983) (but not sufficient) to keep quantum gravity effects under control. Our potential now has two free parameters,

$$V = V_0\phi^2(\phi^2 + B\phi + A) . \qquad (3.12)$$

Note that A has dimensions of M_P^2 and B has dimensions of M_P. We set $A \geq 0$ to ensure that the extremum at the origin is a minimum[4]. We apply a slightly stricter version of condition (5), namely that the minimum is *stable*, eliminating ambiguity in the definition of "long-lived". Now we can impose $V(\phi) \geq 0$, since the inflaton must evolve toward the lowest-lying point of the potential, which is assumed to be at $V(0) = 0$ (via condition (3)). Hence $V(\phi) = 0$ has at most two solutions for a valid potential, one of which is at the origin, so that $\phi^2 + B\phi + A = 0$ has at most one solution, for which we need an imaginary or vanishing determinant and thus

$$-2\sqrt{\frac{A}{M_P^2}} \leq \frac{B}{M_P} \leq 2\sqrt{\frac{A}{M_P^2}} . \qquad (3.13)$$

We construct candidate potentials by randomly sampling A and B with logarithmic priors within the given bounds and numerically testing to determine whether conditions (2) and (3) are met. After imposing these rules, we can only populate a subset of the

---

[3]The inflaton potential is the *effective* potential of ϕ and thus includes all loop corrections. Consequently, it can include terms that would be non-renormalizable if they appeared at tree level, and the restriction to the quartic form is effectively a further prior based on algebraic simplicity. Note that this cut excludes potentials of the Coleman-Weinberg form (Albrecht and Steinhardt, 1982; Linde, 1982a).

[4]For computational reasons, we also impose $A < 10^5 M_P^2$, but this makes no difference to our results.

(n_s, r) plane, as the assumed functional form of the potential and the selection conditions constitute a strong theoretical prior on the slow-roll parameters, and hence the observables. Such restrictive theoretical priors can lead to a "detection" of a tensor-scalar ratio even when the likelihood with respect to the data shows no evidence for a deviation from $r = 0$ (Destri et al., 2008). This point has also been made in Valkenburg et al. (2008). Our results are in agreement with Hoffman and Turner (2001); Peiris et al. (2003); Boyle et al. (2006). We stress that this cut is independent of any "weighting" function proposed to assess the degree of tuning of potentials that do comply with the above conditions.

3.2.3 Fine-Tuning Criteria

We will discuss two fine-tuning criteria. The first is that proposed by BST, which measures the fine-tuning in the parameters of the potential. Our second measures fine-tuning in the initial configuration of field space, as shown by the sensitivity to an initial kinetic term.

BST Fine-tuning Criterion

The fine-tuning criterion proposed by BST counts the number of "unnecessary features" occurring in the potential during the last 60 e-folds of inflation (Z_η). A feature is a zero of some derivative of η with respect to ϕ, so

$$Z_\eta = \sum_{p=0}^{P} \chi \left(\frac{d^p \eta}{d\phi^p} \right) , \tag{3.14}$$

where $\chi(f)$ counts the number of zeros of f within a given range; in this case the last 60 e-folds of inflation. Since all derivatives of η may be non-trivial, the value of P is arbitrary; we take $P = 10$. Our implementation differs slightly from that of BST, who take $\chi(\eta_{,N})$ in the place of $\chi(\eta_{,\phi})$: derivatives of N can introduce artificial features into the potential (via the $\phi_{,N}$ term), over-estimating the amount of fine-tuning. Our change weakens BST's conclusion that $n_s < 0.98$ to $n_s < 0.99$ for non-fine-tuned potentials, but has no other impact. BST stipulate that models with $Z_\eta > 1$ are fine-tuned, but their criterion is not intended to quantify the degree of fine-tuning in such cases; models with (say) $Z_\eta = 12$ and $Z_\eta = 30$ would be placed on the same footing. Our graphs therefore show all points with $Z_\eta \geq 12$ as equivalent.

Initial Conditions Fine-tuning Criterion

The question of whether inflation itself requires a fine-tuned initial field configuration has been discussed many times (Belinsky et al., 1985; Piran and Williams, 1985; Belinsky and Khalatnikov, 1986; Piran, 1986; Gibbons et al., 1987; Albrecht et al., 1987; Belinsky et al., 1988; Kung and Brandenberger, 1989, 1990; Goldwirth, 1990; Goldwirth and Piran, 1990; Goldwirth, 1991; Goldwirth and Piran, 1992). This is similar to a criterion applied to chaotic inflation in Belinsky et al. (1985); Piran and Williams (1985) and to new inflation in Goldwirth (1990). More generally, one can ask whether inflation generically begins with initial conditions that are anisotropic, inhomogeneous or have non-zero spatial curvature (Moss and Sahni, 1986; Albrecht et al., 1987; Kung and Brandenberger, 1989,

1990; Goldwirth and Piran, 1990; Goldwirth, 1991). We ignore these possibilities, but including them would in general strengthen the conclusions we obtain by working only with an initially arbitrary kinetic term. The inflationary Universe driven by a single scalar field is described by a second-order system of non-linear differential equations. We consider how sensitive this system is to its initial conditions, Π_i and ϕ_i, defining the criterion R_i for a given potential as the fraction of initial conditions leading to valid inflationary trajectories. That is:

$$R_i = \frac{\int F(\phi_i, \Pi_i) d\phi_i d\Pi_i}{\int d\phi_i d\Pi_i}, \tag{3.15}$$

where $F(\phi_i, \Pi_i)$ is an indicator function which is unity where (Π_i, ϕ_i) lead to successful inflation, and zero otherwise. A model which is minimally fine-tuned has $R_i = 1$, and one which is maximally fine-tuned has $R_i = 0$.

Notice that $d\phi_i d\Pi_i$ is not a dynamically invariant measure; the Hamiltonian phase space measure contains an extra factor $a^3(t)$. With this weighting, it has been shown that inflation is disfavored by e^{-3N}, where N is the number of e-folds (Gibbons et al., 1987; Gibbons and Turok, 2008) (for further discussion on whether inflation is generic, see also Hollands and Wald (2002); Kofman et al. (2002)). We avoid the complexities of the measure problem by assuming *a priori* that inflation occurs, considering only the relative weighting of one inflationary model to another; we are not attempting to see whether inflation itself is natural, but instead to compare the degrees of naturalness of different inflationary potentials.

To evaluate F, we must check whether the system satisfies conditions (1)–(5) for a given (Π_i, ϕ_i). In some cases, inflation will end in a false vacuum – a minimum other than $\phi = 0$. These trajectories violate condition (5) and so are discarded. In a flat, expanding, Friedmann-Robertson-Walker Universe, the Hubble parameter is monotonic and fixes the total energy. We can use this to determine when the inflaton will be unable to cross a maximum in the potential without actually evolving the equations of motion. If the system expands through 300 e-folds, yet the eventual destination is still indeterminate, because the total energy exceeds that at the maximum, we discard the trajectory. This situation is rare, and we impose this cut to avoid spending a disproportionate amount of effort evaluating a small set of trajectories. Finally, trajectories where ϕ is not monotonically increasing are discarded.

Computational limitations prevent us from considering every possible initial condition, so we limit ourselves to a small region ~ 60 e-folds before the end of inflation. The choice of the bounds on this region can change the numerical value of R_i significantly. However, our aim is to compare different models, not to provide an absolute indicator of how fine-tuned they are, so we need only enough variation in R_i to distinguish between them. To see how this works, consider the possible reasons a given set of initial conditions can fail to yield sufficient inflation. If there is a feature in the potential (an extremum, or a point where $\epsilon = 1$), most trajectories beyond it will be invalid. If the lower limit on ϕ is too close to zero, and a feature lies beyond it, a model would be unfairly favored. However, for large $|\phi|$, potentials are dominated by the ϕ^4 term and so features of this kind are rare. In practice then, extending the range of ϕ will usually add a non-inflating region to potentials with such a feature, and an inflating region to potentials without one, thus accentuating the difference in R_i.
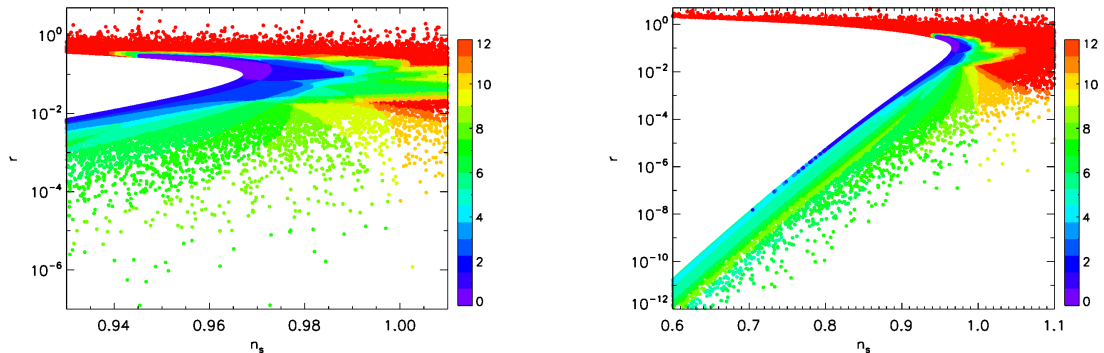
Figure 3.1 The BST fine-tuning criterion with quartic potentials. The color scale shows the value of Z_η in the (n_s, r) plane. The models were selected using a logarithmic prior for A and B. (Left) A limited range shown for comparison with BST. (Right) Z_η in an extended range. The tail of points with low fine-tuning and low r extends beyond this diagram. All points with $Z_\eta \geq 12$ are shown in the same color. See Section 3.3.4 for an explanation of the "forbidden zone" in the middle.

The other major cause of a failing trajectory is overshoot. This occurs when a large Π_i pushes the field through the inflationary region too quickly, leaving insufficient time on the slow-roll attractor to acquire sufficient e-folds of expansion before the inflaton starts oscillating, or driving the field beyond the oscillatory region altogether. This problem is present to some extent in all potentials. When $V \lesssim \Pi^2/2$, the evolution is dominated by the kinetic term; in this region the effect of the potential is negligible, and so evolution does not serve to differentiate between models. We therefore chose as an upper bound for Π the smallest value always within the kinetic dominated regime. We found that $V \lesssim 10^{-10} M_P^4$, so $\Pi_i < 10^{-4} M_P^2$. We also considered various other limits, with no changes in conclusions.

Since we ask that ϕ be increasing we need $\Pi_i \geq 0$. For ϕ_i, we take $\phi_i < \phi_{60}$, the point on the potential 60-e-folds before the end of inflation, as measured by the attractor solution. It is the point where n_s and r are calculated, and defines the smallest value of ϕ from which a valid trajectory can start. Our upper limit on ϕ_i is $\phi_{60} + 20M_P$, which is $\approx 2\phi_{60}$ for large-field models such as $V = \lambda\phi^4$. Our bounds are therefore:

$$\phi_{60} \leq \quad \phi_i \quad \leq (\phi_{60} + 20M_P)\,, \tag{3.16}$$
$$0 \leq \quad \Pi_i \quad \leq 10^{-4}M_P^2\,. \tag{3.17}$$

Finally, because F is evaluated on a finite grid, we ensure that our mesh is fine enough to reduce the resulting error in R_i to less than 1%.
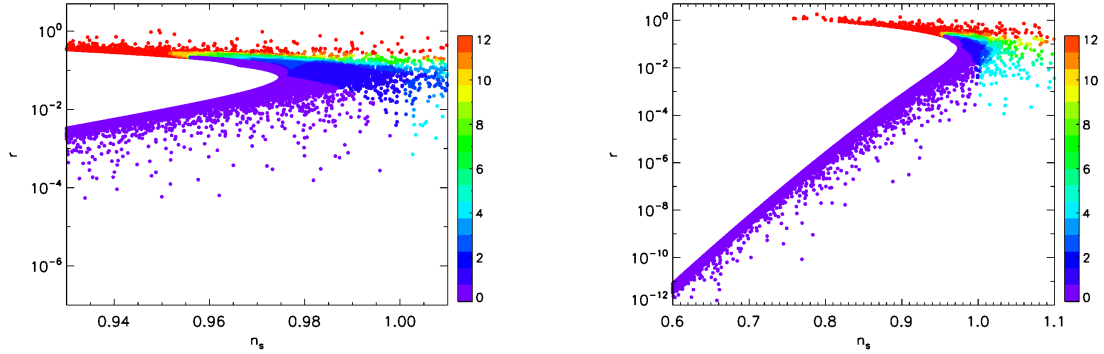
Figure 3.2 The BST fine-tuning criterion with cubic potentials, relaxing selection conditions (4) and (5). The color conventions are the same as those of Figure 3.1. The right panel shows an extended range compared to the left.
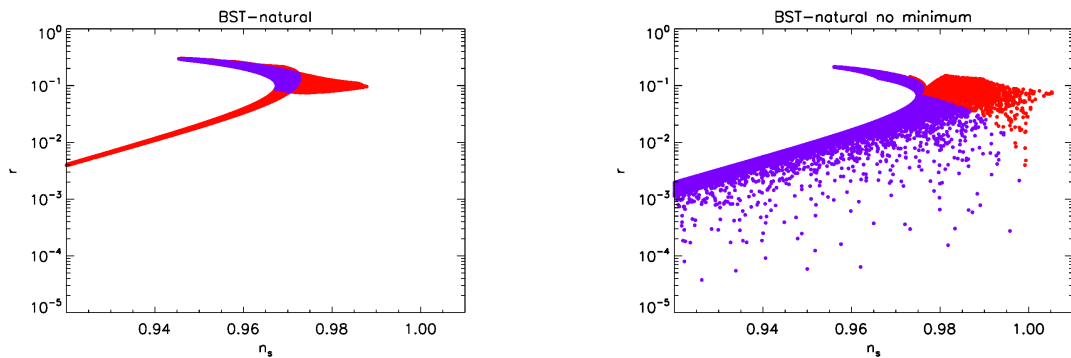
Figure 3.3 Potentials not fine-tuned by the BST definition. Purple points have $Z_\eta = 0$ and red points have $Z_\eta = 1$. (Left) Quartic potentials satisfying all the BST selection conditions. (Right) Cubic potentials relaxing selection conditions (4) and (5).

3.3 Results

3.3.1 BST Fine-tuning Criterion

As the left hand panel of Figure 3.1 shows, we recover Figure 1 of BST. The conclusion of BST that $n_s < 0.98$ has been weakened to $n_s < 0.99$; in addition, there are small differences in the shapes of some of the fine-tuning regions. These changes appear to be due to our use of ϕ derivatives instead of N derivatives when defining Z_η, and will not affect the following discussion. The right hand panel of Figure 3.1 shows a "tail" of untuned models with $r \ll 10^{-2}$, albeit at the cost of very red n_s, which is ruled out by data. This plot demonstrates that the BST criterion does not necessarily "predict" a *small* deviation from scale-invariance, but rather that models with low Z_η have correlated values of n_s and r. If n_s is at the lower limit of the currently permitted range (~ 0.94), then $r \sim 0.01$, but if $n_s \sim 0.97$, $r > 0.01$, and thus observable.

Figure 3.1 has regions with high Z_η for $r > 0.4$ and $r < 10^{-4}$. The corresponding potentials typically have a small region with decreased slope, or an inflection point. This is bracketed by two zeros of V'' (and thus η), which guarantees a high value of Z_η. The boost to Z_η is less efficient in low-r models because the inflaton moves more slowly and the second zero of η is encountered more than 60 e-folds before the end of inflation.

Having reproduced the essential features of BST's results, we now investigate what happens as the selection conditions are relaxed. First, the potential may have parameters which are irrelevant at the values of ϕ which determine n_s and r, but which are important as inflation comes to an end. In this case, we do not expect conditions (4) and (5) to hold, since we are only constraining the form of the potential as the cosmological perturbation spectrum is laid down. We also know that if we have a pure $\Lambda^4 - (\phi/\mu)^4$ potential, $n_s \sim 0.95$ while r can be very small. This potential has the form of equation (3.11) once we drop the positivity condition. We now relax conditions (4) and (5), and consider cubic potentials with a linear term, which includes the "inflection point" potentials found in some string-theoretic constructions (Baumann et al., 2007, 2008; Krause and Pajer, 2008; Linde and Westphal, 2008), with the form

$$V(\phi) = V_0(\phi^3 + A\phi^2 + B\phi + C). \tag{3.18}$$

Without loss of generality, we set $V(0) = 0$, and $V_0 < 0$, so that $V(\phi) \to -\infty$ as $\phi \to \infty$. Once condition (4) is abandoned, inflation is no longer tied to the deepest minimum, so there may be two regions satisfying the inflationary conditions, up to a rescaling of the potential so that the perturbation amplitude is correct. If there are two such regions, one will deposit the field in a (local) minimum, while the other will occur when ϕ is slowly rolling to infinity (or to $V \sim M_P^4$). In this case we include points for n_s and r obtained from both regions, and show our results in Figure 3.2. Strikingly, there are now non-fine-tuned points with a slightly red n_s and with $r < 10^{-4}$. These points occur in a region of the cubic potential with no following minimum.

The whole upper arm of Figure 3.2 contains points where, after inflation finishes, the inflaton ends in a minimum, while the lower arm contains points where it does not. The sharp delineation and curved feature at $n_s \approx 0.975$ and $r \approx 0.05$ marks the boundary between the two regimes, although it becomes much less sharp as n_s increases. It is important to note that the density of points on these planes is controlled by the assumed

priors on the distributions of potential parameters when drawing samples – what matters here is the presence of even a single non-fine-tuned point with small r. Thus the conclusion that large r is correlated with n_s for untuned models reached by applying BST is driven by its implicit assumption that the *endpoint of inflation* is fixed by the same parameters that determine the shape of the potential at ϕ_{60}.

This is more clearly seen in Figure 3.3 which selects out only the points which are not fine-tuned by the BST definition, i.e. the region for which $Z_\eta \leq 1$. The axes of this figure include the 3σ range of n_s allowed by the most recent cosmological data, and values of r which encompass both those accessible to next generation experiments ($r \sim 0.01$) as well as those which are effectively undetectable ($r \sim 10^{-3} - 10^{-5}$). The left panel of this figure, which reproduces the BST results for the full observationally-allowed range, shows that there exist natural models with $r < 10^{-2}$ at the redder end of this range. In BST's analysis, this region extends to $n_s \sim 0.98$, while on our graph the region extends to $n_s \sim 0.99$, due to our use of ϕ derivatives to compute Z_η, rather than N derivatives. In absolute terms, the difference in n_s is small, but the relative breaking of scale invariance differs by a factor of 2, which is a non-trivial distinction and is driven by a small and not unreasonable change in the specification of the fine-tuning criterion.

We also investigated the consequences of imposing the full set of BST selection conditions but adding higher order terms to the potential. With a ϕ^6 term we reproduce the results proposed by BST on qualitative grounds. The only significant new region with $Z_\eta < 2$ was within the formerly blank zone in the middle of the "arms" seen in the plots, and did not contain points with $r < 10^{-2}$. Note that these "forbidden zones" seen in the plots are not due to under-sampling, but arise due to our chosen restrictions on the functional form and shape of the potential. If n_s and r are observed within these zones, inflation cannot be described by a purely quartic potential, as discussed in Section 3.3.4. Figure 3 of Kinney (2002) shows a similar result, but the "forbidden zone" is simply populated more sparsely than the rest of the (n_s, r) plane. This is because they expand in derivatives of the Hubble parameter, instead of using an explicit form for the potential as is done in this work. The differences underscore our main conclusion; how generic an observable appears is often sensitive to details of the inflationary model.

3.3.2   Initial Conditions Fine-tuning Criterion

Figure 3.4 shows the equivalent plots to Figure 3.1, for the initial conditions fine-tuning criterion. The potentials used to generate the plots are the same as Figure 3.1. A model with smaller R_i is more fine-tuned.

We see two distinct regions; a "head" where $n_s \geq 0.96$, which has relatively high R_i, and a "tail" with relatively low R_i. In the head region, failed trajectories are dominated by overshoots, while in the tail the potential contains an extremum in the region of interest. Here, failed trajectories are dominated by those which originate on the far side of the maximum and have insufficient kinetic energy to pass over it. These potentials can appear natural in the BST criterion, because the feature occurs more than 60 e-folds from the end of inflation and does not contribute to Z_η. Figure 3.4 shows that the whole extended tail which the BST criterion considers natural has negligible R_i. This is clarified in Section 3.3.4; the only way to have both low n_s and low r is for ϕ_{60} to be near a maximum.
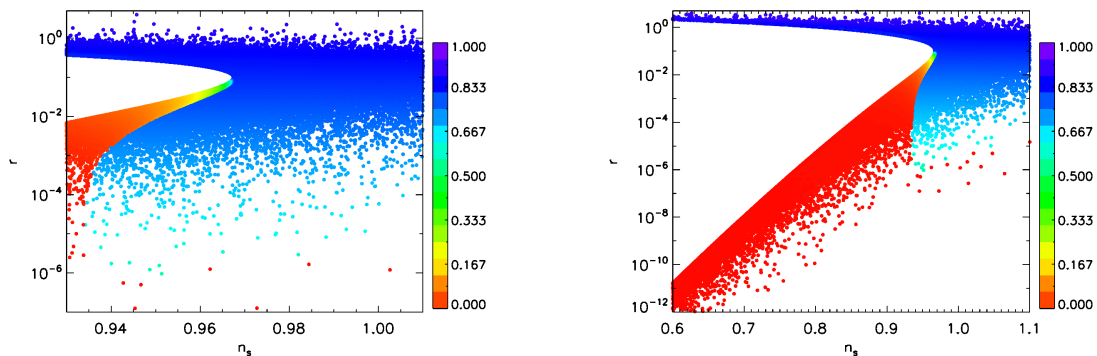
Figure 3.4 The initial conditions fine-tuning criterion for the same potentials as shown in Figure 3.1. The color scale shows the value of R_i in the (n_s, r) plane. Points with a lower R_i are more fine-tuned. (Left) selected region showing observationally relevant values of n_s. (Right) an extended range. See Section 3.3.4 for an explanation of the "forbidden zone" in the middle.

The head contains models which resemble the monomial potentials of chaotic inflation. Such potentials were examined in, e.g., Belinsky et al. (1985), and found to be largely insensitive to their initial conditions; our results support this conclusion. In these potentials, the failed trajectories are almost entirely overshoots. Note that R_i decreases slightly with r. This correlation is not as strong as that between the head and the tail, but it is still significant. For a fixed perturbation amplitude, a larger r implies a larger V' at ϕ_{60}, as the magnitude of V is fixed by condition (1). The system enters the slow-roll regime when $\Pi \sim V'$, so a larger V' will mean an earlier entry into slow-roll, and a corresponding decrease in the amount of overshoot.

The head continues down to $r \approx 10^{-6}$, and potentials with smaller r are very fine-tuned. V'/V must be tiny to achieve such a small r, and there is thus a minimum or inflection point in the vicinity. These are fine-tuned, first because they are sensitive to overshoot, and second because the inflection point can be followed by a steep slope, in which $\epsilon > 1$, prematurely ending inflation. See Chapter 4 for an investigation of the sensitivity of these potentials to initial conditions in the context of brane inflation.

To highlight the complementary information provided by the two fine-tuning criteria, we present two examples in Table 3.1 and Figure 3.5. Case 1 has low R_i, but is not fine-tuned by BST's criterion. Case 2 is fine-tuned according to the BST criterion, but has high R_i. Case 2 has a flat plateau near ϕ_{60}, which leads to the small value of r. Between the plateau and the end of inflation, there is a kink in the potential, during which η changes sign, leading to the high BST fine-tuning.

Figure 3.5 shows that trajectories can travel through three major regimes (Goldwirth and Piran, 1992): a "fast-roll" regime, where the kinetic energy dominates, a "slow-roll" regime, where inflation takes place, and an oscillatory regime. The oscillatory regime is entered as inflation ends, and does not concern us here[5]. Both fast and slow-roll regimes have $V' \ll 3H\Pi$. The condition for fast-roll is $\Pi^2 \gg V$, and for slow-roll, $\Pi^2 \ll V$. The

---

[5]A WKB approximation shows that it begins when $9H^2 \approx 4V''$.

Parameter	Case 1	Case 2
n_s	0.952	0.963
r	2.34×10^{-2}	3.65×10^{-3}
Z_η	1	7
R_i	0.137	0.781
V_0	8.19×10^{-18}	2.38×10^{-17}
A	777.0	148.8
B	55.6	22.86
ϕ_{60}	-11.4	-7.959
ϕ_e	-1.34	-1.264

Table 3.1 Examples of potentials where the two fine-tuning criteria lead to different conclusions. Inflation ends at ϕ_e, when $\epsilon = 1$.
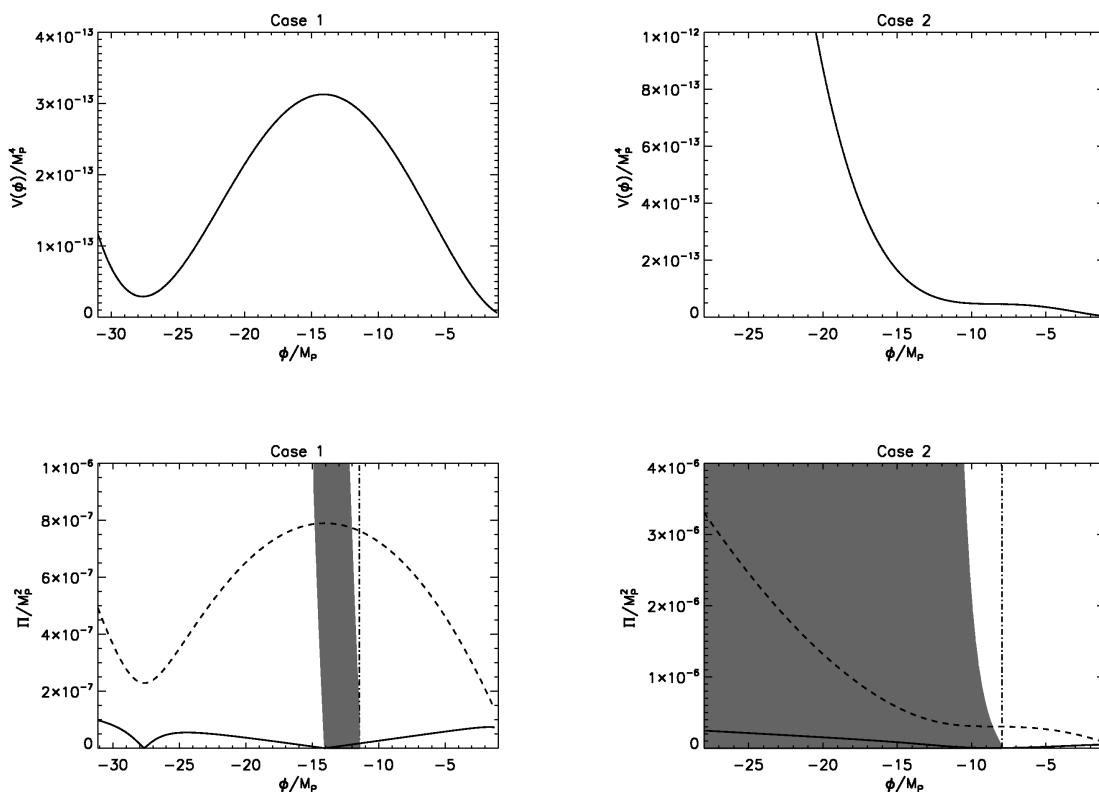
Figure 3.5 (Top) The potentials for Case 1 (left) and Case 2 (right) – see text and Table 3.1. The last 60 e-folds of inflation take place between $-11.4 \lesssim \phi \lesssim -1.3$ for Case 1, and between $-7.96 \lesssim \phi \lesssim -1.26$ for Case 2. (Bottom) The gray region shows the portion of the initial conditions which lead to sufficient inflation in Case 1 (left) and Case 2 (right), while the white region shows a portion which does not. The dashed line shows the boundary between the slow-roll regime and the fast-roll regime, while the solid line shows the slow-roll attractor, and the dot-dashed line shows ϕ_{60}. Only a subset of the Π range tested is shown.

equations of motion are, in the fast-roll regime:

$$\frac{d\Pi}{dN} \sim -3\Pi \qquad \frac{d\phi}{dN} \sim 1, \tag{3.19}$$

and in the slow-roll regime:

$$\frac{d\Pi}{dN} \sim -3\Pi \qquad \frac{d\phi}{dN} \sim \frac{\Pi}{\sqrt{V}}. \tag{3.20}$$

In both regimes, Π is drawn strongly towards the slow-roll attractor. In slow-roll, ϕ rolls gently down the potential, while in fast-roll ϕ increases quickly, but little expansion occurs. If we return to our specific example potentials, trajectories with high initial kinetic energy decelerate rapidly towards the inflationary attractor. In Case 2, if the attractor is reached by a given trajectory, the inflaton will roll smoothly down the potential, yielding sufficient e-folds of inflation. Overshoot trajectories appear when there is sufficient kinetic energy to prevent the system reaching the slow-roll phase with a negative enough ϕ. In Case 1, the system must exit the fast-roll regime in a specific range of ϕ in order to achieve sufficient inflation. If slow-roll starts when $\phi < -15M_P$, the inflaton will be on the wrong side of the maximum, unable to reach $\phi = 0$. If it enters the slow-roll regime with $\phi \gg \phi_{60}$, it will overshoot. The initial Π must therefore be such that the kinetic energy dissipates when $-11M_P < \phi < -15M_P$. This (relatively) narrow range causes the potential to appear fine-tuned.

3.3.3 Common Inflationary Potentials

In addition to the analysis of the quartic potential of equation (3.12), it is interesting to see what fine-tuning criteria say about standard potentials. We computed the values of Z_η and R_i for a set of these, choosing their free parameters to produce observables in agreement with the experimental bounds in Komatsu et al. (2009). The models we investigated are chaotic inflation (Linde, 1983), new inflation (Linde, 1982b; Guth and Pi, 1982), power-law inflation (Lucchin and Matarrese, 1985), hybrid inflation (Linde, 1994), natural inflation (Freese et al., 1990), the Coleman-Weinberg potential (Albrecht and Steinhardt, 1982; Linde, 1982a) and two hilltop models (Easther et al., 2006; Efstathiou and Chongchitnan, 2006). Our results are shown in Table 3.2. Our results agree with similar calculations performed for chaotic (Belinsky et al., 1985) new (Goldwirth, 1990) and natural (Knox and Olinto, 1993) inflation.

For chaotic inflation (and power-law / natural inflation, which have similar dynamical properties), all the fine-tuning criteria agree: they are "natural" in terms of symmetry considerations, turning points, and insensitivity to initial conditions. New inflation and hilltop type models are sensitive to their initial conditions in a way that is not accounted for by the BST fine-tuning criterion. For parameter choices where the log term is significant, Coleman-Weinberg potentials appear fine-tuned in both criteria as the $\log \phi$ term generates a high Z_η, while their sensitivity to initial conditions is well-known (Goldwirth and Piran, 1992). For a sufficiently large value of μ, the log term becomes unimportant and the potential resembles that of chaotic inflation. However, these potentials may be technically natural, in the sense that they are stable against (further) loop corrections.

Model	Potential	μ	n_s	r	Z_η	R_i
Chaotic	$V_0\phi^\mu$	4	0.949	0.273	0	0.868
		2	0.966	0.135	0	0.861
New	$V_0\left[(\phi/\mu)^2 - 1\right]^2$	15	0.955	0.0287	1	0.168
Power-law	$V_0\exp(-\phi\sqrt{2/\mu})$	120	0.983	0.133	0	1.0
Natural	$V_0\left[1 + \cos\left(\phi/\mu\right)\right]$	10	0.966	0.0976	0	0.831
Hybrid	$V_0\left[1 + (\phi/\mu)^2\right]$	1	0.966	0.137	12	0.864
		100	1.01	9.0×10^{-4}	0	0.761
Hilltop	$V_0\left[1 - (\phi/\mu)^2\right]$	10	0.954	0.011	0	0.010
		8.8	0.945	0.007	0	0.06
Hilltop	$V_0\left[1 - (\phi/\mu)^4\right]$	4	0.954	2.2×10^{-4}	1	0.001
		20	0.970	0.018	0	0.60
Stabilized Hilltop	$V_0[C - (\phi/\mu)^4 + 0.01(\phi/\mu)^8]$	0.5	0.955	1.3×10^{-6}	1	0.00
		6	0.963	0.012	3	0.344
Coleman-Weinberg	$V_0[1/4 + (\phi/\mu)^4 \times (\ln\left(\phi/\mu\right) - 1/4)]$	8	0.947	0.0013	10	0.06
		30	0.967	0.057	1	0.83

Table 3.2 Fine-tuning criteria for specific potentials. For the hybrid potentials, we are neglecting the second field until $\phi = -\mu$, at which point we assume inflation ends abruptly. Thus, we are not applying BST selection conditions (3)–(4). Further, we are not applying BST selection condition (4) to the Hilltop potentials. Lower r for the quadratic Hilltop potential comes at the cost of a redder n_s than allowed by observations. Finally, in the case of the natural potential, to compare our μ with the f of Freese et al. (1990), divide by $\sqrt{8\pi}$.

The Coleman-Weinberg potential thus has one parameter, and in the absence of a microscopic theory, no value of μ is favored over any other. However, both Z_η and R_i change significantly between $\mu = 8$ and $\mu = 30$. Thus the tuning criteria will penalize the potential for not being in the "chaotic" regime, even though the underlying physics has not changed.

Hybrid inflation, which can be highly fine-tuned by the BST criterion, in addition to lacking a deterministic end to inflation (since the endpoint of inflation is fixed by the coupling between ϕ and a second field, which is not included in this potential), is very insensitive to initial conditions. In these cases, the BST trend is reversed; models with low r are less fine-tuned. A hybrid model with $\mu \gg 1$ has $V \sim V_0$, $\eta \propto 1/\mu^2$, $Z_\eta = 0$, $n_s \sim 1$, and unobservable r. Conversely, the hybrid case with $\mu \ll 1$ resembles chaotic ϕ^2 inflation, and hence has $Z_\eta = 0$. The intermediate regime $\mu \sim 1$ balances the two asymptotic forms of V and hence has high Z_η, even though r is large and $n_s \sim 0.96$.

Finally, we consider a potential which is stabilized by a higher order term, but is constructed to meet all five of the BST conditions,

$$
V = V_0 \left[C - \left(\frac{\phi}{\mu} \right)^4 + \delta \left(\frac{\phi}{\mu} \right)^8 \right] . \tag{3.21}
$$

The value of C is fixed by condition (4), with $V = 0$ at the minimum, and δ is assumed to be small (10^{-2}).[6] For $\mu < 0.8$ this potential has $Z_\eta = 1$, together with a very low value of r. It provides a counter-example to BST's claim that natural potentials require $r > 10^{-2}$. It is, however, very fine-tuned with respect to initial conditions. Higher values of μ do not involve fine-tuned initial conditions but have a larger r, as well as a large Z_η.

3.3.4 The Forbidden Zone

The blank "forbidden zone" in Figures 3.1-3.4 is not an artifact of the selection priors; it is inaccessible for quartic potentials. Higher order potentials have a smaller forbidden zone, and cubic potentials a larger one (see Figure 3.2). In order to acquire some intuition about why this is the case, we shall now attempt to minimize $n_s - 1$ for fixed r. In what follows, $M_P = 1$. Recall that:

$$
n_s = 1 + 2\eta - 6\epsilon \tag{3.22}
$$

$$
\approx 1 + 2\frac{V''}{V} \qquad \text{(for small } r\text{)}. \tag{3.23}
$$

Thus, $r < 10^{-2}$ implies $\epsilon \lesssim 10^{-3}$, so if in addition $n_s \sim 0.9$, V'' is negative and reasonably large. Now:

$$
V''(\phi_{60}) < 0 \implies 12\phi_{60}^2 + 6B\phi_{60} + 2A < 0 . \tag{3.24}
$$

Since $\phi < 0 < A$, this implies $B > 0$. If V' is small and V'' is large, a zero of V' must be nearby. So, for large $n_s - 1$ and small r, ϕ_{60} will be fairly near a maximum of the potential (ϕ_m)[7]. Furthermore, $\phi_{60} > \phi_m$, because $N \to \infty$ as $\phi_{60} \to \phi_m$. Therefore, we

---

[6]In the original formulation of the BST criterion, polynomial potentials of degree $n > 4$ have $Z_{\text{total}} = Z_\eta + n - 4$, but we will work with a Z_η derived solely from the shape of the potential.

[7]If we desired both η and ϵ to be small, we could achieve this with a flat potential and no necessity for a maximum. This conclusion is borne out by Figure 3.4, where the high R_i for points with $n_s \sim 1$ and small r show that no maxima are nearby.

shall consider all relevant quantities evaluated at ϕ_m, Taylor-expanding around this point when the description is inadequate.

If a maximum exists, it occurs at:

$$\phi_m = \frac{-3B + \sqrt{9B^2 - 32A}}{8}. \tag{3.25}$$

For $A > 0$, this is a strictly decreasing function of B. Recall that $B^2 < 4A$, due to condition (5), so

$$\phi_m > -\frac{\sqrt{A}}{2}. \tag{3.26}$$

A lower bound on V'', therefore, is:

$$
\begin{aligned}
V'' &> V_0(12\phi_m^2 + 6B\phi_m + 2A) \\
&> V_0(12\phi_m^2 + 12\sqrt{A}\phi_m + 2A) \\
&> V_0(3A - 6A + 2A) \\
&> -V_0 A.
\end{aligned}
\tag{3.27}
$$

Since this minimal V'' is at a maximum of V, it also locally minimizes η, so:

$$
\begin{aligned}
V &= \frac{V_0 A^2}{16} \\
\eta &> -\frac{16}{A}.
\end{aligned}
\tag{3.28}
$$

We cannot simply evaluate V' at ϕ_m, because this would tell us that $r = 0$. The leading order of the Taylor expansion is significant here. We compute it via

$$
\begin{aligned}
V' &\approx V'(\phi_m) + V''(\phi_m)\frac{(\phi_{60} - \phi_m)}{2} \\
&\approx (\phi_{60} - \phi_m)\frac{V_0 A}{2}
\end{aligned}
\tag{3.29}
$$

$$\implies \epsilon \approx \frac{(\phi_{60} - \phi_m)^2}{2}\left(\frac{16}{A}\right)^2 \tag{3.30}$$

$$\implies r \approx \frac{2048}{A^2}\left(\phi_{60} + \frac{\sqrt{A}}{2}\right)^2. \tag{3.31}$$

Here our simple analytic approximations can go no further, because ϕ_{60} is a non-linear function of both A and B. However, when $r \sim 10^{-2}$, numerical calculations show that $\phi_{60} - \phi_m \sim 1$. Case 1 (see Figure 3.5 and Table 3.1) has $\phi_{60} - \phi_m \approx 2$. With this final approximation, we obtain:

$$n_s \gtrsim 1 - \sqrt{\frac{r}{2}}. \tag{3.32}$$

Even when the final rough fit is no longer valid, it can be seen that on the lower arm of Figure 3.1, reaching low n_s requires a very low r, and the potential which minimizes n_s for given r will contain a maximum.

For large r, this lower bound on n_s ceases to hold, both because ϵ begins to become significant in the expression for n_s and because ϕ_{60} need no longer be in proximity to a maximum. These models tend to involve a large variation in ϕ during inflation, $\eta \ll \epsilon$, and positive V''. Hence a better approximation for the minimum n_s is:

$$n_s \gtrsim 1 - \frac{3r}{8} \,. \tag{3.33}$$

Matching these two limits together, we obtain a rough description of the forbidden zone.

3.4 Conclusions

We have investigated whether, given the observed limits on the scalar tilt n_s, it is possible to claim that certain ranges of values for the tensor-scalar ratio r are "fine-tuned" in a way that is independent of the definition of fine-tuning. We find that, on the contrary, different criteria are sensitive to different types of fine-tuning, and hence can give differing answers about the naturalness of a given potential. The BST criterion, based on the number of unnecessary "features" in the potential, tells us that models with $r < 10^{-2}$ are fine-tuned for $n_s > 0.95$, while an alternative criterion based on sensitivity to initial conditions shows that models with $n_s < 0.94$ and $r < 10^{-3}$ are fine-tuned, but with $n_s > 0.94$ there are "natural" models with $r \sim 10^{-6}$.

Moreover, BST's conclusions are not robust against relatively minor modifications to their criteria. In particular, the claim that models with $n_s > 0.98$ are fine-tuned is softened by a small technical change in the fine-tuning criterion Z_η: our upper limit on natural values of the spectral index is $n_s < 0.99$. This may seem like a fine distinction, but recall that the breaking of scale invariance is measured by $n_s - 1$, and thus differs by a factor of 2. Moreover, Planck is expected to measure n_s with a 1-σ error of 0.0045 (Planck Collaboration, 2005) and $n_s = 0.99$ cannot be easily distinguished from $n_s = 1$, but $n_s = 0.98$ could be detected with considerable confidence. Likewise, BST's conclusion that $n_s > 0.95$ implies $r > 10^{-2}$ in the absence of fine-tuning only holds if the potential evolves smoothly to an analytic minimum with $V(0) \approx 0$. This is equivalent to requiring that the parameters fixing the shape of the potential as cosmological perturbations are laid down also govern the shape of the potential as inflation ends. While it is not inherently unreasonable, this restriction immediately excludes both hybrid inflation and many stringy constructions, even though their underlying potentials may be technically natural from a particle physics perspective. Pending a better understanding of physics at the energy scale of inflation, this condition therefore seems overly restrictive.

Our alternative criterion is based on the sensitivity of a particular potential to the inflaton's initial conditions. It shows that potentials with $n_s < 0.94$ and $r < 10^{-3}$ are very sensitive to their initial conditions, as are all models with $r < 10^{-6}$. Since inflation was originally designed to solve the initial condition problems of a hot big bang, one should certainly check that inflation itself does not require the Universe to be in a special initial state. For our "scoring" of the initial conditions problem, there is a small preference for models with higher r, but the *specific* results we find depend on the explicit construction of our fine-tuning criterion, so we do not see this as a robust conclusion. Rather, we emphasize that this criterion singles out a different set of potentials from the
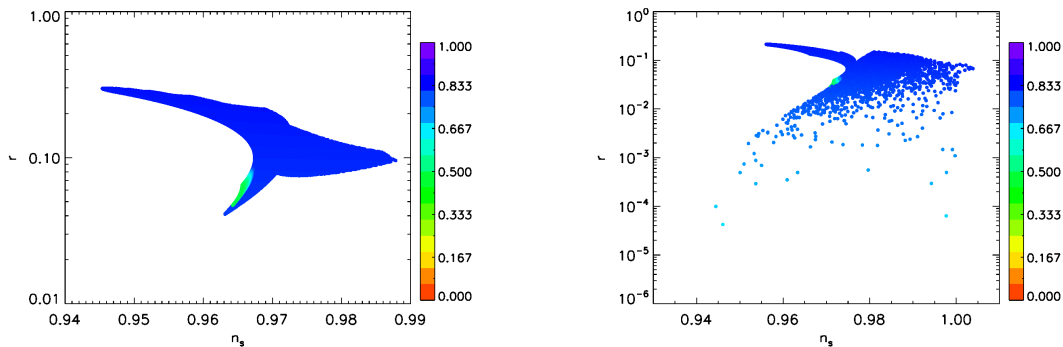
Figure 3.6 The set of models which are deemed natural by both the BST and initial conditions criteria, i.e. $Z_\eta < 2$ and $R_i > 0.5$. The color bar shows R_i. (Left) Imposing all five BST conditions. (Right) Neglecting selection conditions (4) and (5). Potentials which have low $r \sim 10^{-4}$ and $n_s \sim 0.95$ generically contain a flat plateau terminating in a drop, similar to Case 2 in Figure 3.5.

BST criterion, and neither set of potentials is a strict subset of the other. The above conclusions apply to potentials that are well-described by a quartic potential.

We have shown that conclusions drawn from BST-like criteria depend on the explicit assumptions they encode. We also discussed a second type of fine-tuning – sensitivity to the initial field configuration – which is distinct from that described of BST. One may restrict attention to models deemed acceptable by *both* criteria. The left panel of Figure 3.6 indicates that, at face value, this would lead one to conclude that the "natural" lower limit on the tensor amplitude is $r \gtrsim 0.04$, and thus well within the range that is experimentally accessible. However, the right panel shows that this conclusion only holds if condition (4) is imposed. Moreover, there are types of fine-tuning missed by our initial conditions criterion (e.g. initial inhomogeneity or spatial curvature); including these would further reduce the number of models for which inflation can be regarded as generic.

The BST criterion aims to characterize the fine-tuning of the *shape* of the potential. The initial conditions criterion might, at face value, appear to measure only the sensitivity to an initially large kinetic term. However, it is important to realize that it also encodes information about the functional form of the potential: for instance, a potential with a steep hill, a small plateau, and another steep downhill stretch will have difficulty supporting inflation without carefully tuned initial conditions. Thus, R_i penalizes the functional form of the potential, but in a different way to Z_η. Hence, Occam's razor cannot characterize fine-tuning (even of a single property, such as the shape of the potential) in an unequivocal way. These ambiguities will remain after the arrival of data with sufficient precision to select a small region of the (r, n_s) plane. If this region is disfavored according to a specific fine-tuning criterion, should we interpret this as evidence against inflation, or even (in a more limited sense) as evidence against a class of inflationary models? Absent a detailed understanding of the physics of inflation, the answer is unclear, since we would not know whether the fine-tuning identified by that criterion is relevant to the underlying theory.

As is well known (Lyth, 1997), simple single field models of inflation have a direct correlation between the total excursion of the inflaton field and the value of r. If values of $r > 10^{-3}$ are ruled out by experiment, this would also eliminate all single field models with $\Delta\phi > M_P$, with dramatic consequences for inflationary model building, as well as for alternative theories of the early Universe (Khoury et al., 2002; Boyle et al., 2004). Many proposed inflationary models predict signals in this range but only a handful of models live in the lower part, the so-called "tensor desert" at $10^{-3} < r < 10^{-2}$ (Smith et al., 2008; Alabidi, 2007; Efstathiou and Chongchitnan, 2006). Conversely, many other models have $r \ll 10^{-3}$. Consequently, it is clear that the inflationary paradigm would not be ruled out by a failure to detect primordial tensors, although a large and important class of models would be conclusively eliminated.

Our experience with different fine-tuning criteria suggests that it is likely to be impossible to construct a set of criteria which defines a "generic inflationary model" that is simultaneously (a) robust against small changes, (b) does not eliminate apparently reasonable models from consideration, and (c) leads to a definitive conclusion for the likely value of the tensor-scalar ratio r. Rather, the value of analyses like that of BST lies in the extent to which they identify a category of inflationary models which would be falsified if future experiments fail to detect tensors. However, determining that $r \lesssim 0.001$ – or any other value likely to be experimentally accessible – cannot be construed as a definitive test of inflation.

"Alice sighed wearily. 'I think you might do something better with the time,' she said, 'than waste it in asking riddles that have no answers."'

– Alice in Wonderland, by Lewis Carroll

4

Brane Inflation and the Overshoot Problem

Abstract

We investigate recent claims that brane inflation solves the overshoot problem through a combination of microphysical restrictions on the phase space of initial conditions and the existence of the Dirac-Born-Infeld attractor in regimes where the slow-roll attractor does not apply. Carrying out a comprehensive analysis of the parameter space allowed by the latest advances in brane inflation model-building, we find that these restrictions are insufficient to solve the overshoot problem. The vast majority of the phase space of initial conditions is still dominated by overshoot trajectories. We present an analytic proof that the brane-inflationary attractor must be close to the slow-roll limit, and update the predictions for observables such as non-Gaussianity, cosmic string tension and tensor modes.

4.1 Introduction

The inflationary paradigm postulates a period of nearly exponential expansion in the early Universe. This ameliorates various initial conditions problems of the classical hot big bang model and explains the smoothness of the Universe when averaged on large scales. Furthermore, according to inflation, initial inhomogeneities originated as quantum fluctuations in the inflaton field and the metric, sourcing scalar (Hawking, 1982; Guth and Pi, 1982; Starobinsky, 1982; Bardeen et al., 1983) and tensor (Grishchuk, 1975; Starobinsky, 1979) perturbations which are adiabatic, scale-invariant and almost Gaussian. (Bardeen et al., 1983). These broad-brush predictions of inflation are consistent with the latest observations (Komatsu et al., 2009; Peiris and Easther, 2006a; Kinney et al., 2006; Peiris and Easther, 2006b; Martin and Ringeval, 2006; Alabidi and Lyth, 2006; Lesgourgues and Valkenburg, 2007; Peiris and Easther, 2008; Alabidi and Lidsey, 2008; Kinney et al., 2008; Baumann et al., 2009b).

However, while acknowledging the substantial successes of inflation as a phenomenological model of the early Universe, several theoretical challenges remain. Chief among these are the lack of a single compelling particle physics motivation for inflation and an understanding of the initial conditions leading to successful inflation. Attempts have recently been made to address both of these challenges within the context of string theory (see Baumann and McAllister (2009); Linde (2006); Cline (2006); Burgess (2007); Tye (2008); Kallosh (2008); McAllister and Silverstein (2008) for recent reviews). In this Chapter, we shall investigate 'warped brane inflation' (Kachru et al., 2003), one of the most well-developed string theory models of inflation. Here the inflaton field is identified with the position of a D-brane in a warped higher-dimensional space, with effective four-dimensional dynamics governed by the Dirac-Born-Infeld (DBI) action, and an inflaton potential that is in principle completely derivable from fundamental theory (Baumann and McAllister, 2009).

For inflation to be insensitive to initial conditions one typically hopes that the inflaton field will naturally flow toward an attractor solution (Belinsky et al., 1985; Felder et al., 2002; Underwood, 2008), on which the inflaton slow-rolls and the Universe inflates. However, should the energy of the inflaton be initially kinetic dominated, it is possible for the end of inflation to be reached before the solution has approached the attractor. The statement that a large initial kinetic energy spoils inflation is often called the *overshoot problem* (Brustein and Steinhardt, 1993). Due in part to somewhat broad limits on the initial kinetic energy in models of inflation based on effective field theory, this occurs for a large proportion of possible inflationary trajectories. This poses a theoretical problem, because one of the central aims of inflation is to explain the special initial conditions required for the hot big bang. For inflationary models suffering from the overshoot problem, it seems like we have simply exchanged one initial conditions problem for another.

The overshoot problem seems particularly acute for small-field models of inflation, where the inflationary region only occupies a small area of phase space. The brane inflation models that we discuss in this Chapter are of the small-field type (Baumann and McAllister, 2007) and so seem susceptible to a generic overshoot problem. However, Underwood (2008) recently observed that brane inflation may not suffer from the overshoot problem, based on a calculation for a specific set of model parameters. This was due to a combination of two effects: a restriction on the allowable initial conditions due to microphysical consistency conditions, and an increase in the domain of applicability of the attractor trajectory due to the DBI action. Since this claim constitutes an important theoretical argument supporting brane inflation, we will check it for the full range of parameters, including those identified by recent advances in string theory model-building (Baumann et al., 2009a, 2006; Baumann et al., 2008).

Our treatment closely parallels the effective field theory analyses of the overshoot problem that have been performed for most of the classic inflationary models (*e.g.* Belinsky et al. (1985); Piran and Williams (1985)). Related work on the parameter space for brane inflation models consistent with the latest experimental bounds has appeared in Bean et al. (2007); Peiris et al. (2007); Bean et al. (2008); Gmeiner and White (2008); Lorenz et al. (2008); Alabidi and Lidsey (2008); Agarwal and Bean (2009); Tzirakis and Kinney (2009); Chen and Gong (2009); Hoi and Cline (2009); Cline et al. (2009). Here, we add a systematic study of the overshoot problem and the fine-tuning of inflationary initial conditions for brane inflation.
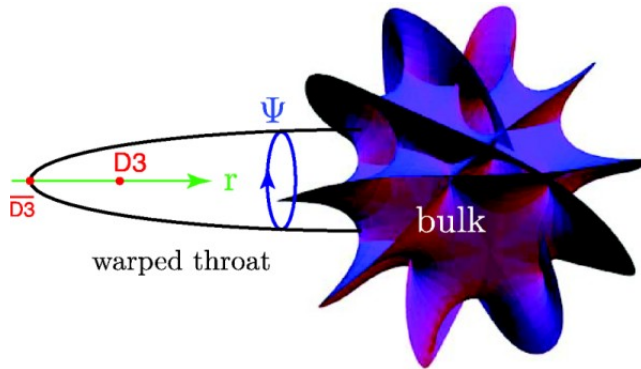
Figure 4.1 Warped D3-brane inflation (figure reproduced from Baumann and McAllister (2009)). The 3-branes fill the four-dimensional spacetime and are pointlike in the extra dimensions (shown in this figure). The *warped throat* has a radial dimension r and a five-dimensional base with coordinates Ψ. At r_{UV} the throat attaches to the bulk space. This introduces a cut-off to the throat solution, making the total space finite. The radial coordinate of the D3-brane in the warped throat plays the role of the inflaton field.

We perform a comprehensive analysis of the entire parameter space, and determine whether, in general, these models solve the overshoot problem. We find that they do not; almost all possible trajectories are still dominated by overshoots. Furthermore, we reiterate two generic predictions of these models already given in Firouzjahi and Tye (2005); Bean et al. (2007): a potentially observable level of cosmic strings generated after the end of inflation and negligible levels of primordial non-Gaussianity. Finally, we examine which regions of the parameter space seem to require less tuning of the initial conditions, and consider the extent to which microphysical restrictions on the phase space of initial conditions would have to tighten before the overshoot problem for brane inflation can be solved.

4.2 D-brane Inflation

4.2.1 D3-brane motion in warped geometries

Since the original brane inflation proposal by Dvali and Tye (1999) many variants of inflation sourced by D-brane motion through a higher-dimensional spacetime have been studied, *e.g.* Dvali et al. (2001); Alexander (2001); Burgess et al. (2001, 2002); Garcia-Bellido et al. (2002); Jones et al. (2002); Shiu and Tye (2001); Burgess et al. (2004); DeWolfe et al. (2004); Iizuka and Trivedi (2004); Burgess et al. (2007); Pajer (2008); Krause and Pajer (2008). In this work we consider a specific, and, arguably, the best understood scenario: D3-brane motion in warped throat geometries (Kachru et al., 2003).

In these scenarios inflation is driven by a mobile D3-brane which fills the four-dimensional spacetime $\mathbb{R}^{3,1}$ (with metric $g_{\mu\nu}$), and is pointlike in the extra-dimensional space M_6 (with metric g_{mn}). The ten-dimensional background spacetime is defined by

the following warped line element

$$ds^2 = h^{-1/2}(y)g_{\mu\nu}dx^\mu dx^\nu + h^{1/2}(y)g_{mn}dy^m dy^n \,, \tag{4.1}$$

where the warp factor $h(y)$ is a function of the coordinates on the internal space M_6. Local approximations to such warped spaces are ubiquitous in the flux compactifications of type IIB string theory[1] (Douglas and Kachru, 2007). Exact solutions to the supergravity equations may be found (Klebanov and Strassler, 2000) that take the structure of a cone for the extra dimensions

$$g_{mn}dy^m dy^n = dr^2 + r^2 ds_{X_5}^2 \,, \tag{4.2}$$

where r parametrizes the radial direction and X_5 is a certain base manifold, for example, a 5-sphere, S^5, or the Einstein space $T^{1,1}$ (arising as the limit of the Klebanov-Strassler geometry). These have $\mathrm{Vol}(S^5) = \pi^3$ and $\mathrm{Vol}(T^{1,1})= \frac{16}{27}\pi^3$, respectively. The warp factor at large r is approximately

$$h(r) \approx \left(\frac{R}{r}\right)^4 . \tag{4.3}$$

In this limit the spacetime (4.1) becomes $AdS_5 \times X_5$. We define the constant curvature (or warping) scale

$$\frac{R^4}{(\alpha')^2} \equiv 4\pi g_s N \frac{\pi^3}{\mathrm{Vol}(X_5)} \,, \tag{4.4}$$

in terms of the string coupling g_s, the string length $\ell_s = \sqrt{\alpha'}$, a flux integer N and the dimensionless volume of X_5 with unit radius. $\mathrm{Vol}(X_5)$ generically takes a value of order π^3.

The motion of a D3-brane in this warped geometry is governed by the DBI action

$$S = -\int d^4x \sqrt{-g}\left[\frac{M_{\mathrm{pl}}^2}{2}\mathcal{R} + P(X,\phi)\right] \,, \tag{4.5}$$

where $\mathcal{R}$ is the four-dimensional Ricci scalar and

$$P(X,\phi) = f(\phi)^{-1}\left(\sqrt{1 - 2f(\phi)X} - 1\right) + V(\phi) \,. \tag{4.6}$$

Here we have defined the canonical inflaton $\phi \equiv r\sqrt{T_3}$ in terms of the brane tension

$$T_3 = \frac{1}{(2\pi)^3}\frac{1}{g_s}m_s^4 \,, \tag{4.7}$$

where $m_s \equiv 1/\sqrt{\alpha'}$ is the string scale. The quantity $X \equiv -\frac{1}{2}g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi \approx \frac{1}{2}\dot\phi^2$ is the corresponding kinetic term, and $V(\phi)$ is the inflaton potential, whose derivation we shall sketch in §4.2.2. The function $f(\phi) \equiv T_3^{-1}h(\phi)$ is the (rescaled) warp factor in terms of the canonical inflaton

$$f(\phi) = \frac{\lambda}{\phi^4} \,, \tag{4.8}$$

---

[1] Readers unfamiliar with this terminology may find a useful Stringlish-to-English dictionary in Hertzberg et al. (2007).

with λ given by

$$\lambda \equiv \frac{\pi}{2} \frac{N}{\mathrm{Vol}(X_5)}\,. \tag{4.9}$$

In the slow-roll limit, $f(\phi)X \ll 1$, the DBI action (4.6) reduces to the canonical slow-roll action

$$P(X,\phi) = -X + V(\phi)\,. \tag{4.10}$$

Slow-roll corresponds to non-relativistic motion of the brane. The non-slow-roll limit associated with relativistic motion of the brane is characterized by a Lorentz-like factor

$$\gamma \equiv [1 - f(\phi)\dot{\phi}^2]^{-1/2}\,. \tag{4.11}$$

For $\gamma \gg 1$ inflation can occur even for relatively steep potentials. However, large values of γ correspond to small values of the effective speed of sound associated with inflaton fluctuations. This leads to large non-Gaussianity in the primordial scalar fluctuations, so that the parameter region of viable relativistic DBI models is significantly constrained by the observational data (Baumann and McAllister, 2007).

The Hubble parameter, H, is given by the sum of potential and kinetic contributions

$$H^2 = \frac{1}{3M_{\mathrm{pl}}^2}\left[V(\phi) + \frac{\gamma - 1}{f}\right]\,. \tag{4.12}$$

4.2.2 Parametrization of the inflaton potential

The computation of the potential $V(\phi)$ for D3-brane inflation has a long and painful history (see Baumann and McAllister (2009) for a recent review). Approximations commonly employed in string theory computations have proved to be insufficient when trying to capture all effects relevant for inflation.

This challenge is illustrated by the *eta problem*: small corrections, $\Delta V(\phi)$, to the inflationary potential often lead to large corrections to the second derivative of the potential or the inflaton mass, causing inflation to end prematurely. This is an example of the UV sensitivity of inflation where high-energy effects, in this case dimension-6 Planck-suppressed corrections, affect the inflationary dynamics.

We now give a brief summary of all relevant contributions to the D3-brane potential (more details may be found in Baumann and McAllister (2009)). At zeroth order, the inflaton potential is given by the Coulomb potential between a D3-brane and an anti-D3-brane,

$$\hat{V}_0(\phi) = D\left(1 - \frac{3D}{16\pi^2\phi^4}\right)\,, \tag{4.13}$$

where the constant D sets the scale of the brane-antibrane interaction

$$D \equiv \frac{2T_3}{h_{\mathrm{IR}}}\,. \tag{4.14}$$

Here $h_{\mathrm{IR}} \equiv h(\phi_{\mathrm{IR}})$ is the warp factor at the tip of the throat (see Fig. 4.1). This warped scale D exponentially suppresses the D3-$\overline{\mathrm{D3}}$ interaction, making the potential (4.13) extraordinarily flat. In addition there may be distant energy sources in the bulk space,

which shift the scale of the potential but do not significantly affect the slope or the curvature, such that

$$\hat{V}_0(\phi) = D\left(s - \frac{3D}{16\pi^2\phi^4}\right), \tag{4.15}$$

where $s \gtrsim 1$ parametrizes the contribution of bulk energy to the cosmological constant during inflation.

So far, the potential is flat, and eminently suitable for supporting a large amount of inflation. In particular, the inflaton mass is small and $\eta \ll 1$. However, we have ignored crucial interactions with moduli-stabilizing ingredients. Kachru et al. (2003) identified that the leading moduli-stabilization effect leads to a generic eta problem by adding a $H^2\phi^2$ term to the potential, which makes $\eta = \frac{2}{3}$. However, they also pointed out that further corrections to the potential are expected from additional moduli-stabilization effects and from coupling of the throat to the bulk geometry. It has recently become possible to systematically analyse the leading corrections to the brane inflation potential (Baumann et al., 2009a, 2010), using insights from the AdS/CFT correspondence (Maldacena, 1998; Gubser et al., 1998; Witten, 1998). The potential acquires a correction, ΔV, of the form

$$\Delta V(\phi, \Psi) = \sum_\Delta c_\Delta \left(\frac{\phi}{\phi_{\mathrm{UV}}}\right)^\Delta f_\Delta(\Psi), \tag{4.16}$$

where Ψ denotes the angular direction, and ϕ_{UV} is the value of the inflaton at the large-radius end of the throat. The scaling dimension Δ takes a discrete spectrum of values [2] (Baumann et al., 2009a, 2010):

$$\Delta = 1, \; \frac{3}{2}, \; 2, \; \frac{5}{2}, \; \dots \tag{4.17}$$

Sufficiently deep inside the throat, $\phi \ll \phi_{\mathrm{UV}}$ and so only a few of the lowest Δ's will be relevant. Furthermore, symmetries can project out some of the Δ modes. Notice that the $\Delta = 1$ mode does not affect η, but only corrects the slope of the potential. We will absorb it into the small-eta contribution of the potential by defining

$$V_0(\phi) \equiv \hat{V}_0(\phi) + a_1 f_1(\Psi)\frac{\phi}{M_{\mathrm{pl}}}. \tag{4.18}$$

In general, the potential implied by (4.16) has a highly nontrivial angular dependence given by the functions $f_\Delta(\Psi)$. We will not solve the angular problem in full generality in this work, but study various approximate solutions that we believe capture the qualitative features. Only in the case when a single Δ mode is dominant has the effective radial potential been computed analytically (Baumann et al., 2009a). In that case, the resulting potential is

$$V(\phi) = V_0(\phi) + M_{\mathrm{pl}}^2 H^2 \left[\left(\frac{\phi}{M_{\mathrm{pl}}}\right)^2 - a_\Delta \left(\frac{\phi}{M_{\mathrm{pl}}}\right)^\Delta\right]. \tag{4.19}$$

---

[2]The scale of the individual corrections (parametrized by the coefficients c_Δ) is hard to compute without more global information about the compact space and its effects on perturbations at ϕ_{UV}, the large-radius end of the throat. The constants c_Δ are therefore treated as free (tunable) parameters.

The minus sign in Eq. (4.19), which is crucial for achieving the correct phenomenology, arises because f_Δ is a harmonic function. Any nontrivial harmonic function must have a negative region; thus, if one angular mode dominates, minimizing over the angular directions will lead to negative f_Δ. This argument is only valid when a single mode of Δ dominates, hence we consider only that case and small perturbations thereof.

We believe this approximation to be valid in the inflationary region (in particular, during the last 60 e-folds of expansion), but for large values of the inflaton field higher angular modes may become relevant and complicate the angular motion. When considering the overshoot problem, however, the evolution in this region will usually be dominated by kinetic energy, and hence largely insensitive to the detailed shape of the potential.

Consider, for example, the case where the inflaton potential is determined by the combined effect of the $\phi^{3/2}$, the ϕ^2, and possible higher ϕ^Δ corrections. Importantly, the two correction terms have different angular dependencies

$$\Delta V(\phi, \Psi) = a_{3/2} \left(\frac{\phi}{M_{\rm pl}} \right)^{3/2} f_{3/2}(\Psi) + a_2 \left(\frac{\phi}{M_{\rm pl}} \right)^2 f_2(\Psi) . \tag{4.20}$$

At some critical value of the field, $\phi_{\rm c}$, (depending on the coefficients $a_{3/2}$ and a_2) both terms will be equally important. For $\phi < \phi_{\rm c}$, the $\phi^{3/2}$ term dominates and the potential has a unique angular minimum. For $\phi > \phi_{\rm c}$, the ϕ^2 term dominates and the location of the angular minimum changes. For a critical region near $\phi_{\rm c}$, both the $\phi^{3/2}$ term and the ϕ^2 term are dynamically important, so the exact radial motion is hard to compute in a simple way. In this Chapter, we will make the simplifying assumption that a single mode ϕ^Δ determines the potential shape throughout the throat region. We know this to be invalid for large values of ϕ, but since the dynamics in this regime will usually be dominated by kinetic energy, we do not expect a distortion of the potential shape beyond the inflationary region to significantly affect our overshoot analysis.

The corrections to the D3-brane potential were computed in Baumann et al. (2009a) in the 'probe approximation', i.e. the most general UV perturbations to the throat supergravity solution are parametrized and the D3-brane is then used as a probe of that perturbed geometry. It is true that in the relativistic limit the brane kinetic energy can lead to additional corrections to the throat geometry. It is quite possible that these corrections again fall into the classification of Baumann et al. (2009a), but the full nonlinear problem has not been studied and is beyond the scope of the present Chapter.

In any case, once the dynamics has become potential-dominated, our branes always slow-roll (see Section 4.5). The potential corrections for a slow-rolling (nonrelativistic) brane derived in Baumann et al. (2009a) are therefore sufficient. Furthermore, in the relativistic limit the dynamics is insensitive to the form of the potential, but the (corrected) shape of the warp factor would be important.

Two qualitatively different scenarios can now arise, depending on whether the fractional $\Delta = \frac{3}{2}$ mode is part of the spectrum or is projected out.

Quadratic models. If the $\Delta = \frac{3}{2}$ mode is projected out of the spectrum, the effective radial potential is

$$V(\phi) = V_0(\phi) + \beta H^2 \phi^2 . \tag{4.21}$$

The phenomenology of these types of potentials was first studied analytically by Kachru et al. (2003) and Firouzjahi and Tye (2005), and numerically by Shandera and Tye (2006).

Inflection point models. If the fractional mode $\Delta = \frac{3}{2}$ is present, it leads to inflection point models (Baumann et al., 2009a; Baumann et al., 2008, 2007; Krause and Pajer, 2008). If $\phi^{3/2}$ was the only correction term, the effective radial potential would be (Baumann et al., 2009a)

$$V(\phi) = V_0(\phi) + M_{\mathrm{pl}}^2 H^2 \left[\left(\frac{\phi}{M_{\mathrm{pl}}} \right)^2 - a_{3/2} \left(\frac{\phi}{M_{\mathrm{pl}}} \right)^{3/2} \right]. \tag{4.22}$$

We also allow a small ϕ^2 term to slightly perturb the potential in the inflationary region

$$V(\phi) = V_0(\phi) + M_{\mathrm{pl}}^2 H^2 \left[\beta \left(\frac{\phi}{M_{\mathrm{pl}}} \right)^2 - a_{3/2} \left(\frac{\phi}{M_{\mathrm{pl}}} \right)^{3/2} \right], \tag{4.23}$$

although this does not lead to qualitative changes in our results. If $\beta \sim O(1)$ then the potential (4.23) may be viewed as a small perturbation of the inflection point model of Eq. (4.22) (Baumann et al., 2009a; Baumann et al., 2008).

To simplify calculations, we now set $M_{\mathrm{pl}} = 1$ and replace $H(\phi)$ by its value at ϕ_{UV}

$$H_{\mathrm{UV}}^2 \equiv H^2(\phi_{\mathrm{UV}}) \approx \frac{V_{\mathrm{UV}}}{3} = \frac{Ds}{3} + \mathcal{O}(Ds\phi^2). \tag{4.24}$$

This leads to an error in $V(\phi)$ of $\mathcal{O}(D\phi^4)$, the same as the overall uncertainty in the potential, which can be removed by a small redefinition of a_Δ.

Thus the quadratic case is

$$V(\phi) = Ds \left(1 - \frac{CD}{\phi^4} + \alpha\phi + \frac{\beta\phi^2}{3} \right), \tag{4.25}$$

and the inflection point case is

$$V(\phi) = Ds \left(1 - \frac{CD}{\phi^4} + \alpha\phi + \frac{\beta}{3} \left[\phi^2 - a_{3/2}\phi^{3/2} \right] \right), \tag{4.26}$$

where for convenience we have defined

$$C \equiv \frac{3}{16\pi^2} \frac{1}{s}. \tag{4.27}$$

Parameters of the potential. To be sure that the argument regarding the minimization of the angular function is valid, and that the sign of the leading coefficient is such that it can cancel the inflaton mass term, one single power of ϕ must be dominant in Eq. (4.16). The linear term (which cannot solve the eta problem) must therefore be smaller than the leading a_Δ term, *i.e.*

$$\alpha\phi < \frac{a_\Delta \phi^\Delta}{2} \tag{4.28}$$

throughout inflation. We do not know the exact location of the tip, so we impose this at the end of inflation, when $\epsilon = 1$, ensuring that it is the case for all larger ϕ. For the inflection point case, a_2 must be small compared to $a_{3/2}$ (although it need not be positive), and so we also impose $0.9 < \beta < 1.1$.

4.2.3 Initial conditions and overshoot

The basic objective of this Chapter is to examine the overshoot problem for models of brane inflation. To quantify the effect of an initially large kinetic energy, we define the conjugate momentum

$$\Pi \equiv \frac{\partial \mathcal{L}}{\partial \dot{\phi}} = \frac{\dot{\phi}}{(1 - f(\phi)\dot{\phi}^2)^{1/2}} = \gamma \dot{\phi} \,, \tag{4.29}$$

where $\mathcal{L}$ is the Lagrangian density corresponding to the DBI action (4.6). We can redefine the Lorentz factor (4.11) in terms of the conjugate momentum (4.29)

$$\gamma \equiv \sqrt{1 + f(\phi)\Pi^2} \,. \tag{4.30}$$

We wish to analyze the effect of the initial field value ϕ_i and a possibly large initial momentum Π_i on classical inflationary trajectories in a flat, homogeneous Universe. We shall apply a criterion similar to that applied to, for example, chaotic inflation in Belinsky et al. (1985); Piran and Williams (1985). The criterion R_i for a given potential is the proportion of the total phase space which leads to successful inflation, and is defined by

$$R_i \equiv \frac{\int F(\phi_i, \Pi_i) d\phi_i d\Pi_i}{\int d\phi_i d\Pi_i} \,, \tag{4.31}$$

where $F(\phi_i, \Pi_i)$ is an indicator function unity where (ϕ_i, Π_i) leads to successful inflation, and zero otherwise. We define a successful trajectory to be one which satisfies all microphysical and observational constraints (see below). Minimal fine-tuning of initial conditions corresponds to $R_i = 1$, and maximal fine-tuning to $R_i = 0$. Here we have used the Hamiltonian measure of Gibbons et al. (1987), canceling a factor of $a^3(\phi_i)$ which is the same for all parameters. Our main results do not depend on the details of the measure, but on the upper bound for Π derived below. Hence our conclusions should be largely unaltered should an alternative measure be used.

This is a well-defined criterion, because there are microphysical constraints on the allowable initial conditions that make the total phase space finite.[3] A lower limit is set upon ϕ_i because the branes must initially be separated by at least a string length, so that inflation does not end immediately. We also assume that $\Pi_i \geq 0$, so that ϕ evolves monotonically. We shall now review briefly two other important constraints on the phase space for brane inflation, arising from microscopic consistency requirements, which bound the initial conditions phase space from above.

Field range constraint. A restriction on the maximal field range of the inflaton for brane inflation models was derived in Baumann and McAllister (2007); heuristically, this bound arises because the inflaton now has a geometrical interpretation, and cannot move a distance that is larger than the size of the compactified space. At the same time the four-dimensional Planck mass (the unit in which the field excursion of the canonical inflaton field is measured) also scales with the volume of the compact space. By considering only the volume of the space inside the warped throat (omitting any motion within the bulk) one obtains an upper bound on the inflaton field range for a given warp factor. For

---

[3]This is to be contrasted with the corresponding effective field theory analysis, where ad-hoc limits on the phase space need to be introduced.

the AdS warp factor considered in this work, the field range bound ultimately has the following simple form (Baumann and McAllister, 2007):

$$\left(\frac{\Delta\phi}{M_{\text{pl}}}\right)^2 < \frac{4}{N}\,. \tag{4.32}$$

For the throat to be warped we require that $N \geq 1$, and the validity of the supergravity description requires that $N \gg 1$. We shall take $N = 10$ to be the lower limit (the exact number makes little difference to the results), and thus obtain a lower bound on the warping and an upper bound on the field range. We see that in the regime of control the field range is always sub-Planckian.

Backreaction constraint. To constrain the initial inflaton momentum we consider the requirement that the D3-brane not contain enough energy to trigger the production of significant amounts of Kaluza-Klein modes, which would back-react on the geometry. This is required for the effective four-dimensional description to be valid, and implies that the Hubble parameter is everywhere less than the local warped string scale (Underwood, 2008; Frey et al., 2006; Becker et al., 2007; Kobayashi et al., 2008). At the tip, where the warped string scale is smallest, we impose the condition

$$H \leq \frac{m_s}{h_{\text{IR}}^{1/4}}\,. \tag{4.33}$$

Using Eq. (4.14), we get

$$H < \sqrt{2}\pi^{3/4}\,D^{1/4}g_s^{1/4}\,. \tag{4.34}$$

Finally, this can be written as a function of the conjugate momentum

$$\frac{\Pi}{\sqrt{2(\gamma+1)}} \quad < \quad \sqrt{6\pi^{3/2}\,D^{1/2}g_s^{1/2} - V(\phi)}\,. \tag{4.35}$$

For concreteness, we shall take $g_s = 0.01$.

4.3 Dynamics and Observables

In the preceding section, we examined the model of D-brane inflation, and the theoretical consistency conditions that attach to it. We shall now discuss the evolution of the inflaton, define observables, and briefly review the relevant observational constraints, completing the definition of successfully inflating trajectories.

4.3.1 Equations of motion

The inflationary dynamics are controlled by the DBI action (4.5). For our time coordinate we shall use the number of e-folds of inflation,

$$dN = Hdt\,. \tag{4.36}$$

Using the canonical momentum of Eq. (4.29), we can write the equations of motion in the following form

$$\frac{d\phi}{dN} = \frac{\Pi}{H\gamma}, \tag{4.37}$$

$$\frac{d\Pi}{dN} = -3\Pi - \frac{V'(\phi)}{H} + \frac{f'(\phi)}{2H\gamma}\left(\frac{\gamma-1}{f}\right)^2, \tag{4.38}$$

where a prime denotes a derivative with respect to ϕ. Our conventions are chosen such that ϕ decreases towards the end of inflation, in agreement with the sign convention used, for example, in Peiris et al. (2007). We solve this system of equations numerically in the following analysis.

4.3.2 Cosmological observables

Power spectra. We shall calculate observables to first order in the Hubble slow-variation parameters

$$\epsilon = \frac{2M_{pl}^2}{\gamma}\left(\frac{H'}{H}\right)^2, \tag{4.39}$$

$$\eta = \frac{2M_{pl}^2}{\gamma}\left(\frac{H''}{H}\right), \tag{4.40}$$

$$\kappa = \frac{2M_{pl}^2}{\gamma}\left(\frac{\gamma'H'}{\gamma H}\right). \tag{4.41}$$

The power spectra of scalar and tensor fluctuations are then given (to first order in the slow-variation parameters) by

$$P_s(k) = \frac{\gamma}{8\pi^2 M_{pl}^2}\frac{H^2}{\epsilon}\bigg|_{k=aH\gamma}, \tag{4.42}$$

$$P_t(k) = \frac{2}{\pi^2}\frac{H^2}{M_{pl}^2}\bigg|_{k=aH}. \tag{4.43}$$

Notice that scalar perturbations freeze out at $k = aH\gamma$, when exiting the sound horizon, while tensor modes freeze out at $k = aH$. We can parametrize small deviations from scale-invariance using the scalar spectral index n_s, and measure the magnitude of gravitational waves using the tensor-scalar ratio r

$$n_s - 1 = \frac{d\ln P_s}{d\ln k} = 2\eta - 4\epsilon - 2\kappa, \tag{4.44}$$

$$r = \frac{P_t}{P_s} \approx \frac{16}{\gamma}\epsilon. \tag{4.45}$$

Because P_t, the tensor power spectrum, is evaluated at freeze-out for the tensor modes $k = aH$, the second equality in (4.45) holds only in the slow-roll regime, where $\gamma \approx 1$. When we calculate n_s and r, we shall always start the inflaton on the DBI attractor

trajectory, as described by Eq. (4.52) in Section 4.5. When we consider the fine-tuning of the initial conditions we shall quantify the effect of starting the inflaton off the attractor. In the following analysis, we assume that these observables are measured at the typical physical scales corresponding to cosmic microwave background (CMB) measurements.

4.3.3   Observable constraints

The most important observational constraints on inflation are provided by the CMB. In order to predict the observables for brane inflation, we must compute the evolution of the inflaton, find the end of inflation, and then work backward to determine the e-folds corresponding to the CMB scales, taking reheating into account.

The end of inflation. Within a string length of the tip of the throat, string modes between the D3-brane and the anti-D3-branes will become tachyonic (Kachru et al., 2003). The resulting instability will annihilate the branes and end inflation. However, near the tip, the warp factor deviates from the form of Eq. (4.8) in a model-dependent way. This makes it difficult to calculate the value of ϕ when the tachyonic transition occurs. Fortunately, it is not necessary; inflation will end when $\epsilon > 1$, so we need simply make sure that this happens before the tip of the throat.

While we cannot calculate the value of ϕ at the tip exactly, we can find an approximation to it by combining Eqs. (4.8) and (4.14). This is a very conservative upper bound, because the true form of the warp factor is much flatter than this (Klebanov and Strassler, 2000). We found that this upper bound is generally slightly larger than ϕ_{end}, at which inflation ends via $\epsilon = 1$, but that the differences between the two quantities do not significantly change the observables. Thus, we can be reasonably confident that the tip of the throat is not reached before the breakdown of slow-roll.

We encountered a small class of models for which γ became very large in the region of the tip, preventing $\epsilon > 1$. With the AdS warp factor, these models are all ruled out by the field range bound. However, since inflation here is likely to continue right up to the tachyonic transition, it is possible that model-dependent corrections to the warp factor near the tip could weaken the field range bound sufficiently to prevent the model from being ruled out. This could lead to distinctive observational signatures, such as primordial black holes, but because γ only becomes large near the end of inflation, is unlikely to lead to large non-Gaussianity on CMB scales. The calculation of such warp factor corrections is beyond the scope of this work. We simply assume the AdS warp factor throughout, and thus consider such models ruled out by the field range bound.

Reheating uncertainty. Reheating after brane inflation is very model-dependent (Kofman and Yi, 2005). Here, we make the simplifying assumption of instantaneous reheating. From the point of view of the potential, more sophisticated reheating models would simply modify the number of effective e-folds needed for inflation to solve the horizon problem, thus shifting the part of the potential to which the CMB scales correspond. This is roughly equivalent to a small change in the value of s in Eq. (4.15). Because we already allow s to vary, this should not change the results.

Horizon scales. If we assume instantaneous reheating, a scale k exits the horizon N

e-folds before the end of inflation;

$$N(k) - \log \left[H(N) \right] = 67.9 - \frac{1}{2} \log \left[H(\phi_{\text{end}}) \right] . \tag{4.46}$$

Here ϕ_{end} is the value of the inflaton at the end of inflation. We can use Eq. (4.46) to deduce the location of the CMB scales in terms of N. This depends slightly on the potential, but in practice it is usually around $N = 58$. In order to solve the horizon problem, inflation must last at least this long.

WMAP normalization. The amplitude of the power spectrum has been accurately determined by experiments such as WMAP (Komatsu et al., 2009) to be, at the 68% confidence level, $P_{\text{s}} = (2.445 \pm 0.096) \times 10^{-9}$ at $k = 0.002$ Mpc^{-1}. We impose the constraint that P_{s} is within 1.5σ of its observed value. This essentially fixes the overall scale of the potential, Ds in Eq. (4.19).

The only other measurement currently strong enough to significantly constrain the model is that on n_{s}, but this is not imposed because we would like to see the full range of observables predicted by the self-consistent model parameter space.

4.3.4 Priors on the potential parameters

In the quadratic case, β lies between zero and one. Since all models are equivalent for a sufficiently small β, we take $\beta > 10^{-5}$ as a lower limit.

In the inflection point case, a fine-tuning is required in the parameter $a_{3/2}$ to achieve inflation, so we simply place broad limits upon this parameter and allow the sufficient inflation constraint to set effective limits.

In order to avoid complications with reheating, we assume that s, which measures the contribution to the total potential energy from the bulk, does not dominate enormously over the potential energy from the brane-antibrane pair. Concretely, we assume $s < 10$. Table 4.1 summarizes our priors for the entire parameter space.

Parameter	Lower limit	Upper limit
$\log \beta$ (quadratic)	-5	0
β (inflection)	0.9	1.1
$a_{3/2}$	0	1
$\log s$	0	1
α	0	Eq. (4.28)
$\log \lambda$	-0.3	Eq. (4.32)
ϕ_{i}	ϕ_{end}	Eq. (4.32)
Π_{i}	0	Eq. (4.35)

Table 4.1 Summary of parameter space. Uniform priors were placed on the quantities in the leftmost column. The lower bound on $\log \lambda$ comes from Eqs. (4.9) and (4.32).
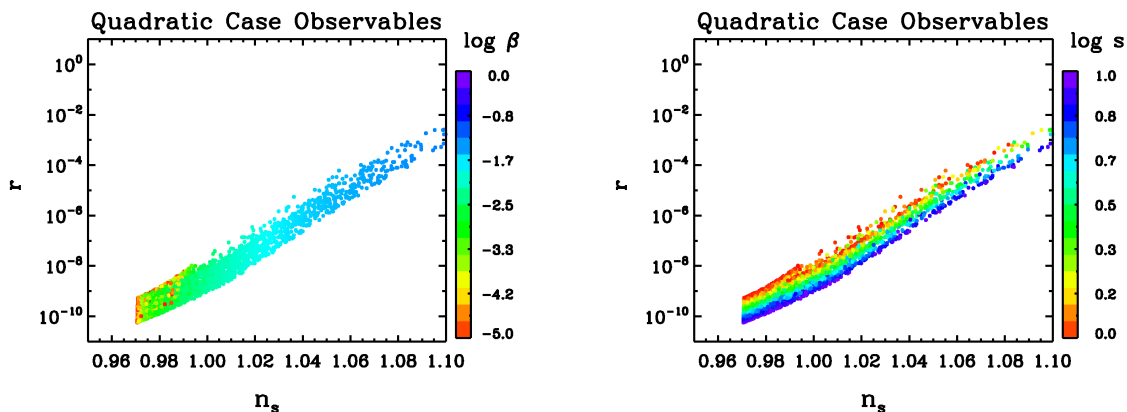
Figure 4.2 The scalar spectral index $n_{\rm s}$ vs the tensor-scalar ratio r for the quadratic case. Left: color scale shows $\log\beta$, the size of the mass correction to the potential. Right: color scale shows $\log s$, the magnitude of the bulk energy in units of the warped string scale at the tip.

4.4 Results

4.4.1 Observables

In this section we consider the cosmological observables, ignoring the overshoot problem for the moment. We shall choose the initial conditions specifically in order to minimize the amount of overshoot; this implies a large $\phi_{\rm i}$, and $\Pi_{\rm i}$ chosen so that the evolution starts on the inflationary attractor. Each dot in Figs. 4.2 and 4.3 corresponds to a set of parameters which generated a successful trajectory (as defined in the previous sections) with these special initial conditions, and the values of $n_{\rm s}$ and r plotted for each dot are those for this trajectory. Bean et al. (2007) performed this analysis for the quadratic case, and we reproduce their results where our analyses overlap.

Figure 4.2 plots the observables $n_{\rm s}$ and r against values of the potential parameters, for the quadratic case. The scatter in r is largely due to variation in the values of the parameter s, while β controls n_s. A significant value of α would boost the tensor-scalar ratio, without affecting $n_{\rm s}$. Since we are not allowing α to dominate the dynamics, it just leads to additional scatter in r. The equivalent graphs for the inflection point case follow a similar pattern, although $n_{\rm s}$ is generically much bluer than in the quadratic case, reflecting the increased difficulty of circumventing the η problem.

In neither case are observable tensor modes generated. This is consistent with the Lyth bound in combination with the field range limit of Baumann and McAllister (2007).

Points with the red tilt favored by WMAP have small β. Red tilt requires $\eta < 0$, and, since we always have $\gamma \approx 1$ (see Section 4.5), this implies $V'' < 0$. This implies

$$2\beta\phi < \frac{20\,CD}{\phi^6}\,. \qquad (4.47)$$

Equation (4.28), the limit on α, is in terms of $a_2 = 1 - \beta$, and so a larger value of α is allowed for such small β models. Thus the scatter due to α is more pronounced for parameters giving rise to a red tilt.

When $\beta \lesssim 10^{-5}$, the quadratic term is dynamically unimportant and the potential is

$$V(\phi) = Ds \left(1 - \frac{CD}{\phi^4} \right) . \tag{4.48}$$

Thus the predictions for all models with very small β are essentially identical, as derived analytically in Appendix C of Kachru et al. (2003).

The Lorentz factor γ is very close to unity for all models, quadratic and inflection point. We found that it was not possible to generate an observable degree of primordial non-Gaussianity with this potential if all constraints are satisfied, in agreement with the results of Bean et al. (2007). A proof is presented in Section 4.5 that, for the potential (4.19) and the AdS warp factor, the inflationary attractor slow-rolls. The proof makes use of the field range bound of Eq. (4.32) and significant non-Gaussianity can be generated if this restriction is ignored (Bean et al., 2007). A successful trajectory will in general be close to the attractor for most of its evolution, and thus close to the slow-roll limit. We found no trajectories, for any initial conditions, which generated observable non-Gaussianity on CMB scales.

Because all models are slow-roll, the nonstandard kinetic term does not have a significant effect on the dynamics, and so the value of λ is uncorrelated with n_s and r.

The brane-antibrane annihilation which ends inflation will generate cosmic strings, at a dimensionless amplitude given by Firouzjahi and Tye (2005)

$$G\mu = \sqrt{\frac{1}{32\pi g_s}} \left[\frac{1}{f_{IR}} \right]^{1/2} . \tag{4.49}$$

Here $g_s < 1$ is the string coupling, and f_{IR} is the (rescaled) warp factor at the tip of the throat. Taking a fiducial value of $g_s = 0.01$ and using Eq. (4.14), we can calculate $G\mu$. If we assume a red spectral tilt, we have $5 \times 10^{-9} < G\mu < 2 \times 10^{-8}$ for the quadratic case, and $7 \times 10^{-11} < G\mu < 2 \times 10^{-9}$ for the inflection point case. A significantly larger value of s would lower these values. Models with a redder spectral tilt need a smaller warp factor at the tip to match the CMB amplitude, and thus lead to a lower value of $G\mu$. Where they overlap, these results are in agreement with Firouzjahi and Tye (2005); Bean et al. (2007).

The current best observational bounds on cosmic strings come from searching for their signatures in the CMB (Lo and Wright, 2005), and from nondetections of the gravitational radiation that would result from the decay of string networks. Millisecond pulsar timing surveys provide the best constraints of the latter type, but are dependent on assumptions about the overall level of background gravitational waves. Constraints vary between $G\mu < 10^{-6}$ and $G\mu < 10^{-7}$ (Battye et al., 2006).

This is insufficient to dramatically limit the model, but future weak lensing surveys or CMB experiments could improve constraints to $G\mu < 10^{-9}$ (Mack et al., 2007; Planck Collaboration, 2005; Fraisse et al., 2008), while the LISA satellite could improve upon that by several orders of magnitude (DePies and Hogan, 2007).

4.5 Non-Relativistic DBI Attractor

For the DBI kinetic term and the potential used in this Chapter, significantly relativistic trajectories with $\gamma > 1.1$ are known to be ruled out by the field range bound (Baumann

and McAllister, 2007) (see Bean et al. (2007) for a numerical study). Here we prove analytically that the inflationary attractor is always close to the slow-roll limit, with $\gamma \approx 1$.

To describe the DBI inflationary attractor found by Underwood (2008), we define a new variable,

$$\chi \equiv -\dot{\phi}\sqrt{f}\,, \tag{4.50}$$

such that

$$\dot{\chi} = \frac{\sqrt{f}}{\gamma^3}\left[V' - 3H\gamma\chi - \frac{f'}{f^2}(\gamma - 1)\right]. \tag{4.51}$$

This has an attractor solution when $d\chi/dt \approx 0$:

$$
\begin{aligned}
\chi_{\text{Attr}} &= \frac{(A+B)\sqrt{1 + A^2 + 2AB} - B}{(1 + (A+B)^2)} \\
&< A + B\left(\frac{\sqrt{1 + A^2 + 2AB} - 1}{1 + (A+B)^2}\right),
\end{aligned} \tag{4.52}
$$

where we have defined

$$A \equiv \frac{V'\sqrt{f}}{3H}\,, \tag{4.53}$$

$$B \equiv \frac{f'}{3Hf^{3/2}}\,. \tag{4.54}$$

If $A^2 + 2AB \ll 1$, then we have $\chi_{\text{Attr}} \sim A$. The definition of γ gives

$$\gamma = (1 - \chi^2)^{-1/2}\,. \tag{4.55}$$

To show that the attractor solution is nonrelativistic, it suffices to show that $A \ll 1$, irrespective of the size of B. To see this, we shall consider only the part of the attractor solution which depends on B, assuming that A is small, and bracket the behavior in the two limiting cases. If B is small enough that AB is also small, then

$$\chi - A < \frac{AB^2}{1 + (A+B)^2} < A\,. \tag{4.56}$$

To derive the final inequality, note that this term is largest when $B \gg 1$.

If, on the other hand, $B > 1/A$, so that AB is also large, we have

$$\chi - A < \frac{\sqrt{2A}B^{3/2}}{1 + (A+B)^2} \sim \sqrt{\frac{2A}{B}} < \sqrt{2}A\,. \tag{4.57}$$

The second equality is because $AB \gg 1$, which implies that $A > B^{-1}$. In the worst case, therefore, the speed on the attractor is bounded by a small multiple of A.

We must now show that A is small. Using (4.53) we have the following upper bound

$$A < \frac{\sqrt{f}V'}{3\sqrt{V}}\,. \tag{4.58}$$

This can clearly be large if f is allowed to grow without limit, but the field range bound prevents this. Eq. (4.32) can be recast in the form (Peiris et al., 2007) (in units where $M_{\rm pl} \equiv 1$)

$$\frac{{\rm Vol}(X_5)}{\pi} \int_{\phi_{\rm end}}^{\phi_{\rm UV}} d\phi \ \phi^5 f(\phi) \ < \ 1 \,, \qquad (4.59)$$

which implies, for the AdS warp factor

$$\frac{\pi^2}{2} \lambda \phi_{\rm UV}^2 < 1 \,, \quad \Rightarrow \quad \sqrt{f} < \frac{\sqrt{2}}{\pi \phi_{\rm UV} \phi^2} \,. \qquad (4.60)$$

Substituting this and Eq. (4.19), the form of the potential, into Eq. (4.58) and dropping terms of order unity gives

$$A \ < \ \frac{\sqrt{Ds}\left(2\phi + \frac{CD}{\phi^5}\right)}{\phi_{\rm UV}\phi^2 \pi \sqrt{1 + \frac{1}{3}\phi^2 - a_\Delta \phi^\Delta - \frac{CD}{\phi^4}}} \,. \qquad (4.61)$$

This is much less than 1 unless $\frac{D}{\phi^4} \sim 1$, by which time inflation has already ended. Thus the attractor is slow-roll, and as long as the inflaton moves on the attractor, we will have $\gamma \approx 1$, in agreement with the numerical results.

4.5.1   Implications for the overshoot problem

We shall now relax the special choice of initial conditions in the last section, and consider the full $(\phi_{\rm i}, \Pi_{\rm i})$ plane compatible with the initial conditions bounds of §4.2.3, checking each successful set of parameters from the previous section for trajectories suffering from overshoot. We find that, for every set of parameters allowed by the constraints, $R_{\rm i} < 5 \times 10^{-3}$. That is, more than 99.5% of initial conditions lead to trajectories suffering from overshoot. This was because, in contrast to the specific case considered by Underwood (2008), the bound on the initial kinetic energy was insufficiently restrictive to solve the overshoot problem. The other effect identified in Underwood (2008), the increased extent of the inflationary attractor, and consequent increase in e-folds, was also found to be ineffective. Both effects required a much stronger warping than was allowed by the field range bound and the observational constraints.

We have determined that the entire "successful" parameter space suffers from overshoot, but would like to know which parameter combinations suffer more than others. Unfortunately, the fact that overshoot trajectories dominate makes this difficult, so we introduce an additional, ad-hoc, kinetic bound of $\Pi_{\rm i} \leq 10^{-7} M_{\rm pl}^2$ purely as a presentation aid for the following results. We call the analogous rescaled initial conditions fine-tuning criterion $\hat{R}_{\rm i}$. The bound was chosen so that the least fine-tuned quadratic model had $\hat{R}_{\rm i}$ of about unity, and then applied to both the quadratic and inflection point cases. We calculated $\hat{R}_{\rm i}$ to an accuracy of about 1%, using a grid-refinement method.

Figure 4.3 shows the results for $\hat{R}_{\rm i}$. For observationally relevant values of $n_{\rm s}$, the quadratic case suffers less from overshoot than the inflection point case. Points with $n_{\rm s} < 1.0$ always have $\hat{R}_{\rm i} \approx 0$. These potentials have small β or $a_{3/2}$, and so are very flat in the region of interest. Because of this, they must dissipate more kinetic energy before
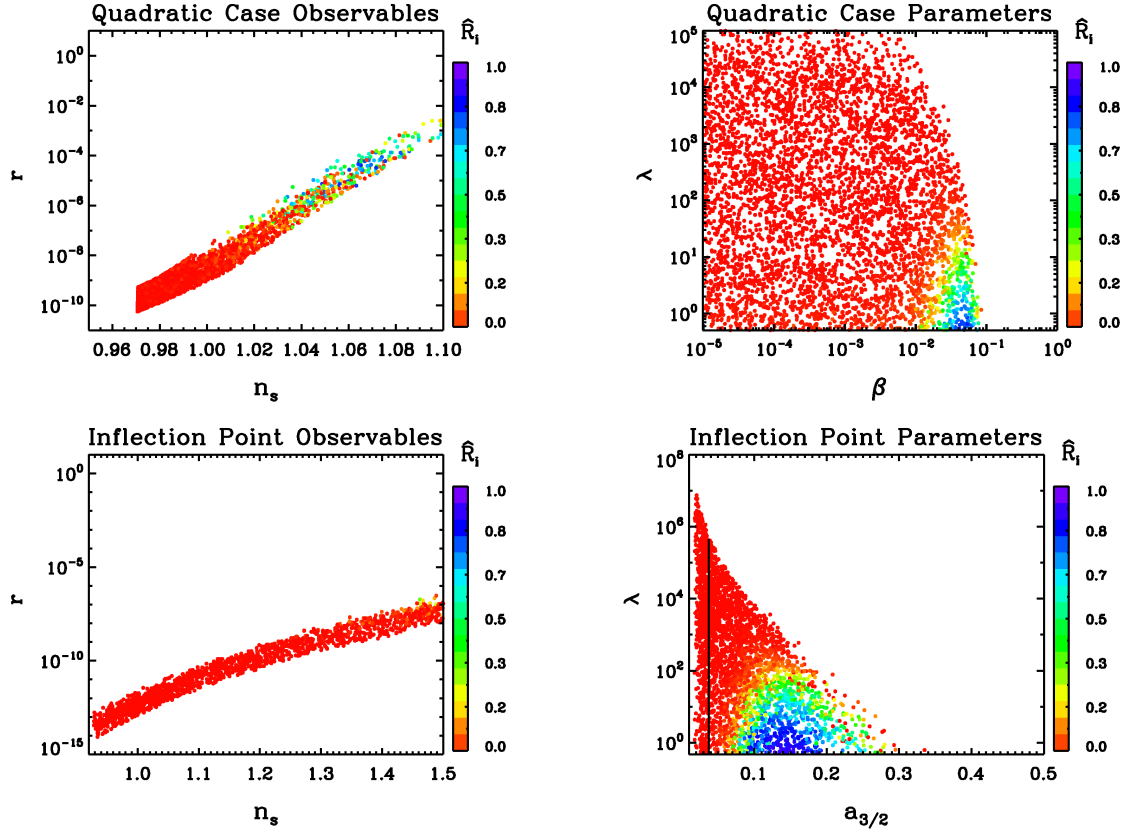
Figure 4.3 The color scale shows the rescaled initial conditions criterion $\hat{R}_i$. Top: results for the quadratic case. Bottom: results for the inflection point case. Left: scalar spectral index n_s plotted against the tensor-scalar ratio r. Right: strength of the warping scale λ plotted against the magnitude of the mass term (β for the quadratic case, $a_{3/2}$ for the inflection point case).

becoming potential–dominated, making them more susceptible to overshoot. The flatter the potential, the less field variation is required to generate sufficient inflation. Thus the flatter models, which also have $n_{\rm s} < 1$, can have a larger warping scale than those with $n_{\rm s} > 1$, without violating Eq. (4.32).

For blue tilted models, the most important parameter controlling $\hat{R}_{\rm i}$ is λ, the warping scale. The warping scale is largely uncorrelated with the observables, which leads to the slightly chaotic appearance of the top-left plot of Fig. 4.3. It is apparent from the right-hand plots of Fig. 4.3, which show $\hat{R}_{\rm i}$ as a function of the potential parameters, that a high value of λ can make it harder to achieve inflation, even though this means a smaller initial conditions phase space. The effect of increasing λ is to tighten the field range bound, Eq. (4.32). A tighter field range bound allows the inflaton less space to dissipate excess kinetic energy, making overshoot more severe.

4.5.2 Successful inflection point inflation

The generic behavior of the potential in the inflection point case involves either super-Planckian field variations (disallowed by the field range bound), when the $a_{3/2}$ correction is too small, or else a metastable minimum of the potential when it is too large. Successful inflation can only occur in a carefully balanced intermediate regime between these two effects, when the minimum has become an inflection point. This requires a fine-tuning of the model parameters and the initial conditions, which we shall now quantify. We shall assume $\alpha = 0$ for simplicity. In order to have an inflection point, we need

$$5 \times 10^{-3} < a_{3/2} < 1.5 \,. \tag{4.62}$$

If $a_{3/2} > 1.5$, a minimum always occurs for $\phi > M_{\rm pl}$. This could be transformed into an inflection point potential if D was sufficiently large, but inflation would then occur in the vicinity of the former minimum, where
$\phi > M_{\rm pl}$, which is always outside the throat. If $a_{3/2} < 10^{-3}$, the potential is insufficiently flat and the field variation is super-Planckian, violating the field range bound.

Once we have chosen an $a_{3/2}$, we need to pick D carefully. It is helpful to reparametrize CD in terms of $a_{3/2}$

$$CD \equiv w \left(\frac{3 a_{3/2}}{4} \right)^{12} \,. \tag{4.63}$$

To get inflation we need $5.3 \times 10^{-3} < w < 6 \times 10^{-3}$. Once we combine this limit with the observed amplitude of the power spectrum, we need:

$$0.01 < a_{3/2} < 0.4 \,. \tag{4.64}$$

The bottom-right plot of Fig. 4.3 shows this fine-tuning graphically, and the further fine-tuning that results when we also require that $n_{\rm s} < 1.1$.

A nonzero α can weaken this tuning, and increase the generic value of r, but it does not significantly change the observables, unlike in the quadratic case.

4.6 Conclusions

Recently, Underwood (2008) proposed an intriguing solution to the overshoot problem for brane inflation.[4] Theoretical limits on allowed initial conditions shrunk the volume of phase space, removing most of the trajectories leading to overshoot, while the existence of the DBI attractor extended to regions where the slow-roll attractor did not, thus allowing more inflation to take place. However, Underwood (2008) studied the dynamics only for a very specific choice of parameters in the inflationary potential and the warp factor, so it remained unclear if these two effects were sufficient to solve the overshoot problem in general.

In this Chapter we performed a comprehensive scan of the parameter space allowed by the latest advances in brane inflation model-building (Baumann et al., 2009a, 2006; Baumann et al., 2008). For each set of parameters we first worked out the observables for a particular set of initial conditions, chosen so that they were least likely to suffer from the overshoot problem. If this trajectory satisfied all microphysical and observational constraints, we quantified the degree of overshoot by calculating the proportion of the homogeneous phase space of initial conditions leading to successful inflation. We found that in all cases this was less than 5×10^{-3}; that is, less than 0.5% of the phase space volume of initial conditions led to successful trajectories. Therefore, brane inflation does not significantly alleviate the overshoot problem.

The qualitative differences in our conclusions compared to Underwood (2008) are easy to understand: Most importantly, the parameters chosen by Underwood (2008) violate the field range bound (4.32), and therefore do not satisfy all microphysical constraints, unless one allows $\mathrm{Vol}(X_5)$ (the dimensionless volume of the base of the warped throat geometry) to be very small. In fact, we find that $\mathrm{Vol}(X_5) \lesssim 10^{-4}$ is required to prevent overshoot solutions. This is to be compared with the generic value of $\mathrm{Vol}(X_5) = \pi^3$ that we have used in our analysis. For parameters which do satisfy the constraints, the extension of the attractor solution does not occur; as proved in Section 4.5, the DBI attractor is close to the slow-roll limit throughout once we impose the field range bound.

We find that known theoretical limits on the initial conditions phase space do not shrink it sufficiently to rule out all the overshoot trajectories. The difference from Underwood (2008) is due to Eq. (3.40) of that work missing a factor of T_3^{-1}, as can be seen on dimensional grounds. If the maximum initial momentum is Π_{max} then, with the revised bound given in Eq. (4.33), we have

$$\Pi_{\mathrm{max}} \propto D^{1/4} \,, \tag{4.65}$$

where D is a measure of the warped string scale at the tip of the throat (which determines the warped 3-brane tension at the tip). On the other hand, the amplitude of the primordial power spectrum is proportional to Ds, as shown by Eq. (4.43). Here, s parametrizes the ratio of energy density localized in the bulk relative to the energy density of the brane-antibrane system in the throat. If the inflationary energy is dominated by the brane-antibrane tension [corresponding to $s \sim 1$ in Eq. (4.19)], the warped string scale at the tip constrains both the range of the initial momentum and the amplitude of the

---

[4]Other creative proposals for solving the overshoot problem of brane inflation have appeared in Itzhaki and Kovetz (2007); Itzhaki (2008); Spaliński (2009).

primordial power spectrum. Because the latter is tightly constrained by observations, we cannot solve the overshoot problem.

If we allow the inflationary energy to be dominated by bulk energy (corresponding to $s \gg 1$ in Eq. (4.19)), we break this correlation. The kinetic energy bound is still given by Eq. (4.65), and allows us to solve the overshoot problem provided $s \gtrsim 10^{20}$ (since the normalization of the power spectrum constrains the combination Ds, this is equivalent to a very small value of D, the warped string scale at the tip). However, the bulk energy remains after the end of inflation. In this case, it may be challenging to explain how the Universe attains a state with small cosmological constant after inflation, *i.e.* reheating would have to be addressed carefully.

The addition of an initially large, negative spatial curvature can mitigate overshoot (Freivogel et al., 2006) by increasing the Hubble friction. In Freivogel et al. (2006), the inflaton gains kinetic energy by rolling down a steep potential after starting at rest through a tunnelling event. In our formulation, we make no specific assumption about the mechanism by which initial conditions are set, only that they have to obey the microphysical bounds (4.32) and (4.33), and the magnitude of the kinetic term allowed by the backreaction bound is typically significantly larger than in the specific mechanism of Freivogel et al. (2006). For large kinetic energy, the friction due to curvature is less effective. We found that the initial contribution of curvature to the Hubble parameter had to be larger than the contribution of the scalar field to have a significant effect. The largest initial kinetic energy for which the curvature term was effective was far below the microphysical bound; about $\Pi \sim 10^{-6}$. The equivalent figure without curvature is $\Pi \sim 10^{-7}$. This is not a significant increase in terms of resolving overshoot, and so we have neglected the effects of spatial curvature.

In order to determine which regions of parameter space suffered most from overshoot, we excluded the parts of the initial conditions space which always led to overshoot, and computed the proportion of the remaining space that led to successful inflation. We found parameter sets with a blue tilt $n_s > 1$ and a small warping scale suffered least from overshoot. In the inflection point case, avoiding overshoot to the same degree as the quadratic case required very blue spectra, inconsistent with observational limits.

Our analysis of the overshoot problem relied on currently available microscopic bounds on the inflaton field value $\phi_{\max}$ (from the geometric field range bound) and momentum $\Pi_{\max}$ (from backreaction considerations). We found that these bounds were too weak to allow a solution of the overshoot problem. However, it is conceivable that a future analysis could reveal stronger constraints on $\phi_{\max}$ and $\Pi_{\max}$. We therefore ask what bounds would be sufficient. Reducing $\phi_{\max}$ while keeping $\Pi_{\max}$ fixed would make the overshoot problem more severe, since this provides less field space for dissipating excess kinetic energy. The most promising avenue for solving the overshoot problem is therefore finding a stronger microscopic bound on the initial inflaton momentum. Quantitatively, our analysis showed that the new bound on $\Pi_{\max}$ would have to be at least 4 orders of magnitude stronger than the currently applied limit. Solving the overshoot problem for parameters giving a red tilt would require an improvement of 5 orders of magnitude. These conclusions are not sensitive to small perturbations of the potential shape away from the ~ 60 observationally-relevant e-folds of inflation.

As a by-product, our analysis recovered known observational predictions of quadratic brane inflation (Bean et al., 2007; Firouzjahi and Tye, 2005), and computed new pre-

dictions for the inflection point case. The model parameter space consistent with microphysical and observational restrictions predicts unobservably small levels of both tensor modes and primordial non-Gaussianity, but a potentially observable cosmic string tension, in both the quadratic and inflection point cases. We have also shown analytically that the DBI attractor solution will always slow-roll, and thus trajectories near it will generate nonobservable levels of non-Gaussianity.

"It is a capital mistake to theorise before one has data. Insensibly one begins to twist facts to suit theories, instead of theories to suit facts. "

  – A Scandal in Bohemia, by Sir Arthur Conan-Doyle

5

Minimally Parametric Power Spectrum Reconstruction from the Lyman-α Forest

Abstract

Current results from the Lyman-α forest assume that the primordial power spectrum of density perturbations follows a simple power-law form, with running. We present the first analysis of Lyman-α data to study the effect of relaxing this strong assumption on primordial and astrophysical constraints. We perform a large suite of numerical simulations, using them to calibrate a minimally parametric framework for describing the power spectrum. Combined with cross-validation, a statistical technique which prevents over-fitting of the data, this framework allows us to reconstruct the power spectrum shape without strong prior assumptions. We find no evidence for deviation from scale-invariance; our analysis also shows that current Lyman-α data do not have sufficient statistical power to robustly probe the shape of the power spectrum at these scales. In contrast, the ongoing Baryon Oscillation Sky Survey will be able to do so with high precision. Furthermore, this near-future data will be able to break degeneracies between the power spectrum shape and astrophysical parameters.

5.1 Introduction

The primordial power spectrum of density fluctuations underpins much of modern cosmology. On large scales, it has been measured with high precision by cosmic microwave background (CMB) experiments (e.g. Komatsu et al. 2011 and references within). In order to improve our knowledge of its scale-dependence, we turn to smaller scales, and astrophysical measurements probing later epochs in the evolution of the Universe. In this Chapter, we shall examine constraints from the data set which has probed the smallest scales to date: the Lyman-α forest.

The Lyman-α forest consists of a series of features in quasar spectra due to scattering of quasar photons with neutral hydrogen. Since hydrogen makes up most of the baryonic density of the Universe, the Lyman-α forest traces the intergalactic medium (IGM), and thus the baryonic power spectrum, on scales from a few up to tens of Mpc. This makes it the only currently available probe of fluctuations at small scales in a regime when the corresponding density fluctuations were still only mildly non-linear, thereby simplifying cosmological inferences. A number of authors have examined the cosmological constraints from the Lyman-α forest in the past (Croft et al. 1998; Theuns et al. 1998; McDonald et al. 2000; Hui et al. 2001; Viel et al. 2002; Gnedin and Hamilton 2002; McDonald et al. 2005a; Viel and Haehnelt 2006; Lidz et al. 2006), while Seljak et al. (2005, 2006) examined constraints combined with other data sets. For a review of the physics of the IGM and its potential for cosmology, see Meiksin (2009).

Previous analyses assumed that the primordial power spectrum on Lyman-α scales is described by a nearly scale-invariant power law – a strong prior – and proceeded with parameter estimation under this assumption. In contrast, in this work we attempt to constrain the shape and amplitude of the primordial power spectrum at these scales using minimal prior assumptions about its scale-dependence.

In view of the observational effort dedicated to the Lyman-α forest, and its promise as a probe of the primordial power spectrum, in this work we shall explore the possibilities of going beyond parameter fitting. To give us insight into the underlying model for the power spectrum shape, which parameter estimation by itself cannot do, our present application to Lyman-α data should therefore ideally assume full shape freedom throughout the analysis. As a nearly scale-invariant primordial power spectrum is a generic prediction of the simplest models of inflation, a minimally parametric reconstruction can be a powerful test of inflationary models. Lyman-α constrains the smallest cosmological scales; thus, it provides the longest lever-arm when combined with the statistical power and robustness of CMB data, yielding the best opportunity presently available to understand the overall shape of the power spectrum.

The main Lyman-α observable, the flux power spectrum, does not have a simple algebraic relationship to the matter power spectrum. By $z \sim 3$, the absorbing structures are weakly nonlinear, and are also affected by baryonic physics. Hence, to establish the relationship between the primordial power spectrum and the flux power spectrum, we must resort to hydrodynamical simulations. The initial conditions used in our simulations allow for considerable freedom in the shape of the primordial power spectrum, and this allows us to recreate the Lyman-α forest resulting from generic power spectrum shapes. Using an ensemble of simulations which sample the parameter space required to describe the flux power spectrum, we construct a likelihood function which can be used to perform minimally parametric reconstruction of the primordial power spectrum, while simultaneously constraining parameters describing IGM physics.

A statistical technique called cross-validation (CV) is used to robustly reconstruct the primordial power spectrum and Markov Chain Monte Carlo (MCMC) techniques are used to obtain the final constraints. The statistical approach parallels Sealfon et al. (2005) and Verde and Peiris (2008), who applied the same method to data from the CMB and galaxy surveys. Peiris and Verde (2010) added the current Lyman-α forest data to the joint analysis with larger scale data, via the derived constraints on the small-scale matter power spectrum from McDonald et al. (2005a). However, these latter constraints

were derived assuming a tight prior on the shape of the primordial power spectrum at Lyman-α scales – an assumption which we drop in this work. In our analysis we consider both the flux power spectrum determined by McDonald et al. (2006) from low-resolution quasar spectra obtained during the SDSS, and simulated data for the upcoming Baryon Oscillation Sky Survey (BOSS: Schlegel et al. 2009).

This Chapter is organized as follows. In Section 5.2 we review the framework for power spectrum reconstruction and describe the details of the simulations and parameter estimation setup. Section 5.3 describes the data, and results are presented in Section 5.4. We conclude in Section 5.5. Technical details of our calculations are relegated to Appendices 2.4 and 5.2.6.

5.2 Methods

In this section, we describe the statistical technique used in this Chapter, and how we built the likelihood function for minimally parametric reconstruction from Lyman-α data. Section 5.2.1 describes the framework for power spectrum reconstruction in general terms, while Section 5.2.2 gives further details of our specific implementation of this framework. Sections 5.2.3 - 5.2.5 detail numerical methods used to extract a flux power spectrum from a given primordial power spectrum. Finally, Section 5.2.7 describes the parameter estimation implementation.

5.2.1 Power spectrum reconstruction

Previous analyses of the Lyman-α forest (McDonald et al. 2005a; Viel et al. 2004) have assumed that the primordial power spectrum is a nearly scale-invariant power law, of the form

$$P(k) = A_{\mathrm{s}} \left(\frac{k}{k_0} \right)^{n_{\mathrm{s}}-1+\alpha_{\mathrm{s}} \ln k} , \tag{5.1}$$

and then constrained A_{s}, n_{s} and α_{s}. In this work we will follow the same spirit as Sealfon et al. (2005); Verde and Peiris (2008), going beyond parameter estimation in an attempt to deduce what the Lyman-α forest data can tell us about the shape of the power spectrum under minimal prior assumptions. A major challenge involved in all such reconstructions is to avoid over-fitting the data; it is of little use to produce a complex function that fits the data set extremely well if we are simply fitting statistical noise. Equally, an overly prescriptive function which is a poor fit to the data should be rejected. To achieve this balance, we add an extra term, $\mathcal{L}_P$, to the likelihood function which penalizes superfluous fluctuations. Schematically, the likelihood function is:

$$\log \mathcal{L} = \log \mathcal{L}(\mathrm{Data}|P(k), \mathrm{Astrophysics}) + \lambda \log \mathcal{L}_{\mathrm{P}} , \tag{5.2}$$

where the form of $\mathcal{L}_{\mathrm{P}}$ will be discussed shortly. We have assumed a model for the IGM thermodynamics, discussed in Section 5.2.4, and marginalised over a number of parameters of this model, including the temperature and temperature-density relation. This model constitutes a physically motivated prior constraint, incorporating information

about the expected physics of the IGM. Part of the purpose of this study is to examine whether this model is sufficient; should we detect features in the primordial power spectrum, we would look at whether a physically motivated extension of this model could account for them.

Equation 5.2 now contains an extra free parameter which measures the magnitude of the smoothing required; the penalty weight λ. As $\lambda \to \infty$ the likelihood will implement linear regression. For particularly clean data, carrying no evidence for any feature in the $P(k)$, λ should be large. Data carrying strong evidence for $P(k)$ features would be best analysed with a small value of λ. We need a method of determining, from the data, the optimal penalty weight. Our chosen technique is called cross-validation (CV) (Green and Silverman 1994), which quantifies the idea that a correct reconstruction of the underlying information should accurately predict new, independent data.

The variant of CV used in this Chapter splits the data into three sets. The function is reconstructed using two of these sets (training sets). The likelihood (excluding the penalty term) of this reconstruction, given only the data in the remaining set (validation set), is calculated. This step is called validation; because the data in each set are assumed to be independent, we now have a measure of the predictivity of the reconstruction. Validation is repeated using each set in turn and the total CV score is the sum of all three validation likelihoods. The optimal penalty is the one which maximizes the CV score.

More generally, CV splits the data into k independent sets, with $2 < k \leq n$, where n is the number of data points. $k - 1$ sets are used for training, and the remaining set for validation. Larger k allows for a bigger training set, and thus better estimation of the function to be validated against, but for most practical problems large k is computationally intractable. We have chosen $k = 3$ as a compromise. We verified that using $k = 2$, following Verde and Peiris (2008), made a negligible difference to our results despite the smaller training set size.

CV assumes that each set is uncorrelated; a mild violation of this assumption will lead to an underestimation of errors, but not a systematic bias in the derived parameters Carmack et al. (2009). Our data include a full covariance matrix, and so we are able to verify that correlations between the sets are weak.

The minimally parametric framework applied in this Chapter follows that of Sealfon et al. (2005), Verde and Peiris (2008) and Peiris and Verde (2010). It uses cubic splines to reconstruct a function $f(x)$ from measurements at a series of points, x_i, called the knots. The function value between each pair of knots is interpolated using a piecewise cubic polynomial. The spline is fully specified by the knots, continuity of the first and second derivatives, and boundary conditions on the second derivatives at the exterior knots (the knots at either end of the spline). The splines have vanishing second derivative at the exterior knots. If the power spectrum is given by smoothed splines, the form of the likelihood function given above is

$$\log \mathcal{L} = \log \mathcal{L}(\text{Data}|P(k)) + \lambda \int_k d\ln k (P''(k))^2, \qquad (5.3)$$

$$\text{where } P''(k) = \frac{d^2 P}{d(\ln k)^2}.$$

Knot	Position	$P(k)$ $(/10^{-9})$	
	(Mpc^{-1})	Minimum	Maximum
A	0.475	0.83	3.25
B	0.75	0.60	3.23
C	1.19	0.60	3.67
D	1.89	0.53	4.16

Table 5.1 Positions of the knots. The maximum and minimum values of $P(k)$ are the extremal values covered by our simulations. Fixed knots are not shown, but are discussed in the text.

5.2.2 Knot placement

The number and placement of the knots is chosen initially and kept fixed throughout the analysis. Once there are sufficient knots to allow a good fit to the data, adding more will not alter the shape of the reconstructed function significantly. In choosing the number of knots, we seek to find a balance between allowing sufficient freedom in the power spectrum, and having few enough parameters that the data are still able to provide meaningful constraints on two sets out of three when subdividing the data into the training and validation sets, as described above. Available computing resources limit us in any case to considering only a few knots. We fit the primordial power spectrum with a four-knot spline for the Lyman-α forest k-range. The flux power spectrum is available in twelve k-bins, so there are three bins per knot, which should allow sufficient freedom. By comparison, Peiris and Verde (2010) used seven knots to cover the k-range spanned by CMB, galaxy surveys and Lyman-α data, with a single knot for the Lyman-α forest.

The SDSS flux power spectrum covers the range of scales, in velocity units, of $k_v = 1.41 \times 10^{-3} - 0.018\,\text{s/km}$. Dividing by a factor of $H(z)/(1+z)$ converts to comoving distance coordinates, so the constraints on the matter power spectrum are on scales of roughly $k = 0.4 - 3\ \text{Mpc}^{-1}$. In this range of scales we place four knots (A–D, from large to small scales) evenly in log space. Numerical details of the knots are shown in Table 5.1. The maximum and minimum values of $P(k)$ given there for each knot are simply the extremal values covered by our simulations. Simulation coverage of $P(k)$ has been expanded where necessary to fully cover the range allowed by the data.

We must specify the primordial power spectrum on scales well outside the range probed by data, even though they have no effect on the Lyman-α forest. This is for two reasons. The first is that when running a simulation we must have a well-defined way to perturb the initial particle grid for all scales included in the simulation. In order to ensure that the scales on which we have data are properly resolved, we also need to simulate larger and smaller scales , and these require a defined power spectrum. The second reason is that our interpolation scheme works best when the perturbations induced by altering one of the knots are reasonably local. Adding extra end knots helps to prevent large secondary boundary effects, which would make interpolation far more difficult.

For numerical stability reasons, we would like the amplitude of fluctuations on these scales to be reasonably constant, but do not wish to make strong assumptions about the amplitude of the power spectrum there. Therefore we add two "follower" knots at each end of the spline. The amplitude is fixed to follow the nearest parameter knot, assuming
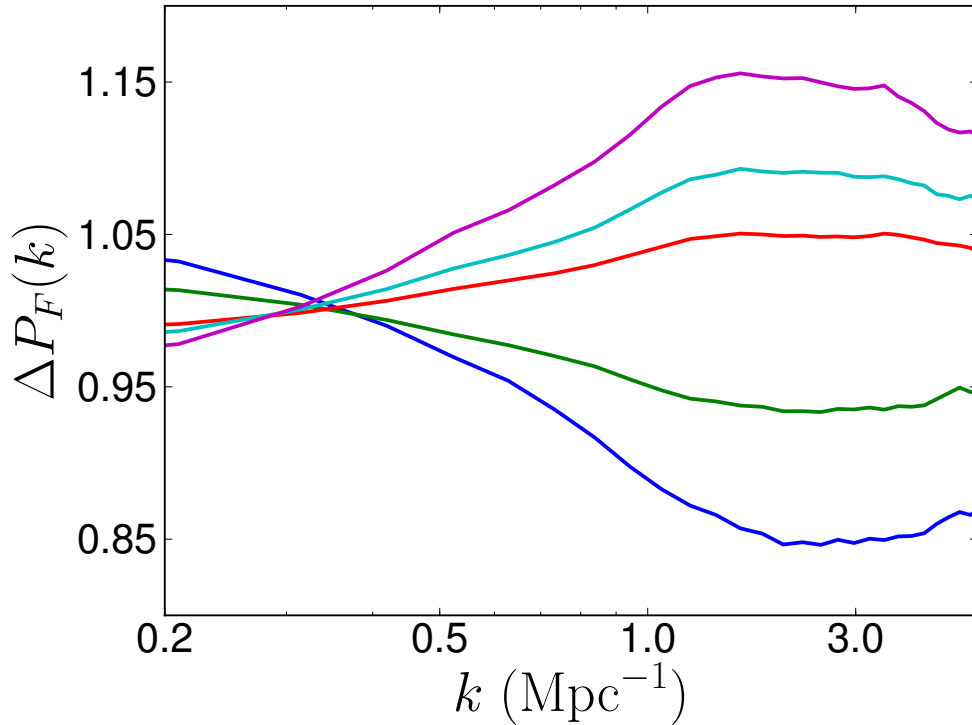
Figure 5.1 Effect on the flux power spectrum of varying the D knot at $z = 3$. On a scale where the best-fitting amplitude is 0.9, the amplitudes of the D knot are, from the lowest line upwards, 0.5, 0.7, 1.1, 1.3 and 1.7. Non-linear growth tends to erase dependence on the initial conditions, so the effect is smaller at lower redshifts.

that between follower and followed, the shape is a power law with $n_s = 0.97$. [1] The two follower knots are at scales of $k = (0.15, 4) \, \mathrm{Mpc}^{-1}$.

We also add a few knots, even more distant from the scales probed by the data, with completely fixed amplitudes consistent with the WMAP best-fitting power spectrum. The amplitude of the primordial power spectrum on these scales does not significantly affect results; we have added knots here so that the initial density field is well-defined on a larger range of scales than probed by the simulation. This allows us to avoid any boundary effects associated with the ends of the spline. These fixed knots are at $k = (0.07, 25, 40) \, \mathrm{Mpc}^{-1}$, with amplitudes of $(2.43, 2.03, 2.01) \times 10^{-9}$. Fig. 5.1 shows the effect of altering the amplitude of the D knot on the flux power spectrum.

5.2.3 Simulations

In this study, full hydrodynamical simulations were run using the parallel TreePM code GADGET-2 (Springel 2005). GADGET computes long-range gravitational forces using a

---

[1]Hence, if the amplitude of the power spectrum at the D-knot is P^D, the power spectrum at the follower knot has the amplitude $P^D(k^D/k^{D+1})^{0.03}$, where k^D is the position of the D knot and k^{D+1} the position of the follower.

particle grid, while the short-range physics are calculated using a particle tree. Hydrodynamical effects are included by having two separate particle types; dark matter, affected only by gravity, and baryons, modelled using smoothed particle hydrodynamics (SPH), where particles are supposed to approximate density elements in the matter fluid. The rest of this section gives technical details of our simulations and the included astrophysics, and may be skipped by the reader interested only in the cosmological implications.

GADGET has been modified to compute the ionisation state of the gas using radiative cooling and ionisation physics as originally described by Katz et al. (1996), and used in Viel and Haehnelt (2006). Star formation is included via a simplified prescription which greatly increases the speed of the simulations, where all baryonic particles with over-density $\rho/\rho_0 > 10^3$ and temperature $T < 10^5 K$ are immediately made collisionless. Viel et al. (2004) compared simulations using this prescription with identical simulations using a multi-phase model, and found negligible difference in the Lyman-α statistics. Additionally, all feedback options have been disabled, and galactic winds neglected; Bolton et al. (2008) found that winds have a small effect on the Lyman-α forest. The gravitational softening length was set to 1/30 of the mean linear inter-particle spacing.

The gas is assumed to be ionised by an externally specified, spatially homogeneous UV background, based on the galaxy and quasar emission model of Haardt and Madau (2001). We follow previous analyses in assuming that the gas temperature is initially in equilibrium with the CMB, that the gas is in ionisation equilibrium, optically thin, and that we can neglect metals and evolution of elemental abundances. Lyman-α absorption arises largely from near mean-density hydrogen, which should undergo little chemical evolution, so using a simplified star formation criterion and neglecting metals is physically well-motivated. Assuming that the gas is optically thin and in ionisation equilibrium will break down during reionisation, but at the redshifts we are interested in, we can model the effect of non-instantaneous reionisation by increasing the photo-heating rate, as described in Viel and Haehnelt (2006).

The fiducial simulation for this Chapter has a box size of 60 Mpc h^{-1} and 2×400^3 gas and dark matter particles, [which we will write as (60, 400) in future], and runs from $z = 199$ to $z = 2$. Snapshots are output at regular intervals between redshift 4.2 and 2.0. Initial conditions were generated using N-GenICs, modified to specify the primordial power spectrum by a spline, and use separate transfer functions for baryons and dark matter, as calculated using CAMB (Lewis et al. 2000).

For knots B and C, we used the above fiducial parameters for box size and particle resolution. For the D knot, we slightly compromised on box size in favour of particle resolution, and used simulations of (48, 400), since we found that the D knot had a negligible affect on the largest scales. To fully capture the behaviour of the A knot, we used larger simulations with (120, 400). We have used different sized simulations to ensure that for each knot, the characteristic scales representing it have very good numerical convergence; this issue is addressed in full in Section 5.2.6. Our ability to do this is one technical advantage of our approach compared with previous studies; if we were to alter the amplitude of the whole power spectrum, we would need to achieve convergence over all the relevant scales at once. In our approach, each simulation only needs strict convergence over the narrow range of scales probed by a single knot.

5.2.4 IGM thermodynamics

Constraints on the thermal history of the IGM are given in terms of the parameters of a polytropic temperature-density relation

$$T = T_0 \left(\frac{\rho}{\rho_0}\right)^{\gamma-1} , \tag{5.4}$$

where a given SPH particle has temperature T and density ρ. T_0 and ρ_0 are the average quantities for the whole simulation snapshot. To determine T_0 and γ from a simulation box, a least-squares fit is performed from low-density particles satisfying

$$-1.0 < \log\left(\frac{\rho}{\rho_0}\right) < 0 . \tag{5.5}$$

Regions that are less dense than the lower limit above are ignored because they are poorly resolved in SPH simulations (Bolton and Becker 2009). The simplified star formation criterion means that many overdensities have been turned into stars, and their baryonic evolution not followed; hence they are also neglected. Both γ and T_0 are assumed to follow a power law broken at $z = 3$ by HeII reionisation (Schaye et al. 2000), so that they are given by:

$$\gamma = \begin{cases} \gamma^A \left[(1+z)/4\right]^{d\gamma^S} & \text{if } z < 3, \\ \gamma^A \left[(1+z)/4\right]^{d\gamma^R} & \text{if } z > 3. \end{cases} \tag{5.6}$$

$$T_0 = \begin{cases} T_0^A \left[(1+z)/4\right]^{dT_0^S} & \text{if } z < 3, \\ T_0^A \left[(1+z)/4\right]^{dT_0^R} & \text{if } z > 3. \end{cases} \tag{5.7}$$

When performing parameter estimation, we marginalize over γ^A, T_0^A and $d\gamma^{S,R}$, $dT_0^{S,R}$. The different thermal histories were constructed by modifying the fiducial simulation's photo-heating rate as described in Section 2.2 of Bolton et al. (2008).

The effective optical depth is described by a power law, with parameters:

$$\tau_{\text{eff}} = \tau_{\text{eff}}^A \left[(1+z)/4\right]^{\tau_{\text{eff}}^S} . \tag{5.8}$$

Previous studies (McDonald et al. 2005a; Viel and Haehnelt 2006) used the same transfer function for both dark matter and baryon particles; we have used different transfer functions for baryon and dark matter species. At our starting redshifts, the transfer functions for the baryons are about 10% lower than for the dark matter on these scales, because baryon fluctuations have not grown as fast during tight coupling. Once they have decoupled from the photons, the baryons fall into the potential wells of the dark matter, and by $z = 1$, the linear transfer functions are almost identical. At redshifts $2-3$, however, the effect is small but noticeable, and accounts for a 2% scale-independent drop in the power spectrum. This is too small to affect current data, but could be potentially important for analysing BOSS data.

Our model for the thermodynamics of the IGM assumes that the gas is optically thin. By $z = 3$, the neutral fraction of hydrogen is extremely low, making this a good

approximation for describing Lyman-α photons. However, the reionisation of HeII to HeIII at this time means that the gas is optically thick to helium reionising photons. HeII reionisation alters the effective heating rate and changes the temperature-density relation by adding energy to the gas; an effect we account for by marginalising over γ and T_0. However, HeII heating processes are likely to be inhomogeneous, potentially producing significant fluctuations in the IGM temperature. The impact of such fluctuations on the flux power spectrum was examined by McQuinn et al. (2011), using simulations which incorporated radiative transfer effects. They found a negligible effect on the 1D flux power spectrum used here. This was ascribed to two effects; first, a higher temperature region of the IGM has a smaller neutral fraction, leading to exponentially less absorption, and thus a negligible contribution to the flux power spectrum. Furthermore, the large, slightly overdense regions giving rise to the Lyman-α forest are less affected by temperature fluctuations than underdense areas.

5.2.5 The flux power spectrum

In the case of Lyman-α, the observable is not a direct measurement of the clustering properties of tracer objects, as in galaxy clustering, but the statistics of absorption along a number of quasar sightlines. Therefore we define the flux, $\mathcal{F}$, as

$$\mathcal{F} = \exp(-\tau), \tag{5.9}$$

where τ is the optical depth. We define the flux power spectrum as

$$P_F(k) = |\tilde{\delta_F}(k)|^2,$$
$$\delta_F = \frac{\mathcal{F}}{\bar{\mathcal{F}}} - 1 \,. \tag{5.10}$$

Here $\bar{\mathcal{F}}$ is the mean flux. The tilde denotes a Fourier transformed quantity, where our Fourier conventions, used throughout, are:

$$\tilde{f}(k) = \int f(x)e^{ikx}\mathrm{dx} \,. \tag{5.11}$$

To aid the eventual understanding of our results, we digress slightly here to review the physical effects of the various thermal parameters on the flux power spectrum. The mean flux, essentially a measure of the average density of neutral hydrogen, has a large impact on the amplitude of the flux power spectrum. Cosmological information from the Lyman-α forest is obtained through examining the power spectrum shape and its redshift dependence. The effect of a higher temperature, as preferred by the flux power spectrum, is to suppress power predominantly on small scales, as a higher temperature wipes out small-scale structure in the baryons. The exponent of the temperature-density relation, γ, controls the temperature difference between voids and overdensities. A higher γ makes voids cooler and overdensities hotter. At high redshifts, where much of the Lyman-α absorption comes from voids, the effect of an increased γ is to decrease the temperature of the Lyman-α emitting regions, so there is relatively more small-scale structure. At low redshifts, however, most of the Lyman-α absorption comes from near mean density material, and so an increased γ increases the temperature, decreasing the

amount of small-scale structure. For further details of the physical effects of the various parameters, see Section 4.2.1 and Fig. 3 of Viel and Haehnelt (2006), as well as Fig. 11-13 of McDonald et al. (2005a).

Current constraints on P_F are given by McDonald et al. (2006), determined from ~ 3000 SDSS quasar spectra at $z = 2 - 4$.

Each simulation snapshot was processed to generate an averaged flux power spectrum as follows. First, 8000 randomly placed simulated quasar sightlines were drawn through the simulation box. For a 60 Mpc h^{-1} box, this constitutes an average spacing between sightlines of 670 h^{-1}kpc, corresponding to scales of roughly $k = 10$ Mpc h^{-1}, far smaller than the scales probed by the Lyman-α forest. We verified that doubling the number of sightlines to 16000 made a negligible difference to the resulting power spectra.

When calculating absorption, particle peculiar velocities were included, which increases the (non-rescaled) magnitude of the power spectrum by approximately 10%.

To generate the flux power spectrum, the absorption due to each SPH particle near the sightline is calculated, giving us a number of simulated quasar spectra, which are smoothed with a simple boxcar average. Each spectrum is rescaled by a constant so that the mean flux across all spectra and absorption bins matches that observed by Kim et al. (2007). This rescaling hides our ignorance of the amplitude of the photo-ionizing UV background. The mean over all the rescaled spectra is then used as the extracted flux power spectrum for the box. For further details of how we computed the absorption, see Section 2.4.

We follow previous work in not attempting to model continuum fitting errors. The Si III contamination found by McDonald et al. (2006) is modelled by assuming a linear bias correction of the form $P'_F = [(1 + a^2) + 2a \cos(vk)] P_F$, with $a = f_{\text{SiIII}}/(1 - \bar{\mathcal{F}})$, $f_{\text{SiIII}} = 0.011$, and $v = 2271 \,\text{km/s}$.

Finally, since high-density, damped Lyman-α systems (DLAs) are not modelled by our simulations, we add a correction to the flux power spectrum to account for them, of the form calculated by McDonald et al. (2005b). The amplitude of this correction is a free parameter, and will be discussed further in Section 5.2.7.

We checked the convergence of our simulations with respect to box size and particle resolution. Here we give only a brief summary of the results; further details may be found in Section 5.2.6. For the highest redshift bins at $z = 4.2$, 4.0 and 3.8, increasing the particle resolution had a large effect on the flux power spectrum. Achieving numerical convergence for the Lyman-α forest at high redshift is challenging, because most of the signal for the Lyman-α forest is coming from poorly resolved underdense regions. In addition, current data at high redshifts are much more noisy than at low redshifts, and future surveys will not probe these redshifts at all. Accordingly, we follow Viel and Haehnelt (2006) and do not use the three highest redshift bins in our analysis.

At lower redshifts, and except in the smallest and largest k-bins, the change with increased particle resolution was small. On the smallest scales, however, there was a change of around 5% in each bin. This increase is systematic, and so we correct for it as described in Section 5.2.6. The larger box increased power on the largest scales by around 5%, due to sample variance in the simulation box. The methodology we used to correct for this effect is again detailed in Section 5.2.6.

The above figures were the dominant errors in our modelling of the flux power spectrum. Uncorrected modelling errors are therefore $\lesssim 2\%$ of the flux power spectrum in

each bin, far below the current measurement error of $\sim 12\%$ in each bin of the flux power spectrum, and on the order of the expected statistical errors for the BOSS survey, which are $\sim 1.5\%$. A significant decrease in modelling errors would require the use of simulations with improved particle resolution, which are beyond the computational resources available to us.

5.2.6 Convergence Checks

In this appendix, we detail the checks we have performed to ensure that our simulations are properly converged with respect to box size and particle number. We have usually been comparing the relative change in the flux power spectrum when changing a parameter, making strict convergence not essential.

To check box-size convergence, we compared the flux $P_F(k)$ for our fiducial simulation with a large box size simulation ("L") which had size (75, 500), and otherwise identical parameters to the fiducial simulation ("F"). This isolates the effect of box size by having identical particle resolution to simulation F. To test convergence with respect to particle number, we used a high resolution simulation ("H"), with (60, 500). In order to isolate the change due to numerical effects, we did not rescale the mean optical depth for the plots in this section.

The left hand plot of Fig. 5.2 shows the change in the flux power spectrum with increased box size; simulation L divided by simulation F. The flux power spectrum is converged with respect to box-size; however, there is a systematic increase on large scales. This is due to sample variance: the specific realisation of cosmic structure we are using is biased slightly low on the largest scales probed by the box. The larger box recovers the input power spectrum much better on these scales, because it contains far more modes, and hence shows an increase in power. Because we use the same realisation of structure for all our simulations, this effect will be constant, and is easily corrected for by altering the best-fitting power spectrum. Once this is done, convergence of the flux power spectrum is very good.

The right hand plot of Fig. 5.2 shows the change in the flux power spectrum with increased particle resolution. The effect is small, except on small scales or at high redshift. Achieving numerical convergence for the Lyman-α forest at high redshift is challenging, because most of the signal for the Lyman-α forest is coming from poorly resolved underdense regions. In addition, current data at high redshifts is much more noisy than at low redshifts, due to a paucity of quasar spectra. Accordingly, we follow Viel and Haehnelt (2006) and do not use the three highest redshift bins at $z = 4.2, 4.0$ and 3.8 in our analysis.

Our initial base for the central power spectrum was the output of simulation H. We corrected for box size and resolution following the method proposed by McDonald (2003). To correct for sample variance on the largest scales, we ran two additional simulations with box size $120\,\mathrm{Mpc}\,h^{-1}$, and 400^3 and 200^3 particles. The smaller simulation has identical particle resolution to our fiducial simulations, and should thus be directly comparable to them. We then corrected our best-guess power spectrum by the ratio between the two,

$$C_S(k) = P_F(N_P = 400)/P_F(N_P = 200)\,. \tag{5.12}$$

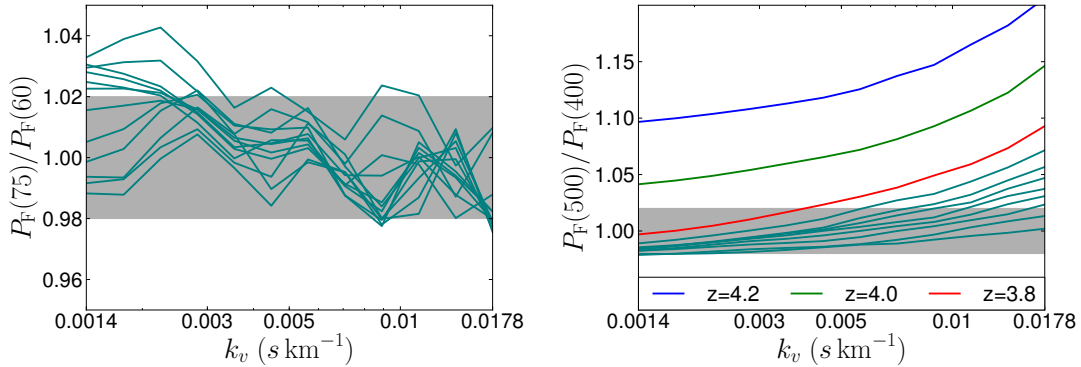The smallest scale bin was excluded from this correction as it was clearly being affected

Figure 5.2 (Left) The change in the flux power spectrum due to increasing the box size at fixed particle resolution. Each green line shows the effect on a different redshift bin, from 4.2 to 2.0. The effect is generally around 2% (grey box), with a systematic increase on large scales, for which we correct (see text). (Right) The effect on the flux power spectrum due to increasing the particle number from 2×400^3 to 2×500^3. Each line shows the effect on a different redshift bin, from 4.2 to 2.0. The grey box show a variation of 4%. Redshift bins with greater variation are labelled.

by poor resolution convergence.

Because simulation H is nearly converged, we had to be careful when correcting for resolution; if the error in the correction is larger than the correction itself, accuracy is definitely not increased. Therefore, we ran two simulations, with box size of $24\,\mathrm{Mpc}\,h^{-1}$; $T1$ and $T2$. $T1$ has the same particle resolution as simulation H, and thus 200^3 particles. $T2$ has 400^3, giving it much increased particle resolution. These simulations do not resolve the largest scales probed by the Lyman-α forest at all, but Fig. 5.2 shows that these scales are not affected by poor resolution convergence. Simulation H is corrected by

$$C_R(k) = P_F^{T2}/P_F^{T1} \,. \tag{5.13}$$

To avoid our correction itself being biased by a small box size, we ignore those bins on scales greater than a quarter of the box-size of $24\mathrm{Mpc}\,h^{-1}$. We also ignore any correction for redshift bins where the correction for the smallest scale bin is less than the uncertainty in the simulations, which we take to be 1%, on the grounds that these are already fully converged. We are left with a slight increase in power on the smallest scales.

5.2.7 Parameter estimation

So far we have given a formula for the primordial power spectrum, and described how we use it to extract a flux power spectrum to compare with observational data. In this section, we shall describe how we actually performed that comparison. First we describe a scheme for robustly interpolating the parameter space to obtain flux power spectra corresponding to parameter combinations which we have not simulated, following Viel and Haehnelt (2006). Secondly, we describe the parameters of the Monte Carlo Markov chains we used for parameter estimation. For more details of MCMC, see, for example Lewis and Bridle (2002).

Parameter interpolation

Directly calculating a flux power spectrum from a given set of primordial fluctuations requires a hydrodynamical simulation. This makes it impractical to directly calculate P_F for every possible set of input parameters. Instead, simulations are run for a representative sample and other results are obtained from these via interpolation. We assume that the flux power spectrum varies smoothly around the best-fitting model, parametrize this variation with a quadratic polynomial for each data point, and then check that this accurately predicts new points. If we have some simulation with a parameter vector which differs from a 'best-guess' simulation by δp_i, the corresponding change in the flux power spectrum, δP_F, is given by

$$\delta P_F = \sum_j \alpha_j \delta p_j + \beta_j \delta p_j^2 \,. \tag{5.14}$$

The coefficients of this polynomial are constrained by performing a least-squares fit to flux power spectra generated by numerical simulations. We experimented with including cross-terms (of the form $p_i p_j$), but found that this did not significantly improve the accuracy of the interpolation.

To estimate the interpolation coefficients, we used seven simulations for each of our four power spectrum parameters, one of which was used to test the accuracy of the interpolation. To check for correlation between parameters, we simulated varying two neighbouring knots at once. As the greatest effect of each knot on the flux power spectrum is over a localized range of scales, our interpolation errors should be maximal here. We needed only four simulations per thermal history parameter, and checked we could accurately predict δP_F for a very different thermal history. As a final interpolation verification, we performed a simulation where all six parameters were changed simultaneously. Fig. 5.3 shows the interpolation errors for one of our tests, which are around 1% of the total change for each bin. This is smaller than the expected statistical errors for BOSS, and was replicated by our other test simulations.

MCMC methodology

To perform parameter estimation, we use a version of the publicly available CosmoMC Lewis and Bridle (2002) code, with a modified likelihood function as described in Section 5.2.1.

We marginalize over four parameters for the four knots, with priors as specified in Table 5.1, and over eight parameters of the thermal history, as described in Section 5.2.3.

We follow the advice of McDonald et al. (2006), and add a number of nuisance parameters to the SDSS data, all with Gaussian priors. To parametrize uncertainty in the resolution of the spectra, we add a parameter α^2 with prior 0 ± 49, and multiply the flux power spectrum by $\exp\left(-k^2\alpha^2\right)$. The effect of an increased α^2 is therefore to damp power on the very smallest scales. Each redshift bin has one parameter, f_i, to describe uncertainty in the subtraction of background noise, with a prior of 0 ± 0.05. To marginalize over the uncertainty in the effect of DLAs, we add A_{damp}, with a prior of 1 ± 0.3. The effect of this correction is to increase the slope of the flux power spectrum.

We also marginalize over residual uncertainties in the Hubble parameter, h and Ω_M, using flat priors of $0.2 < \Omega_M < 0.4$, and $0.5 < h < 0.9$. For the rest of our background
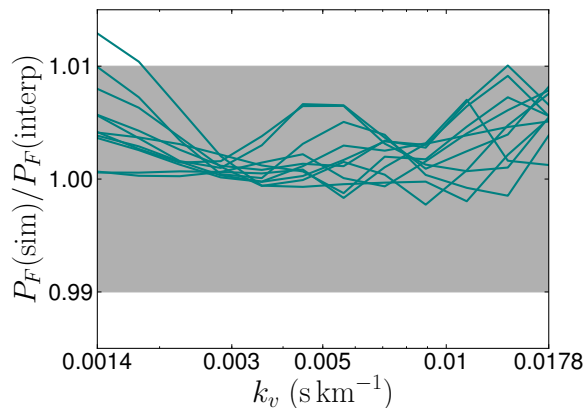
Figure 5.3 The difference between the flux power spectrum as obtained from interpolation, and directly by simulation. Here only the C and D knots have been changed from their initial values. Each line represents simulation output at a different redshift bin, between $z = 2.0$ and $z = 4.2$. The grey band shows 1% error bars.

cosmology, we assume parameters in agreement with those preferred by WMAP 7, including negligible gravitational waves and spatial curvature. The priors on h and Ω_M make a negligible difference to our results, because both these parameters only weakly affect the Lyman-α forest. We assume $T_0 < 50000$ K and $0 < \gamma < 5/3$ on physical grounds; the temperature-density relation of the IGM cannot be steeper than the perfect gas law, and very high temperatures would contradict independent measurements of the IGM temperature by Schaye et al. (2000).

Cross-validation methodology

Cross-validation (CV) requires the splitting of the data set into n independent sets. For best results, these sets should be as uncorrelated as possible. We choose to use alternating bins in k for each set. For data with n k bins, the first set would consist of bins $1, 4, 7 \ldots$, the second bins $2, 5, 6 \ldots$ and the third similarly.

To calculate the CV score, we estimate the best-fit from the two training sets, using a MCMC. The CV score for the remaining, validation, set is the likelihood of this best-fit. The total CV score for a given penalty is the sum of the CV scores for each set.

5.3 Datasets

5.3.1 Current data from SDSS

The SDSS data used in this study consist of a best-fitting flux power spectrum in 12 k-bins and 11 redshift bins, together with a covariance matrix and a set of vectors describing the foreground noise subtraction. It was analysed by McDonald et al. (2006), and comes from 3000 quasar spectra. Of these, ~ 2000 are at redshift $2.2 - 3$, and ~ 1000 above that. We use the 8 redshift bins at $z < 3.8$ only.

We have chosen not to include any additional small-scale information based on high-resolution quasar spectra. In principle, this can help break degeneracies and should be included in future analyses. Currently, however, systematic error from such data sets is hard to quantify, and the optimal method for extracting the thermal state of the IGM is not yet clear. Our focus in this work has been robustness, and so we have limited ourselves to a single data set, whose properties have been extensively studied and are relatively well-understood.

5.3.2 Simulated data from BOSS

In this section we will describe our simulated data for forecasting constraints from BOSS, an ongoing-future survey which will acquire 1.6×10^5 quasar spectra (Schlegel et al. 2009), between $z = 2.2 - 3.0$. We need to simulate both a covariance matrix and a flux power spectrum.

We have assumed that the noise per spectrum of the BOSS data will be approximately the same as they were for SDSS. This is a simple assumption, but broadly justified because both surveys use similar instruments (Schlegel et al. 2009). Truly accurate modelling of the covariance matrix is impossible until the release of the final data, however, we expect our modelling of the BOSS covariance matrix to be completely adequate for a forecast. Our simulated BOSS covariance matrix is simply the SDSS covariance matrix scaled to account for the increase in statistical power resulting from the much greater number of quasar sightlines. There are roughly 2000 quasar sightlines in the SDSS sample below $z = 3$, so the scale factor is $2000/160000 = 1/80$.

To generate the flux power spectrum, we used cosmological parameters consistent with the best-fitting results from WMAP 7, and thermal parameters consistent with theoretical expectations: $\gamma \sim 1.45$ and $T_0 = 2.3 \times 10^3 \left[(1+z)/4\right]^{0.2}$ K. The effective optical depth was $\tau = 0.36 \left[(1+z)/4\right]^{3.65}$. The power spectrum amplitude was selected to match a spectrum with $\sigma_8 = 0.8$ and $n_s = 0.96$.

We then added uncorrelated Gaussian noise with a variance given by the diagonal elements of the simulated BOSS covariance matrix. As BOSS will only take data at $z \leq 3$, we dispense with the thermal parameters for higher redshifts. The foreground noise properties of the BOSS data are expected to be similar to those of the SDSS data; we therefore leave the priors on the parameters measuring uncertainty in the noise subtraction and the parameter measuring resolution uncertainty, α^2, unchanged.

BOSS is also expected to determine the transverse flux power spectrum. Simulating the larger scales needed to properly model the effect of this is beyond the scope of this Chapter, and we refer the interested reader to Slosar et al. (2009); White et al. (2010).

5.4 Results

5.4.1 Current constraints

Fig. 5.4 shows the CV-reconstruction of the primordial power spectrum from SDSS data. We have shown plots for two penalties: one high, one low. This was because we have been unable to determine an optimal penalty from current data; the CV score shows no significant variation, even when the penalty is having negligible impact on the likelihood.
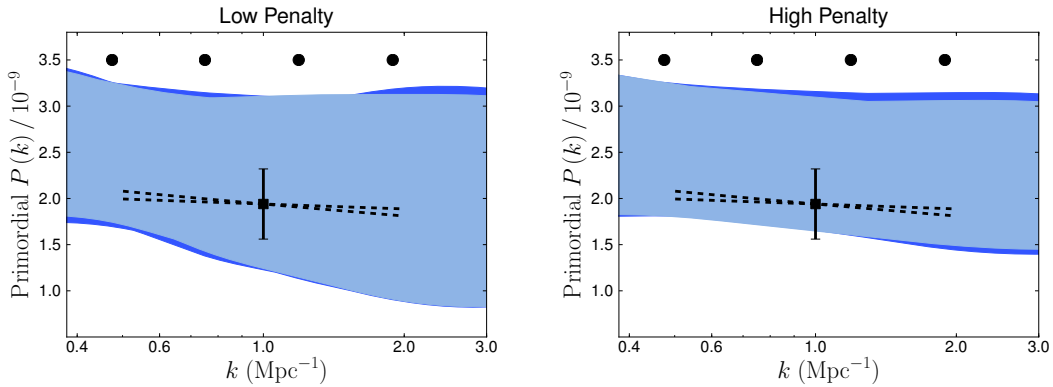
Figure 5.4 Constraints on the primordial power spectrum from SDSS data from CV, for low (left) and high (right) penalties. Black circles show the positions of the knots, with arbitrary normalisation. The light blue regions show the top 68% of likelihoods for SDSS data, while the dark blue regions show the 95% likelihood range. We discuss the similarity between these two regions in the text. The black error bar shows the results of previous analyses (Viel et al. 2009) assuming a power-law power spectrum at $k = 1\,\mathrm{Mpc}^{-1}$. The dashed lines show limits on the slope from that work.

We interpret this to mean that the shape constraints on the primordial power spectrum from current Lyman-α data are very weak.

We have emulated confidence limits by plotting the envelope of samples which have a likelihood in the top 68% and the top 95%. At the 95% level, the power spectrum is allowed to oscillate more within the allowed envelope, but the size of the overall constraint on the amplitude does not greatly change. In this work, the low penalty plot is being affected by the priors on the A-knot; our constraints have become so weak that at the largest scales probed they are affected by the range probed by our simulation suite (see Fig. 5.5 for these limits). It is also possible that the shape of the primordial power spectrum on much larger scales, assumed for numerical reasons, is having a small effect on the constraints by anchoring the amplitude of the power spectrum. Neither effect would be important if it were not for the extreme weakness of the constraints from the data. However, the similarity of the 68 and 95% contours is a generic feature of using the cross-validation technique for reconstruction, as found in Verde and Peiris (2008). It occurs because we have marginalised over the power spectrum shape in a non-trival way, meaning that the shape of the likelihood function is not particularly Gaussian in the power spectrum parameters.

Previous analyses assumed a power-law prior for the shape of the primordial power spectrum, and constrained this slope and the overall normalisation from the same data used above. While such parameter estimation leads to tight constraints from the data (assuming the underlying shape prior is correct), relaxing this tight prior leads to the loss of ability to constrain the scale-dependent shape of the power spectrum. The current data can still be used as part of a minimally parametric primordial power spectrum if one exploits the extended range in scales that can be probed in combination with other data sets (Peiris and Verde 2010).

The black error bar in Fig. 5.4 shows a comparison with Viel et al. (2009). Our
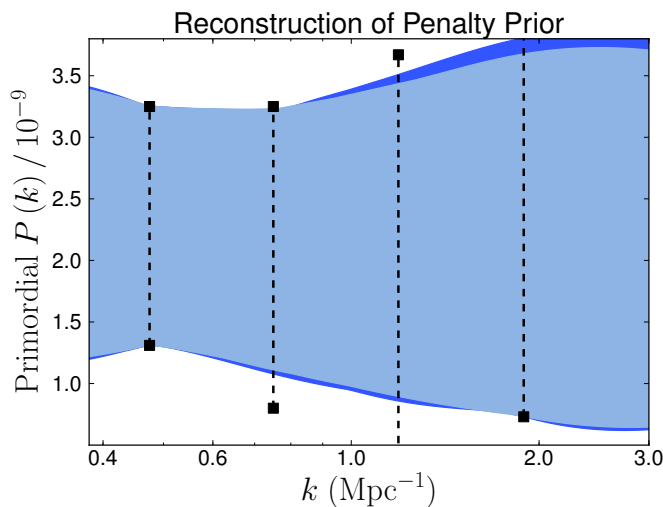
Figure 5.5 Constraints on the primordial power spectrum from the penalty term alone, using the value in the "low penalty" plot of Fig. 5.4 above. Dashed lines show the power specrum range sampled by the simulations.

method gives results for the amplitude of the primordial power spectrum at Lyman-α scales which are completely consistent with that work, but somewhat weaker. This is to be expected; we are removing a tight prior on the shape of the power spectrum. For a very high penalty, i.e. the limit at which the implicit prior in our analysis approaches a power-law spectrum, we can reproduce the error bars of Viel et al. (2009). We are also in agreement with the results of an earlier analysis of the Lyman-α forest, McDonald et al. (2005a), which constrained $\sigma_8 = 0.85 \pm 0.13$.

The corresponding constraints on n_s from our reconstruction are extremely weak, especially for the low penalty: $n_s \sim 0.2 - 1.2$. The constraints on n_s in Viel et al. (2009), in addition to the power-law prior, were greatly assisted by the fact that the pivot scale k_0 in Eq. 5.1 was chosen to be $k_0 = 0.002 \, \text{Mpc}^{-1}$; a small change in the slope of the power spectrum at k_0 leads to a large change in power spectrum amplitude by $k = 1 \, \text{Mpc}^{-1}$. Here, we are trying to constrain a scale-dependent $n_s(k) = 1 + \frac{d \ln P}{d \ln k}$ using only the interval of scales sampled by Lyman-α forest. We find that, while the current Lyman-α data are able to constrain the amplitude of the power spectrum at these scales, they are not powerful enough on their own to significantly constrain the shape of the spectrum in a robust manner. At no penalty do we see any evidence in the current data against a scale-invariant power spectrum.

We can explicitly demonstrate that the current Lyman-α data add little information to a weak prior on the shape of the power spectrum in the following way. Fig. 5.5 shows a minimally parametric reconstruction assuming the penalty designated "low". These constraints were generated *without using any data whatsoever*, and are similar to those obtained with the Lyman-α forest data. This figure shows clearly that our SDSS constraints are affected by the prior even for the low penalty. Since the penalty is proportional to $P''(k)$, it cannot determine the power spectrum amplitude. Instead, the allowed power spectrum amplitude is simply the minimal range probed by our simulations.

The CV part of our method involves reconstructing the optimal penalty, and thus the

strength of the shape prior justified by the data. CV is essentially a method to reconstruct the most favoured prior correlation between knots; since the prior is reconstructed from the data, prior-driven constraints would not necessarily be a problem. However, here we are finding that no particular prior is favoured over any other. Thus, the width of the envelopes in Fig. 5.4 are actually arbitrary and should not be used to draw conclusions about the amplitude of primordial fluctuations at Lyman-α scales.

We performed a number of checks to determine the cause of our failure to find an optimal penalty. Changing our methodology for splitting the data into CV bins did not affect the results. A flux power spectrum simulated in the same way as our BOSS data, and using the same parameters, but with error bars of the same magnitude as the current data showed no preference for a particular penalty, despite, as we shall see, there being a well-defined optimal penalty for BOSS simulations. Fixing the thermal history parameters γ and T_0 to fiducial values was also sufficient to allow us to reconstruct a penalty. Therefore statistical error and systematic uncertainty in the thermal history are the significant factors preventing us from robustly reconstructing a minimally parametric power spectrum shape from current data.

Constraints on the thermal history parameters are as follows. For the low penalty we found $0.8 < \gamma < 1.7$ at 1σ (recall that this upper limit is imposed as a physical prior), while for the high penalty $0.2 < \gamma < 1.7$. The corresponding constraint from Viel et al. (2009) is $\gamma = 0.63 \pm 0.5$. There is a noticeable decrease in the best-fitting value of γ with an increased penalty (i.e. a stronger shape prior). We find it intriguing that we prefer an inverted temperature-density relation with $\gamma < 1.0$ only for a high penalty, but the constraints are so weak that we cannot draw any solid conclusions from them.

Constraints on the other parameters at 1σ were similar for both penalties. Those for the low penalty were: $50000 > T_0 > 35000$ K, $\tau_{\mathrm{eff}} = 0.33 \pm 0.03$, a slope of $\tau_{\mathrm{eff}}^S = 3.3 \pm 0.3$, $h = 0.7 \pm 0.15$ and $\Omega_M = 0.25 \pm 0.04$. Finally, constraints on the noise parameters largely reproduce the priors (listed in Sec 5.2.7). Our results mirror those of Viel et al. (2009); we have therefore verified that those results are not biased by a shape prior on the power spectrum. The constraints of this work and Viel and Haehnelt (2006) on the IGM temperature, T_0, prefer a larger central value than that obtained by McDonald et al. (2005a). However, McDonald et al. (2005a) imposed a prior of $T_0 = 20000 \pm 2000K$, derived from analysis of the flux probability distribution function of high-resolution quasar spectra, so a direct comparison is not possible. For further discussion of this intriguing result, we refer the interested reader to Section 5 of Viel et al. (2009).

5.4.2 Simulated constraints from BOSS

Unlike current data, our simulated BOSS data shows a well-defined maximum in the CV score. In Fig. 5.6 we show the constraints using this optimal penalty, together with our input power spectrum. The input data is reconstructed very well, within an envelope of roughly 0.4×10^{-9}; a precision comparable to that of a CV reconstruction from WMAP data (Verde and Peiris, 2008). Even though our simulated power spectrum is nearly scale-invariant, we do not recover a very high optimal penalty. This is a feature of our approach; unless the data is noiseless, not all oscillations in the power spectrum will be ruled out, and the optimal penalty is one which allows for them while being consistent
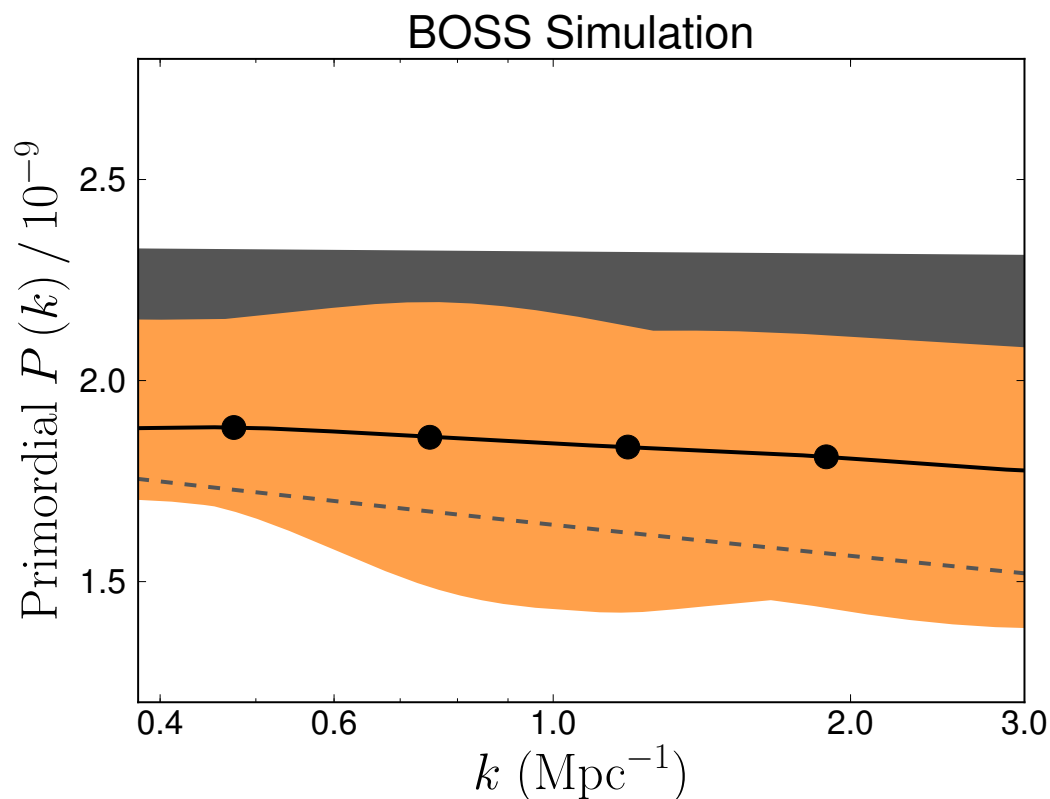
BOSS Simulation

Figure 5.6 Constraints on the power spectrum for simulated BOSS data. Black circles show the positions of the knots, normalized to match the input power spectra (black line). The orange region shows the top 68% of likelihoods from BOSS-quality Lyman-α data. The grey region shows an extrapolation of the 1σ results from WMAP data to these scales, and the grey dashed line shows its lower extent.
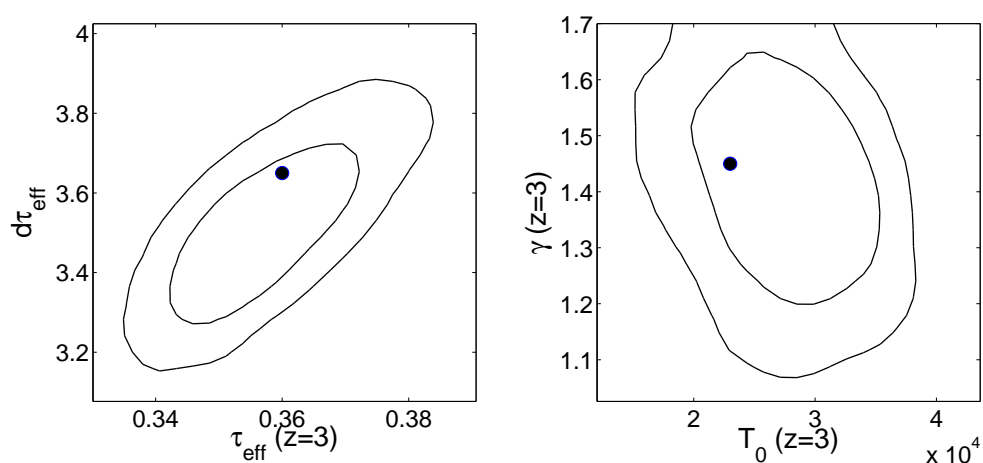

Figure 5.7 Joint 2D posterior constraints on the thermal history using forecast BOSS data. Input parameters are marked by black dots. Contours are drawn at 68 and 95 percent CL. See Section 5.2.3 for definitions of the thermal parameters.

with experimental noise.

Our method was designed to extract $P(k)$, and so the penalty may not be entirely optimal for the derivative. Even given this, our constraints of $0.7 < n_s < 1.2$ are still comparatively weak. However, even this constraint could be useful to test for potential systematics, or in combination with other data sets. One other important data set will be the power spectrum of the cross-correlation of the flux (Viel et al. 2002; McDonald and Eisenstein 2007; Slosar et al. 2009), which BOSS is expected to measure for the first time. Estimating the power of combined constraints is beyond the scope of this Chapter, but it could be considerable.

Fig. 5.7 shows the thermal parameters as reconstructed from BOSS data. We have correctly reconstructed our input, as marked by the black dots. The reconstructed h and Ω_M were also consistent with their input values; $\Omega_M = 0.27 \pm 0.02$ (input: 0.267), $h = 0.74 \pm 0.05$ (input: 0.72).

Marginalized constraints on the thermal and noise parameters are almost a factor of two better for BOSS than for current data. We have assessed the impact that further information about the thermal history of the IGM would have on cosmological constraints, imposing priors corresponding to present and reasonable near-future measurements:

$$\tau_{\text{eff}} = 0.36 \pm 0.11 \,, \qquad\qquad \tau_{\text{eff}}^{S} = 3.65 \pm 0.25 \,,$$
$$T_0 = 23000 \pm 3000K \,, \qquad\qquad \gamma = 1.45 \pm 0.2 \,.$$

Constraints on the mean optical depth are from Kim et al. (2007). For the temperature of the IGM, we follow Becker et al. (2011) and assume a future IGM study has determined γ to the required precision.

The effect of the primordial power spectrum evolves with redshift in a different way to T_0 and γ. Hence, sufficiently accurate data can break degeneracies between them. For τ_{eff}, the constraints from BOSS are already much tighter than our prior from Kim et al. (2007), so this prior provides no additional information. Overall, therefore, the extra information provided by our thermal priors has no significant effect on our reconstruction of the primordial power spectrum.

5.5 Discussion

In this work, we have performed a minimally parametric reconstruction of the primordial power spectrum, using Lyman-α data. This is an extension of McDonald et al. (2005a); Viel and Haehnelt (2006), who used Lyman-α data to measure the amplitude and slope of the primordial power spectrum on small scales, assuming that it had a power-law shape. Using a highly prescriptive model to fit data, even if it is physically motivated, can hide systematic effects, which may bias the recovered parameters in a manner which is hard to detect unless the bias is extremely large. Further, it is vital to go beyond parameter estimation and test the underlying model of the primordial power spectrum. This can in principle be achieved with a minimally parametric reconstruction framework coupled with a scheme for avoiding over-fitting the data.

Peiris and Verde (2010), who attempted such a reconstruction including Lyman-α data, assumed that the power spectrum could be well approximated by an amplitude, a power-law slope and its running across the scales probed by the Lyman-α forest. In their
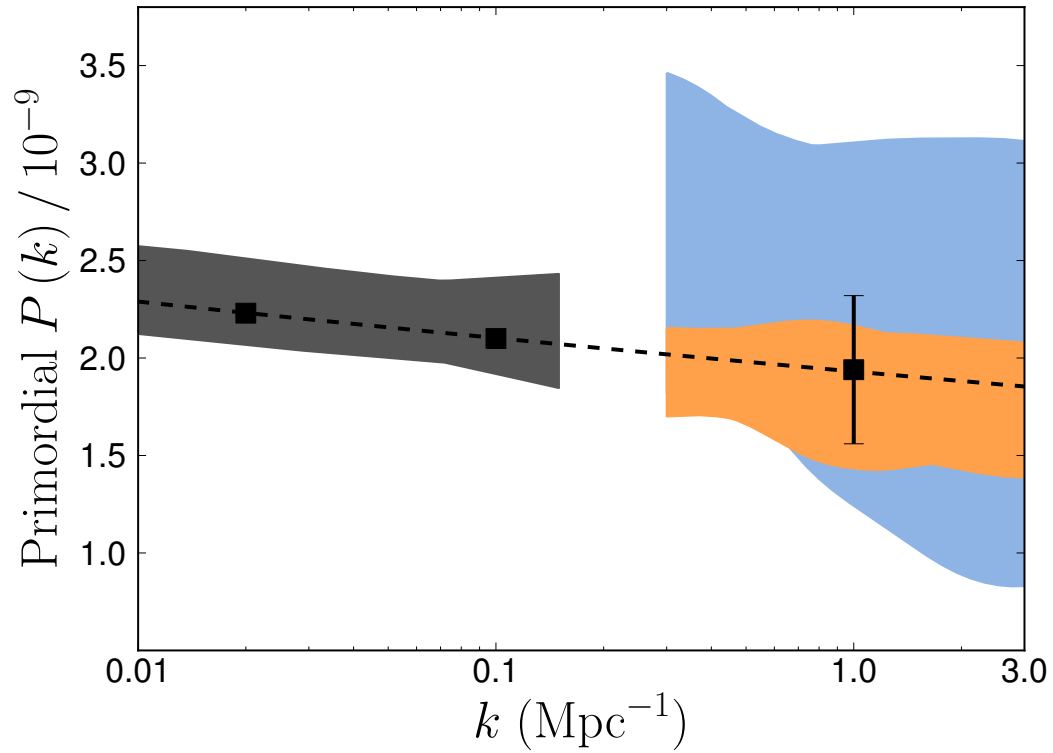
Figure 5.8 Comparison of our constraints. Blue is from current data; orange is our BOSS forecast. The grey region shows part of a reconstruction using both the CMB data and galaxy clustering measured by SDSS (Peiris and Verde 2010). The black squares show two knots used in the earlier reconstruction, while the black error bar shows 1σ constraints on power spectrum amplitude from parameter estimation (Viel et al. 2009). The dashed line shows the extrapolated WMAP best-fitting power spectrum.

analysis this assumption was justified as the Lyman-α data was treated as a single point and combined with CMB and galaxy survey data to reconstruct the power spectrum over a wide range of scales.

However, the only likelihood function available up to now contained a power-law assumption about the primordial power spectrum shape, making it impossible to treat the Lyman-α data in a fully minimally parametric manner. We remedy this, performing a large suite of numerical simulations to construct a new likelihood function. The primordial power spectra thus emulated have considerable freedom in their shapes, specified by cubic smoothing splines. This provides the first ingredient for a minimally-parametric reconstruction scheme.

The second ingredient, as mentioned above, is to avoid fitting the noise structure of the data with superfluous oscillations. To this end, our method uses cross-validation (CV) to reconstruct the level of freedom allowed by the data. CV is a statistical technique which quantifies the notion that a good fit should be predictive. Schematically, it is a method of jackknifing the data as a function of a "roughness" penalty. A small penalty thus allows considerable oscillatory structure in the power spectrum shape, while a larger penalty specifies a smoother shape. This penalty term thus performs the same function as a prior on the smoothness of the power spectrum. Jackknifing the data then tests the predictivity of the smoothing prior, choosing as the optimal penalty the one that maximizes predictivity. For technical details see Section 5.2.1.

For the Lyman-α current data from SDSS (McDonald et al. 2006), CV yields no significant preference for any particular penalty. In the context of CV, this indicates that no penalty is more predictive or favoured over any other; in other words, the data are not sufficiently powerful to accurately reconstruct the strength of the shape prior.

The minimally parametric method thus provides no evidence for features in the power spectrum in the current data, and our results are fully consistent with a scale-invariant power spectrum. The best-fitting amplitude of the power spectrum is, as in previous work, slightly higher than that extrapolated from WMAP (Komatsu et al. 2011). However, because the data do not contain sufficient statistical power to reconstruct the power spectrum shape, our error bars are extremely large. An analysis that uses different statistical techniques, such as Bayesian evidence (Jeffreys 1961), could provide further insight, but is beyond the scope of this Chapter.

In the not so distant future, the first data from a new Lyman-α survey, BOSS (Schlegel et al. 2009), will be made available. We simulate a flux power spectrum and covariance matrix for BOSS, with an 80 fold increase in statistical power over the current data. In this case we successfully reconstruct the power spectrum, using CV to find an optimal penalty. The parameters we extract using CV are completely consistent with the inputs to the simulation, and the resulting constraints are comparable to those achieved by performing CV reconstruction using WMAP data (Verde and Peiris 2008). We verify that statistical error is the factor preventing us from finding an optimal penalty for current data by simulating a power spectrum identical to BOSS, but with wider error bars, again failing to find an optimal penalty.

Finally, we show that adding plausible future data on the smallscale thermodynamics of the IGM to BOSS does not significantly improve constraints on the primordial power spectrum. The simulated BOSS data are sufficiently powerful on their own to break degeneracies between the IGM and cosmological parameters, and are limited by statistical

error rather than systematic uncertainty.

We have not considered the impact of the information BOSS is expected to provide on the transverse flux power spectrum. This will probe larger scales than our current work, offering a longer baseline and thus better sensitivity to the overall shape of the power spectrum. However, applying the present technique to the improved data set would require simulations probing much larger scales, hence greatly increasing the numerical requirements.

Fig. 5.8 shows the constraints from BOSS in comparison to those Peiris and Verde (2010) obtained by reconstructing the power spectrum using the CMB and the matter power spectrum from SDSS. By combining BOSS data with other probes (Seljak et al. 2005, 2006), such as galaxy clustering, the CMB, and the transverse flux power spectrum, we will be able to accurately reconstruct the shape of the power spectrum on scales of $k = 0.001 - 3\,\mathrm{Mpc}^{-1}$, probing ten e-folds of inflation.

"On two occasions I have been asked [by members of Parliament]: 'Pray, Mr. Babbage, if you put into the machine wrong figures, will the right answers come out?' I am not able rightly to apprehend the kind of confusion of ideas that could provoke such a question."

– Charles Babbage

6

Massive Neutrinos and the Non-linear Matter Power Spectrum

Abstract

We perform an extensive suite of N-body simulations of the matter power spectrum, incorporating massive neutrinos in the range $M_\nu = 0.15 - 0.6$ eV, and probing deep into the non-linear regime. We compare our results with the widely used HALOFIT approximation, and find that in the strongly non-linear regime it significantly over-predicts the suppression due to the free-streaming of neutrinos. The maximal discrepancy occurs at $k \sim 1 h\,\mathrm{Mpc}^{-1}$, and is at the level of $2 - 5\%$ of the matter power spectrum. Most published constraints are not affected, as they have used HALOFIT only in the linear or mildly non-linear regime. However, future galaxy and weak lensing surveys will probe these scales, and will also require extremely accurate theoretical predictions of the matter power spectrum in order to put robust constraints on the neutrino mass.

6.1 Introduction

A variety of solar, atmospheric, reactor and accelerator neutrino experiments have firmly established the existence of neutrino flavour oscillations, implying that neutrinos have at least two non-zero mass eigenstates (Becker-Szendy et al., 1992; Fukuda et al., 1998; Ahmed et al., 2004). These experiments yield a lower limit on the sum of the neutrino masses; $M_\nu > 0.05$ eV. Classical β-decay experiments provide an upper limit on the mass of the electron neutrino of 2.2 eV (Lobashev, 2003), while future experiments may reach constraints of 0.2 eV (Wolf and KATRIN collaboration, 2010). However, particle physics constraints on the masses of the other neutrino species are much weaker; $\lesssim 200$ keV for the muon neutrino (Assamagan et al., 1996) and $\lesssim 20$ MeV for the tau neutrino (Barate and ALEPH collaboration, 1998).

To distinguish the neutrino mass hierarchy, which is of particular interest when formulating theories of particle physics beyond the standard model, we need far stronger constraints on the total neutrino mass. Fortunately, these can be obtained by weighing the cosmic neutrino background. Simply requiring $\Omega_\nu < 1$ and fixing the Hubble constant implies $M_\nu \lesssim 60$ eV (Gershtein and Zel'dovich, 1966). The cosmic microwave background (CMB; e.g. Komatsu et al. (2011)) and measurements of the matter power spectrum provide more stringent constraints, presently around $M_\nu \lesssim 0.6$ eV (Reid et al., 2010). The sensitivity of the matter power spectrum to neutrino mass comes about because the free-streaming of the neutrino background suppresses the growth of structure. On large scales, where the growth of structure is still linear, this effect is relatively straightforward to quantify (see, e.g. Bond et al. (1980); Ma and Bertschinger (1994)) and thus can be done using standard Boltzmann codes (Lewis et al., 2000). On smaller scales it is also necessary to account for the non-linear growth of structure. This is best studied using numerical simulations, but is often approximated analytically to facilitate marginalizing over a wide range of cosmological parameters.

One such analytic approximation, widely used in recent years, is HALOFIT (Smith et al., 2003). HALOFIT models the non-linear matter power spectrum by assuming that the growth of structure is due to a combination of the density structure of individual DM haloes and the clustering of these DM haloes relative to each other. It has been calibrated against a wide range of CDM simulations, superseding earlier prescriptions based on the stable clustering hypothesis (Hamilton et al., 1991; Peacock and Dodds, 1996). HALOFIT is claimed to be accurate to $\lesssim 5 - 10\%$ (Smith et al., 2003); however, it has not been calibrated for non-zero neutrino masses. See also the purely analytic description of Abazajian et al. (2005).

Future matter power spectrum surveys aim to make measurements accurate to a few percent. At this level of precision, systematic error from theoretical predictions must be tightly controlled. Fortunately, simulation technology has advanced to the point where it is possible to model the impact of neutrinos to these accuracies, despite the difficulties presented by the large thermal dispersion of light neutrinos (Brandbyge et al., 2008; Viel et al., 2010). It appears thus timely to directly compare HALOFIT predictions of the non-linear matter power spectrum with full numerical simulations including massive neutrinos. For this purpose we use the recently developed version of the Tree-SPH code GADGET-3 (Viel et al., 2010), which can include massive neutrinos. We will discuss our methods in Section 6.2, present our results in 6.3, and conclude in Section 6.4.

6.2 Modelling the matter power spectrum

There has been significant work over the years developing numerical simulations and numerically calibrated fitting formulae to model the non-linear matter power spectrum. In Section 6.2.1 we will review the predicted effect of neutrinos in linear theory. Then in Section 6.2.2 we will discuss the strategies used for simulating the effect of massive neutrinos, describing both particle and grid-based methods. In Section 6.2.3, we will summarise HALOFIT, a widely used analytic model for the non-linear matter power spectrum (Smith et al., 2003). Finally, Section 6.2.4 gives details of our numerical simulations.

6.2.1 The effect of massive neutrinos in linear theory

In this Section, we will briefly review linear theory, which is discussed in more depth in, for example Kolb and Turner (1990); Bond et al. (1980); Eisenstein and Hu (1999); Lesgourgues and Pastor (2006) (and references therein). Neutrinos in the very early universe are relativistic and tightly coupled to the primordial plasma. They decouple while still relativistic and then redshift adiabatically. Once non-relativistic, massive, weakly-interacting neutrinos behave qualitatively as a species of hot/warm dark matter, suppressing fluctuations on scales smaller than their thermal free-streaming length.

This suppression has a distinctive shape in the linear regime. Consider neutrinos with a total energy density Ω_ν and contribution to the dark matter f_ν, given by

$$\Omega_\nu = \frac{M_\nu}{93.8\,h^2\mathrm{eV}} \,, \qquad\qquad f_\nu = \frac{\Omega_\nu}{\Omega_M} \,, \qquad (6.1)$$

where Ω_M is the present day matter density, M_ν is the total neutrino mass in all species, and h is the Hubble constant.

The residual thermal velocity of the neutrinos at redshift z is

$$v_{\mathrm{th}} \approx 150(1+z)\left[\frac{1\,\mathrm{eV}}{M_\nu}\right]\,\mathrm{km/s}\,. \qquad (6.2)$$

By analogy with the Jeans length, we can define a free-streaming scale

$$k_{\mathrm{FS}} = \sqrt{\frac{2}{3}}\frac{H(t)a(t)}{v_{\mathrm{th}}(t)}\,. \qquad (6.3)$$

Neutrinos cannot cluster on scales smaller than k_{FS}, and the matter power spectrum is suppressed by a constant factor proportional to the neutrino energy density, f_ν. The Boltzmann equation can be solved numerically to show that the fractional change in the power spectrum for $f_\nu < 0.07$ is

$$\frac{\delta P}{P} \approx -8f_\nu\,. \qquad (6.4)$$

On larger scales, $k < k_{\mathrm{FS}}$, the neutrinos have free-streamed only for a portion of the cosmic history, and so there is a smaller amount of suppression. The largest scales at which the neutrinos have an effect is defined by the wave-number for which they free-streamed only immediately after becoming non-relativistic

$$k_{\mathrm{nr}} \sim 0.018\sqrt{\Omega_M\left[\frac{M_\nu}{1\,\mathrm{eV}}\right]}\,h\,\mathrm{Mpc}^{-1}\,. \qquad (6.5)$$

Today, neutrinos in the observationally acceptable range have $k_{\mathrm{nr}} \sim 10^{-3}h\,\mathrm{Mpc}^{-1}$, placing the free-streaming effect at the scales probed by galaxy and Lyman-α measurements of the matter power spectrum. A number of bounds on the neutrino mass have been published using this effect. See, for example Elgaroy et al. (2002); Seljak and Slosar (2006); Tegmark et al. (2006); Mantz et al. (2010); Reid et al. (2010).

6.2.2 Incorporating massive neutrinos into numerical simulations

Numerical simulations incorporating massive neutrinos have a long history (Klypin et al., 1993; Ma and Bertschinger, 1994), but only recently has the perturbative effect of massive neutrinos as a component in a mainly ΛCDM universe been examined (Brandbyge et al., 2008; Brandbyge and Hannestad, 2009, 2010; Viel et al., 2010). We shall build on the work of Viel et al. (2010), who incorporated massive neutrinos into GADGET-3, an extended version of the code described in Springel (2005). The methods used to simulate massive neutrinos have been described in detail in Viel et al. (2010). Here we shall briefly review the important points, and refer to that work for details and validation tests.

Particle-based simulations (Brandbyge et al., 2008) model neutrinos as a separate low-mass particle species, with an initial distribution identical to that of the dark matter, but significant thermal dispersion. The simulation box is evolved under gravity as normal, except that the short-range tree force is not computed for the neutrino component, greatly increasing the speed of the simulations. This does not significantly impact accuracy, because their thermal dispersion prevents neutrinos clustering on small scales.

The neutrino species is not allowed to alter the timestep of the evolution. Thus, the timestepping does not rigourously obey the Courant conditions for the neutrinos. The effect of this approximation on results was examined by Brandbyge et al. (2008) and found to be negligible. This is because structure growth is being driven by the dominant cold dark matter component, so it is sufficient to use the timestep associated with this species.

One potential problem with particle based simulations is thermal shot noise from the finite sampling of the neutrino distribution (Wang and White, 2007), especially for the higher thermal velocities encountered at high redshift or for particularly low mass neutrinos. We deal with this by explicitly checking the effect of shot noise, and increasing the neutrino particle number should it be non-negligible.

Alternatively, the effect of neutrinos on the dark matter can be modelled by adding their gravitational force to that of the dark matter on the particle-mesh grid (Brandbyge and Hannestad, 2009). In this grid-based approach, the neutrinos are not followed directly, and their clustering is assumed to be well-described by linear theory. This allows us to model the effects of neutrinos on the non-linear growth in the dark matter, while avoiding shot noise and the need to deal directly with large thermal velocities. However, because the neutrinos are evolved using linear theory, grid-based simulations neglect any non-linear evolution in the neutrino component itself, including backreaction from non-linearities in the dark matter.

One limitation of both strategies is that the effective mass of the neutrino is constant throughout the simulation. At high redshift neutrinos may be partially relativistic, decreasing their effective mass. If the initial redshift is too large, simulations such as ours which do not account for this effect will over-estimate the suppression due to neutrinos by a constant factor in the linear regime at high redshift. This places an upper bound on the initial redshift which depends on the neutrino mass. Our simulations aim to saturate this bound, so that perturbations are initially in the linear regime as far as possible. Some mechanism for incorporating relativistic corrections to our neutrinos could be useful for future work.

6.2.3 The HALOFIT model for the non-linear matter power spectrum

HALOFIT is a widely used formula for estimating non-linear growth from a given linear theory power spectrum. It has been calibrated on a wide range of numerical simulations (Smith et al., 2003), without, however, explicitly incorporating neutrinos. In this Section we will give a brief summary; for full details we refer to Smith et al. (2003).

HALOFIT models non-linear growth in the spirit of the stable clustering hypothesis (Hamilton et al., 1991; Peacock and Dodds, 1996). It assumes that the non-linear growth deep in the nonlinear regime is independent of the details of the linear power spectrum; a reflection of the principle that the growth of halos should depend only on the local physics at the scale of the halo, and not the internal details of the pre-collapse material, or the larger-scale clustering of matter. Furthermore, it assumes that all redshift information is contained within the linear growth function. The non-linear correction is assumed to be a function of the linear power spectrum at the present, and not to depend on the historical growth rate. Both of these approximations have been shown to be reasonably accurate for cold dark matter by the success of fitting formulas inspired by them. However, neutrinos may evolve separately to the CDM, meaning that the combined power spectrum should strictly depend on non-linear growth in both components.

The ideas of stable clustering are not the only possible approach. Saito et al. (2008) proposed a formula based on one-loop perturbation theory, while Lesgourgues et al. (2009) have modelled the effects of neutrinos with an approach based on a time renormalisation group flow. However, these formulae are designed for when growth is still only weakly non-linear, $k < 0.3h\,\mathrm{Mpc}^{-1}$, while HALOFIT has the advantage that it may be extended deep into the non-linear regime.

HALOFIT splits the non-linear power spectrum into two components; a quasilinear term Δ_{Q}^2, which is analogous to the halo correlation term, and a non-linear term Δ_{H}^2. The dimensionless power spectrum is defined as

$$\Delta^2 = \frac{k^3 P(k)}{2\pi^2}\,. \tag{6.6}$$

We can define the non-linear scale, k_σ, to be the wavenumber where the perturbation amplitude becomes greater than unity. Non-linear growth will be important when $k > k_\sigma$, with an amplitude supposed to depend on $y = k/k_\sigma$, the slope of the power spectrum, and the matter fraction. The form of Δ_{H} used by HALOFIT is

$$\Delta_{\mathrm{H}}^2(k) = \frac{\Delta_{\mathrm{H}}^{2\prime}(k)}{1 + \mu(n)/y + \nu(n)/y^2} \tag{6.7}$$

$$\Delta_{\mathrm{H}}^{2\prime}(k) = \frac{a(n)y^{3f_1(\Omega)}}{1 + b(n)y^{f_2(\Omega)} + [c(n)f_3(\Omega)y]^{3-\gamma(n)}} \tag{6.8}$$

$$\text{Here} \quad 3 + n = \left.\frac{d\ln\sigma^2}{d\ln R}\right|_{k=k_\sigma}. \tag{6.9}$$

n is the effective spectral index of the power spectrum, and a, b, c, μ and ν are functions of n. Note that $\nu(n)$ is not related to neutrinos. $f_{1,2,3}$ are functions of Ω_M, the total matter energy density. We refer readers interested in the exact values of these functions to Appendix C of Smith et al. (2003).

As well as modelling the non-linear collapse of halos, HALOFIT incorporates a quasi-linear term, for the regime where non-linear growth has started, but does not yet dominate. This is given as an enhancement to the linear power spectrum, and incorporates an exponential cut-off, ensuring it is negligible on small scales. The form of the quasilinear term is

$$\Delta_{\mathrm{Q}}^2(k) = \Delta_{\mathrm{L}}^2(k) \left[\frac{(1 + \Delta_{\mathrm{L}}^2(k))^{\beta(n)}}{1 + \alpha(n)\Delta_{\mathrm{L}}^2(k)} \right] \exp[-f(y)] , \tag{6.10}$$

where β and α are again functions of n and $f(y) = y/4 + y^2/8$. Δ_{L}^2 is the dimensionless linear theory power spectrum. For the values of these functions, we again refer to Appendix C of Smith et al. (2003). The total power is given by summing the two contributions

$$\Delta_{\mathrm{NL}}^2 = \Delta_{\mathrm{Q}}^2 + \Delta_{\mathrm{H}}^2 . \tag{6.11}$$

Massive neutrinos suppress the linear power spectrum, increasing the non-linear scale and retarding non-linear growth. HALOFIT estimates the magnitude of this effect, , but neglects any effect of the neutrinos on the process of halo collapse, as well as the effects of non-linear growth on the neutrinos themselves.

6.2.4 The numerical simulations

Our fiducial simulations have 512^3 dark matter particles in a box of $150\mathrm{Mpc}\,h^{-1}$. We tested $M_\nu = 0.15$ eV, 0.3 eV and 0.6 eV, running 4 simulations for each value of M_ν with different structure seeds, to suppress cosmic variance. To check we had correctly sampled the large scale modes, to explore the linear regime and to compare directly to Viel et al. (2010), we ran simulations with box sizes of $512\mathrm{Mpc}\,h^{-1}$. To facilitate comparison with Brandbyge et al. (2008) and Viel et al. (2010), our fiducial cosmology had $\Omega_M = 0.3$, a spectral index $n_s = 1.0$, the Hubble parameter $h = 0.7$, and amplitude of the primordial power spectrum $A_{\mathrm{s}} = 2.27 \times 10^{-9}$. This implies that a simulation without neutrinos has $\sigma_8 = 0.87$ at $z = 0$, but a simulation with $M_\nu = 0.3$ eV has $\sigma_8 = 0.8$. The effects of varying the cosmology was checked with three pairs of grid-based simulations. Each pair had one simulation without neutrinos, and one with $M_\nu = 0.6$ eV. Both simulations in each pair had one of $A_{\mathrm{s}} = 2.0 \times 10^{-9}$, $n_s = 0.9$ or $h = 0.75$.

Grid-based simulations used a grid of 512^3. The initial redshifts of our simulations were $z_i = 99$ for $M_\nu = 0.6$ eV, $z_i = 49$ for $M_\nu = 0.3$ eV and $z_i = 24$ for $M_\nu = 0.15$ eV, for the reasons discussed in Section 6.2.2. Our initial conditions were generated using the Zel'dovich approximation, and the linear theory power spectrum as calculated by CAMB (Lewis et al., 2000). The transfer function used for our cold dark matter (CDM) particles was a weighted average of the linear theory transfer functions for CDM and baryons, to account for the slight difference between them. Neutrinos and CDM had identical initial random phase information, to simulate adiabatic initial conditions. The realisation of neutrino structure for the grid-based simulations was identical to that in the initial conditions for the particle simulations, and the power spectrum was given by linear theory, calculated using CAMB.

To check the numerical convergence of our results, we ran two simulations with 2×640^3 particles; one without neutrinos, and one with $M_\nu = 0.6$ eV. The effect of shot noise on the particle based simulations was checked with two simulations using 1024^3 neutrino
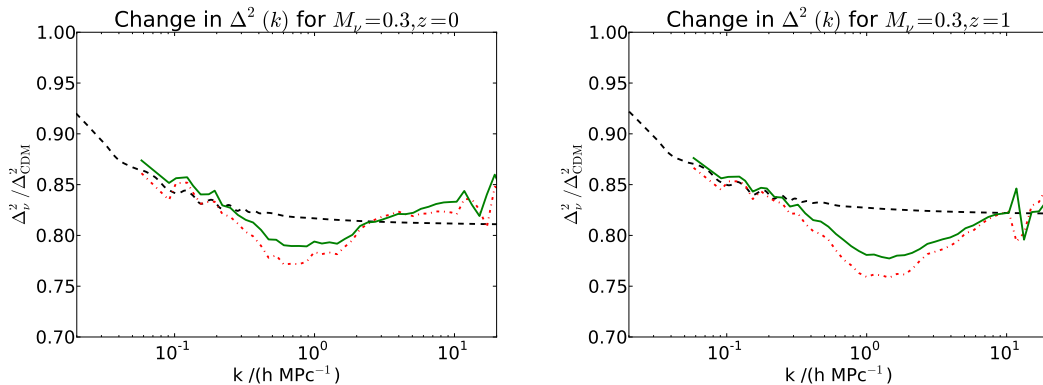
Figure 6.1 The difference in the matter power spectrum at $z = 0$ (Left) and $z = 1$ (Right) between simulations with and without massive neutrinos ($M_\nu = 0.3$ eV). Solid green lines show results for simulations using particles, while dotted red lines show results from grid-based simulations. The black dashed lines show the predicted effect from linear theory. Box size was $150 \mathrm{Mpc}\,h^{-1}$, and the initial redshift was $z = 49$.

particles, one with $M_\nu = 0.15$ eV, and one with $M_\nu = 0.6$ eV. The impact of the resolution of the particle mesh grid on the grid-based simulations was checked by running one simulation with a 1024^3 grid.

We checked the effect of baryonic physics with two grid-based simulations incorporating baryons, which modelled radiative cooling and star formation using the default model of GADGET-3, described in Katz et al. (1996). Recently van Daalen et al. (2011) found that AGN feedback can affect the matter power spectrum at the level of a few percent for $k \approx 2h\,\mathrm{Mpc}^{-1}$. We did not include this effect; however, they found it to be independent of cosmological parameters, and thus presumably neutrino mass, so it should not affect our results.

6.3 Results

Figure 6.1 compares our results to linear theory for a neutrino mass of $M_\nu = 0.3$ eV and a box size of $150 \mathrm{Mpc}\,h^{-1}$. The numerical results shown are the average from 4 simulations using different structure seeds. As discussed before, linear theory predicts an increasing suppression culminating in an almost flat plateau on small scales. In the numerical simulations, the flat plateau is replaced by a dip, with a minimum at $k \sim 1h\,\mathrm{Mpc}^{-1}$. This distinctive "spoon shape" was also found by Brandbyge et al. (2008) and Viel et al. (2010).

We can understand this shape by considering the effect of neutrinos on non-linear growth. The suppression caused by massive neutrinos delays the onset of non-linear growth and increases the wave-number of the non-linear scale. Between the non-linear scale for the simulation without neutrinos and the non-linear scale for the simulation incorporating neutrinos, the suppression will be enhanced by the absence of non-linear growth. On the smallest scales, the non-linear growth lost as a consequence of the massive neutrinos becomes an ever-smaller fraction of the total, and the extra suppression decreases. Eventually non-linear growth comes to dominate the linear effect, and the
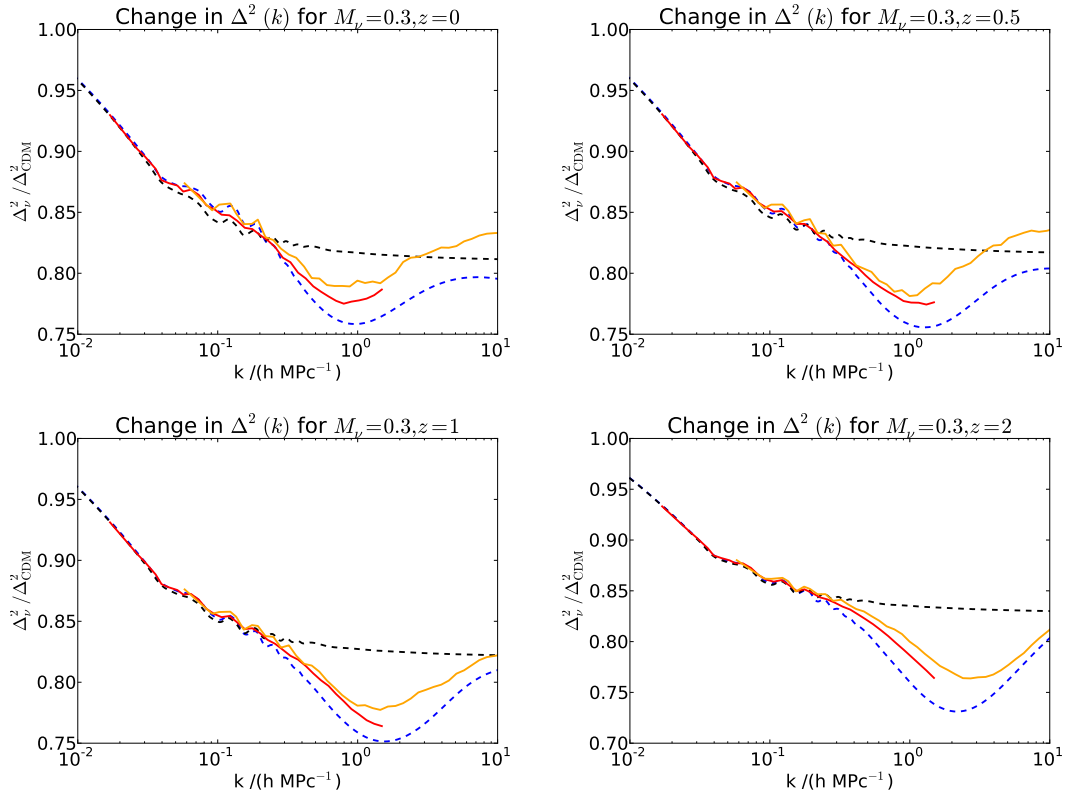
Figure 6.2 The effect of massive neutrinos on the matter power spectrum for a neutrino mass of $M_\nu = 0.3\,\mathrm{eV}$. Solid lines show the difference between simulations with and without massive neutrinos, for both $512\mathrm{Mpc}\,h^{-1}$ (red) and $150\mathrm{Mpc}\,h^{-1}$ (orange) boxes. Initial redshift was 49. The blue dashed line shows the estimated difference using HALOFIT, while the black dashed line shows the prediction from linear theory.
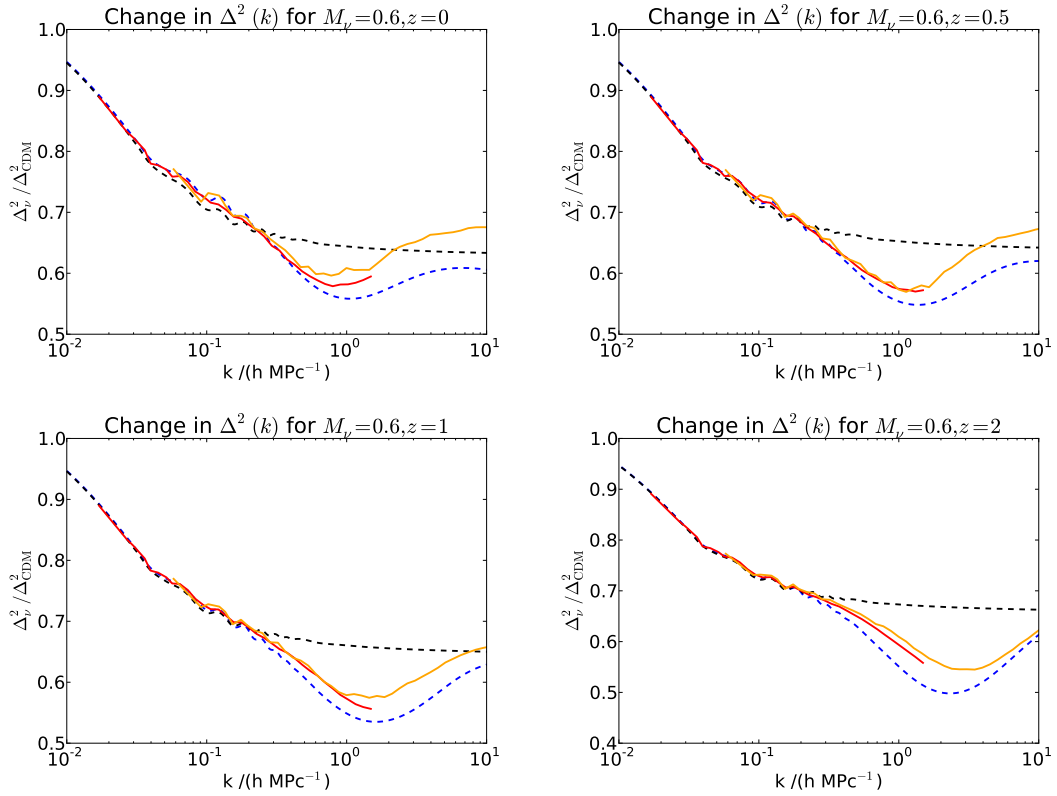
Figure 6.3 The effect of massive neutrinos on the matter power spectrum for a neutrino mass of $M_\nu = 0.6$ eV. Solid lines show the difference between simulations with and without massive neutrinos, for both $512\mathrm{Mpc}\,h^{-1}$ (red) and $150\mathrm{Mpc}\,h^{-1}$ (orange) boxes. Initial redshift was 99. The blue dashed line shows the estimated difference using HALOFIT, while the black dashed line shows the prediction from linear theory.
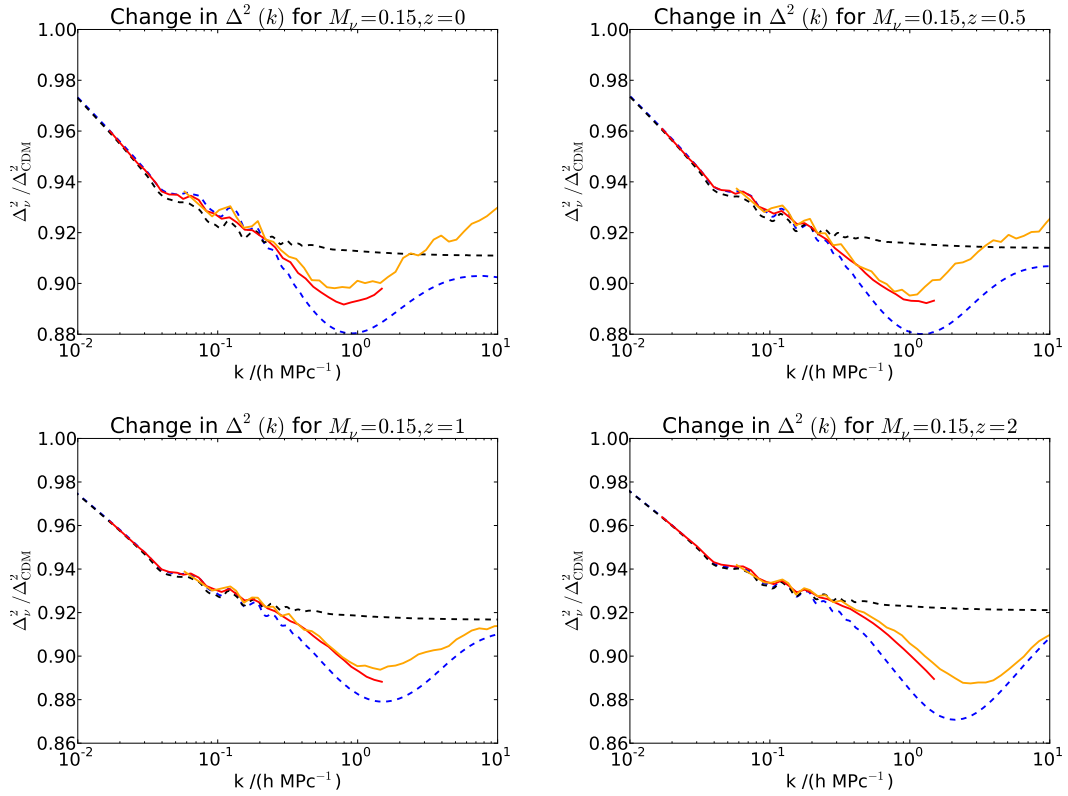
Figure 6.4 The effect of massive neutrinos on the matter power spectrum for a neutrino mass of $M_\nu = 0.15$ eV. Solid lines show the difference between simulations with and without massive neutrinos, for both $512 \mathrm{Mpc}\, h^{-1}$ (red) and $150 \mathrm{Mpc}\, h^{-1}$ (orange) boxes. Initial redshift was 24. The blue dashed line shows the estimated difference using HALOFIT, while the black dashed line shows the prediction from linear theory.

suppression is less than predicted by linear theory.

Our results show a slightly greater maximal suppression than found by Viel et al. (2010) for their fiducial 60Mpc h^{-1} boxes. This is due to our use of a larger box; a 60Mpc h^{-1} box, while preferred for the Lyman-α forest, does not include the largest non-linear scales at $z = 0$, and so misses some of the suppression of the matter power spectrum. Furthermore, we have used a somewhat higher initial redshift than that work, which also contributes.

As mentioned in Section 6.2.4, we performed checks for numerical convergence. Our 150Mpc h^{-1} box simulations are converged to the grid scale of $k = 7\,h\,\mathrm{Mpc}^{-1}$. The effect of resolution, baryons and, for grid-based simulations, the resolution of the particle-mesh grid, all affected the matter power spectrum at the level of $1-2\%$ on these scales. We also simulated two different box sizes, 512 and 150Mpc h^{-1}, which were in good agreement on the largest scales probed by the smaller box.

6.3.1 Particle and Grid Methods

Figure 6.1 shows comparable output for particle and grid-based simulations. There is a somewhat smaller suppression in the particle simulations, particularly near the scales of greatest suppression, which scale roughly linearly in M_ν.

This discrepancy has again a simple physical explanation; structure growth in the dark matter induces structure growth in the neutrino component, as the less energetic neutrinos fall into the gravitational wells created by the dark matter. This decreases the gap between the fluctuations of the neutrinos and those of the dark matter, leading to a reduced neutrino suppression. Essentially this is a backreaction effect, where the dark matter drags the neutrinos with it, and because it results in part from the non-linear growth, it is not being fully accounted for by linear theory neutrinos.

Because this effect is important at the accuracy we wish to achieve, our main results (Figures 6.2, 6.3 and 6.4) use simulations with particles. We have run simulations with a far higher neutrino number to check explicitly that neutrino shot noise has only a minor impact at the relevant scales, for 150Mpc h^{-1} boxes.

6.3.2 Comparison to the HALOFIT model for the non-linear evolution of the matter power spectrum

Figures 6.2, 6.3 and 6.4 show the main results of this Chapter. Each Figure shows the suppression of the matter power spectrum caused by massive neutrinos at four redshift snapshots; $z = 0, 0.5, 1$ and 2 and compares it to the predicted effect from HALOFIT. Figure 6.2 shows $M_\nu = 0.3$ eV, Figure 6.3 has $M_\nu = 0.6$ eV and finally Figure 6.4 has $M_\nu = 0.15$ eV.

HALOFIT clearly over-predicts the suppression of the matter power spectrum due to massive neutrinos in the non-linear regime. The cause is similar to that discussed in Section 6.3.1; HALOFIT includes the neutrinos only through the effect of linear theory neutrinos on the non-linear scale, and neglects any back-reaction from the dark matter. This interpretation is confirmed because HALOFIT agrees very well with the grid-based simulations.

The largest discrepancy, of $2 - 5\%$, occurs at $k \sim 1h\,\mathrm{Mpc}^{-1}$. The location of the maximal suppression in the numerical simulation moves to larger scales at lower redshifts, an effect which is not captured by HALOFIT, although it was present to some extent in our grid-based simulations. Furthermore, the amplitude of the suppression decreases with redshift, which is to be expected from our discussion in 6.3.1; as non-linear growth occurs in the dark matter more neutrinos fall into the gravity wells. A cross-over redshift occurs at $z = 1$; on very small scales the suppression here is equal to that of linear theory, while at lower redshifts it is less. Note also that in the quasilinear regime, $0.05 < k < 0.2h\,\mathrm{Mpc}^{-1}$, the simulations clearly agree with HALOFIT rather than linear theory, at redshifts where linear theory conflicts with HALOFIT. The dependence of our results on M_ν is well-captured by a linear relation, with the maximal suppression for a given redshift being proportional to f_ν, but dependent on redshift.

We have performed some simulations with parameters varied from our fiducial cosmology, as described in 6.2.4. The results of these simulations were similar to those for our fiducial cosmology, and agree with the results discussed in Viel et al. (2010). As in that work, the effect of Ω_M can be incorporated wholly through f_ν.

6.4 Discussion

We have performed a suite of detailed simulations of the matter power spectrum, incorporating the effects of massive neutrinos using both particle and grid-based methods, but focussing on particle simulations. We have compared our results to those predicted by the widely used HALOFIT formula.

Observational constraints on the neutrino mass prior to this work have either used linear theory (Reid et al., 2010) or HALOFIT to estimate the effect of massive neutrinos. Thomas et al. (2010) used photometric redshifts at $z = 0.45 - 0.65$, on scales $k \lesssim 0.2h\,\mathrm{Mpc}^{-1}$. This is in the quasilinear regime, where we find HALOFIT to be accurate. Hence our results should not significantly change their published constraints.

However, while HALOFIT performs well on quasilinear scales, in the fully non-linear regime it significantly over-predicts the suppression due to the neutrino mass. Furthermore, it fails to capture redshift dependence in the amplitude and scale of the peak neutrino suppression. Our improved exploration of the non-linear regime will certainly be of importance for future galaxy (Wang et al., 2005; Schlegel et al., 2009), CMB lensing (Gratton et al., 2008), and weak lensing (Beaulieu et al., 2010) surveys. Abazajian et al. (2011) have published an overview of forecast constraints on the neutrino mass, many of which reach the $M_\nu \sim 0.1$ eV range. This will require an accuracy of around a percent in theoretical predictions of the power spectrum, making the use of a more precise formula than HALOFIT essential. In future work, we will develop an improved fitting formula for the non-linear regime aimed specifically at these future surveys.

"He first, I following; till my straining sense
Glimpsed the bright burden of the heavenly cars
Through a round hole; by this we climbed and thence
came forth, to look once more upon the stars"
    – Inferno, Canto XXXIV, by Dante Alighieri

7

Conclusions

7.1 Summary

In this Thesis, I have discussed a few ways to use cosmological information to constrain fundamental physics. In Chapters 3 and 4, I looked at inflation, while in Chapters 5 and 6, I examined two aspects of the growth of structure.

Chapter 3 examined predictions for the tensor-scalar ratio r from single field models of inflation with a polynomial potential. The motivation for this work was a then-recent paper (Boyle et al., 2006) claiming that a simple notion of fine-tuning, together with a few assumptions about the inflaton potential, could be used to imply that single field inflation models predicted an observable value for r. I examined this claim further, finding it was sensitive both to the assumptions about the potential and the definition of fine-tuning. Thus, an observable value of r should not be regarded as a solid prediction of single-field inflation.

One of the definitions of fine-tuning used in Chapter 3 was the number of inflaton trajectories which overshot the inflationary region. The need to make the initial kinetic energy of the inflaton small to avoid this is a common problem in inflation, called the overshoot problem. D-brane inflation provides limits, derived from the fundamental physics, on the allowed initial kinetic energy, and on the maximal range of the inflaton. In Chapter 4, I examined the restrictions these limits place on the space of initial conditions, and the proportion of overshoot trajectories. I found that ruling out essentially all such trajectories, and "solving" the overshoot problem requires a tightening of the bound on the initial kinetic energy by several orders of magnitude. This is because the tight restrictions on the inflaton range make overshoot more likely, and this effect dominates the decrease in overshoot from bounds on the initial kinetic energy.

In Chapter 5, I made use of the Lyman-α forest to constrain the shape of the primordial power spectrum. In this work, I did not assume any particular model for the generation of the primordial fluctuations, instead performing a minimally parametric model-independent reconstruction of the power spectrum. This work is a consistency

check of the standard model; if the data is consistent with a power law power spectrum, my more general formalism should recover a range of reconstructions containing it. I found that current data is not sufficient to strongly constrain the power spectrum in a minimally parametric way. However, my forecasts indicated that future data should be able to.

Finally, in Chapter 6, I ran numerical simulations to measure the effect of massive neutrinos on the matter power spectrum. I found that the suppression of power they cause in the non-linear regime differs significantly from previous estimates. Future measurements of the power spectrum on small scales will require this information for accurate neutrino mass constraints.

7.2 Outlook

At present, new work on the inflationary potential is in danger of almost immediate obsolescence by results from the Planck satellite. While there are continuing efforts to understand and complete string theory motivated models of inflation, my future work will be directed more towards constraints from large scale structure. In the near future, I intend to use the simulations discussed in Chapter 6 to create a precise fitting formula for the effect of neutrinos on the matter power spectrum in the non-linear regime, suitable for creating forecasts for upcoming surveys.

My work in Chapter 5 used the power spectrum of the Lyman-α flux along the line of sight to the quasar to probe structure on small scales. However, this is not the only way of extracting information from the Lyman-α forest. With sufficient quasar spectra, it is possible to obtain valuable cosmological information from the cross-correlation function; the correlation between absorption in different quasar sightlines. This measurement probes the growth of structure in the $z = 2 - 4$ range on much larger scales than the line of sight Lyman-α power spectrum. In particular, it probes the scales of the Baryon Acoustic Oscillations (BAO) at $\sim 150 \mathrm{Mpc}\, h^{-1}$. BAO are the signature in the late time matter power spectrum of the sound speed of oscillations in the gas of the early Universe, and take the form of a distinctive peak. This location of this peak tracks the evolution of the Hubble flow, providing a universal standard ruler until the present day. This standard ruler is one of the strongest probes of dark energy available. BAO have already been detected at redshift $z < 1$ using galaxy data from the SDSS galaxy clustering survey (Eisenstein et al., 2011), and making a similar measurement in the Lyman-α forest is an important aim of the BOSS survey (Schlegel et al., 2009). The first steps towards this were recently taken by the measurement of the cross-correlation function by Slosar et al. (2011).

Extracting the information from this new statistic will require new simulations with a very large dynamic range, that can simultaneously capture the BAO scale and at least marginally resolve the Jeans scale in order to model the IGM. In collaboration with M. Viel and M. Haehnelt, I intend to perform such simulations, investigating precisely the signal produced in the cross-correlation function by BAO. It has recently been found that large-scale structure, by probing the peaks of the density distribution, is sensitive to any deviation of the primordial statistics from Gaussianity (Dalal et al., 2008). The Lyman-α forest cross-correlation probes the density distribution excluding peaks, and so analytic

estimates suggest that it should be sensitive to non-Gaussianity in a complementary way (Desjacques and Seljak, 2010). I intend to include non-Gaussianity in simulations of the cross-correlation function, in an effort to determine the effect more precisely, and discover how it is affected by degeneracies with other parameters.

The study of the cross-correlation power spectrum from the Lyman-α forest is presently in its infancy, and many more applications are expected. Together with the new information shed by the Planck satellite on inflation and the many upcoming opportunities to constrain the neutrino mass, the outlook for constraining fundamental physics from cosmology is very promising.

Bibliography

Aarseth, S. J.: 2003, *Gravitational N-Body Simulations*, Cambridge University Press, ADS

Abazajian, K., Switzer, E. R., Dodelson, S., Heitmann, K., and Habib, S.: 2005, *Phys. Rev. D* **71(4)**, 043507, arXiv:astro-ph/0411552, ADS

Abazajian, K. N., Calabrese, E., Cooray, A., De Bernardis, F., Dodelson, S., Friedland, A., and et al.: 2011, arXiv:1103.5083 , ADS

Abel, T. and Haehnelt, M. G.: 1999, *Ap. J. Letters* **520**, L13, arXiv:astro-ph/9903102, ADS

Agarwal, N. and Bean, R.: 2009, *Phys. Rev.* **D79**, 023503, arXiv:0809.2798

Ahmed, S. N., Anthony, A. E., Beier, E. W., Bellerive, A., Biller, S. D., Boger, J., Boulay, M. G., Bowler, M. G., Bowles, T. J., Brice, S. J., Bullard, T. V., and Chan, Y. D.: 2004, *Phys. Rev. Lett.* **92(18)**, 181301, arXiv:nucl-ex/0309004, ADS

Alabidi, L.: 2007, *JCAP* **0702**, 012, arXiv:astro-ph/0608287

Alabidi, L. and Lidsey, J. E.: 2008, *Phys. Rev.* **D78**, 103519, arXiv:0807.2181

Alabidi, L. and Lyth, D. H.: 2006, *JCAP* **0608**, 013, arXiv:astro-ph/0603539

Albrecht, A., Brandenberger, R., and Matzner, R.: 1987, *Phys. Rev. D* **35(2)**, 429

Albrecht, A. and Steinhardt, P. J.: 1982, *Phys. Rev. Lett.* **48(17)**, 1220

Alexander, S. H. S.: 2001, *Phys. Rev.* **D65**, 023507, arXiv:hep-th/0105032

Alpher, R. A., Bethe, H., and Gamow, G.: 1948, *Phys. Rev.* **73**, 803, ADS

Amarie, M., Hirata, C., and Seljak, U.: 2005, *Phys. Rev.* **D72**, 123006, arXiv:astro-ph/0508293

Anderson, G. W. and Castano, D. J.: 1995, *Phys. Lett.* **B347**, 300, arXiv:hep-ph/9409419

Assamagan, K., Brönnimann, C., Daum, M., Forrer, H., Frosch, R., Gheno, P., Horisberger, R., Janousch, M., Kettle, P. R., Spirig, T., and Wigger, C.: 1996, *Phys. Rev.* **D53**, 6065

Barate, R. and ALEPH collaboration: 1998, *Eur. Phys. J.* **C2**, 395

Barbieri, R. and Giudice, G.: 1988, *Nuclear Physics B* **306**, 63

Bardeen, J. M., Steinhardt, P. J., and Turner, M. S.: 1983, *Phys. Rev. D* **28(4)**, 679

Battye, R. A., Garbrecht, B., and Moss, A.: 2006, *JCAP* **9**, 7, arXiv:astro-ph/0607339, ADS

Baumann, D., Dymarsky, A., Kachru, S., Klebanov, I. R., and McAllister, L.: 2009a, *JHEP* **03**, 093, arXiv:0808.2811, ADS

Baumann, D., Dymarsky, A., Kachru, S., Klebanov, I. R., and McAllister, L.: 2010, *Journal of High Energy Physics* **6**, 72, arXiv:1001.5028, ADS

Baumann, D., Dymarsky, A., Klebanov, I. R., Maldacena, J., McAllister, L., and Murugan, A.: 2006, *JHEP* **11**, 31, arXiv:hep-th/0607050, ADS

Baumann, D., Dymarsky, A., Klebanov, I. R., and McAllister, L.: 2008, *JCAP* **0801**, 024, arXiv:0706.0360

Baumann, D., Dymarsky, A., Klebanov, I. R., McAllister, L., and Steinhardt, P. J.: 2007, *Phys. Rev. Lett.* **99**, 141601, arXiv:0705.3837

Baumann, D., Jackson, M. G., Adshead, P., Amblard, A., Ashoorioon, A., Bartolo, N., Bean, R., Beltrán, M., de Bernardis, F., Bird, S., Chen, X., Chung, D. J. H., Colombo, L., Cooray, A., Creminelli, P., Dodelson, S., Dunkley, J., Dvorkin, C., Easther, R., Finelli, F., Flauger, R., Hertzberg, M. P., Jones-Smith, K., Kachru, S., Kadota, K., Khoury, J., Kinney, W. H., Komatsu, E., Krauss, L. M., Lesgourgues, J., Liddle, A., Liguori, M., Lim, E., Linde, A., Matarrese, S., Mathur, H., McAllister, L., Melchiorri, A., Nicolis, A., Pagano, L., Peiris, H. V., Peloso, M., Pogosian, L., Pierpaoli, E., Riotto, A., Seljak, U., Senatore, L., Shandera, S., Silverstein, E., Smith, T., Vaudrevange, P., Verde, L., Wandelt, B., Wands, D., Watson, S., Wyman, M., Yadav, A., Valkenburg, W., and Zaldarriaga, M.: 2009b, in S. Dodelson, D. Baumann, A. Cooray, J. Dunkley, A. Fraisse, M. G. Jackson, A. Kogut, L. Krauss, M. Zaldarriaga, & K. Smith (ed.), *American Institute of Physics Conference Series*, Vol. 1141 of *American Institute of Physics Conference Series*, pp 10–120, arXiv:0811.3919, , ADS

Baumann, D. and McAllister, L.: 2007, *Phys. Rev.* **D75**, 123508, arXiv:hep-th/0610285

Baumann, D. and McAllister, L.: 2009, *Annual Review of Nuclear and Particle Science* **59**, 67, arXiv:0901.0265, ADS

Bean, R., Chen, X., Peiris, H., and Xu, J.: 2008, *Phys. Rev.* **D77**, 023527, arXiv:0710.1812

Bean, R., Shandera, S. E., Henry Tye, S. H., and Xu, J.: 2007, *JCAP* **0705**, 004, arXiv:hep-th/0702107

Beaulieu, J. P., Bennett, D. P., Batista, V., Cassan, A., Kubas, D., Fouqué, P., Kerrins, E., Mao, S., Miralda-Escudé, J., Wambsganss, J., Gaudi, B. S., Gould, A., and Dong, S.: 2010, in V. Coudé Du Foresto, D. M. Gelino, & I. Ribas (ed.), *Pathways Towards*

Habitable Planets, Vol. 430 of *Astronomical Society of the Pacific Conference Series*, pp 266–+, arXiv:1001.3349, , ADS

Becker, G. D., Bolton, J. S., Haehnelt, M. G., and Sargent, W. L. W.: 2011, *MNRAS* **410**, 1096, arXiv:1008.2622, ADS

Becker, M., Leblond, L., and Shandera, S.: 2007, *Phys. Rev. D* **76(12)**, 123516, arXiv:0709.1170, ADS

Becker-Szendy, R., Bratton, C. B., Casper, D., Dye, S. T., Gajewski, W., Goldhaber, M., Haines, T. J., Halverson, P. G., Kielczewska, D., Kropp, W. R., and Learned, J. G.: 1992, *Phys. Rev. D* **46(9)**, 3720, ADS

Belinsky, V. A., Grishchuk, L. P., Khalatnikov, I. M., and Zeldovich, Y. B.: 1985, *Phys. Lett. B* **155**, 232

Belinsky, V. A., Ishihara, H., Khalatnikov, I. M., and Sato, H.: 1988, *Progress of Theoretical Physics* **79**, 676, ADS

Belinsky, V. A. and Khalatnikov, I. M.: 1986, *NASA STI/Recon Technical Report N* **88**, 11586, ADS

Bennett, C. L., Bay, M., Halpern, M., Hinshaw, G., Jackson, C., Jarosik, N., Kogut, A., Limon, M., Meyer, S. S., Page, L., Spergel, D. N., Tucker, G. S., Wilkinson, D. T., Wollack, E., and Wright, E. L.: 2003a, *Ap.J.* **583**, 1, arXiv:astro-ph/0301158, ADS

Bennett, C. L., Halpern, M., Hinshaw, G., Jarosik, N., Kogut, A., Limon, M., Meyer, S. S., Page, L., Spergel, D. N., Tucker, G. S., Wollack, E., Wright, E. L., Barnes, C., Greason, M. R., Hill, R. S., Komatsu, E., Nolta, M. R., Odegard, N., Peiris, H. V., Verde, L., and Weiland, J. L.: 2003b, *Ap. J. Supplements* **148**, 1, arXiv:astro-ph/0302207, ADS

Blake, C., Davis, T., Poole, G., Parkinson, D., Brough, S., Colless, M., Contreras, C., Couch, W., Croom, S., Drinkwater, M. J., Forster, K., Gilbank, D., Gladders, M., Glazebrook, K., Jelliffe, B., Jurek, R. J., Li, I., Madore, B., Martin, C., Pimbblet, K., Pracy, M., Sharp, R., Wisnioski, E., Woods, D., Wyder, T., and Yee, H.: 2011, *1105.2862*, arXiv:1105.2862, ADS

Bolton, J. S. and Becker, G. D.: 2009, *MNRAS* **398**, L26, arXiv:0906.2861

Bolton, J. S., Viel, M., Kim, T., Haehnelt, M. G., and Carswell, R. F.: 2008, *MNRAS* **386**, 1131, arXiv:0711.2064, ADS

Bond, J. R., Crittenden, R., Davis, R. L., Efstathiou, G., and Steinhardt, P. J.: 1994, *Phys. Rev. Lett.* **72**, 13, arXiv:astro-ph/9309041, ADS

Bond, J. R. and Efstathiou, G.: 1984, *Ap. J. Letters* **285**, L45, ADS

Bond, J. R., Efstathiou, G., and Silk, J.: 1980, *Phys. Rev. Lett.* **45(24)**, 1980, ADS

Boyle, L. A., Steinhardt, P. J., and Turok, N.: 2004, *Phys. Rev.* **D69**, 127302, arXiv:hep-th/0307170

Boyle, L. A., Steinhardt, P. J., and Turok, N.: 2006, *Phys. Rev. Lett.* **96**, 111301, arXiv:astro-ph/0507455

Brandbyge, J. and Hannestad, S.: 2009, *JCAP* **5**, 2, arXiv:0812.3149, ADS

Brandbyge, J. and Hannestad, S.: 2010, *JCAP* **1**, 21, arXiv:0908.1969, ADS

Brandbyge, J., Hannestad, S., Haugbølle, T., and Thomsen, B.: 2008, *JCAP* **8**, 20, arXiv:0802.3700, ADS

Brustein, R. and Steinhardt, P. J.: 1993, *Phys. Lett.* **B302**, 196, arXiv:hep-th/9212049

Burgess, C. P.: 2007, *Class. Quant. Grav.* **24**, S795, arXiv:0708.2865, ADS

Burgess, C. P., Cline, J. M., Dasgupta, K., and Firouzjahi, H.: 2007, *JHEP* **03**, 027, arXiv:hep-th/0610320

Burgess, C. P., Cline, J. M., Stoica, H., and Quevedo, F.: 2004, *JHEP* **09**, 033, arXiv:hep-th/0403119

Burgess, C. P., Majumdar, M., Nolte, D., Quevedo, F., Rajesh, G., and Zhang, R.-J.: 2001, *JHEP* **7**, 47, arXiv:hep-th/0105204, ADS

Burgess, C. P., Martineau, P., Quevedo, F., Rajesh, G., and Zhang, R.-J.: 2002, *JHEP* **3**, 52, arXiv:hep-th/0111025, ADS

Carmack, P. S., Schucany, W. R., Spence, J. S., Gunst, R. F., Lin, Q., and Haley, R. W.: 2009, *Journal of Computational and Graphical Statistics* **18(4)**, 879

Chen, H. and Gong, J.: 2009, *Phys. Rev. D* **80(6)**, 063507, arXiv:0812.4649, ADS

Chiba, T.: 2008, *JCAP* **8**, 4, arXiv:0805.4660, ADS

Cline, J. M.: 2006, arXiv:hep-th/0612129 , ADS

Cline, J. M., Hoi, L., and Underwood, B.: 2009, *JHEP* **06**, 078, arXiv:0902.0339

Croft, R. A. C., Weinberg, D. H., Katz, N., and Hernquist, L.: 1998, *Ap.J.* **495**, 44, arXiv:astro-ph/9708018

Dalal, N., Doré, O., Huterer, D., and Shirokov, A.: 2008, *Phys. Rev. D* **77(12)**, 123514, arXiv:0710.4560, ADS

DePies, M. R. and Hogan, C. J.: 2007, *Phys. Rev. D* **75(12)**, 125006, arXiv:astro-ph/0702335, ADS

Desjacques, V. and Seljak, U.: 2010, *Advances in Astronomy* 2010, arXiv:1006.4763, ADS

Destri, C., de Vega, H. J., and Sanchez, N. G.: 2008, *Phys. Rev. D* **77(4)**, 043509, arXiv:astro-ph/0703417, ADS

DeWolfe, O., Kachru, S., and Verlinde, H. L.: 2004, *JHEP* **05**, 017, arXiv:hep-th/0403123

Dimopoulos, S., Kachru, S., McGreevy, J., and Wacker, J. G.: 2008, *JCAP* **8**, 3, arXiv:hep-th/0507205, ADS

Douglas, M. R. and Kachru, S.: 2007, *Rev. Mod. Phys.* **79**, 733, arXiv:hep-th/0610102

Dunkley, J., Komatsu, E., Nolta, M. R., Spergel, D. N., Larson, D., Hinshaw, G., Page, L., Bennett, C. L., Gold, B., Jarosik, N., Weiland, J. L., Halpern, M., Hill, R. S., Kogut, A., Limon, M., Meyer, S. S., Tucker, G. S., Wollack, E., and Wright, E. L.: 2009, *Ap. J. Supplements* **180**, 306, arXiv:0803.0586, ADS

Dvali, G., Shafi, Q., and Solganik, S.: 2001, arXiv:hep-th/0105203 , ADS

Dvali, G. R. and Tye, S. H. H.: 1999, *Phys. Lett.* **B450**, 72, arXiv:hep-ph/9812483

Easther, R., Kinney, W. H., and Powell, B. A.: 2006, *JCAP* **0608**, 004, arXiv:astro-ph/0601276

Easther, R. and McAllister, L.: 2006, *JCAP* **0605**, 018, arXiv:hep-th/0512102

Efstathiou, G. and Chongchitnan, S.: 2006, *Progress of Theoretical Physics Supplement* **163**, 204, arXiv:astro-ph/0603118, ADS

Efstathiou, G. and Eastwood, J. W.: 1981, *MNRAS* **194**, 503, ADS

Efstathiou, G., Sutherland, W. J., and Maddox, S. J.: 1990, *Nature* **348**, 705, ADS

Eisenstein, D. J. and Hu, W.: 1999, *Ap.J.* **511**, 5, arXiv:astro-ph/9710252, ADS

Eisenstein, D. J., Weinberg, D. H., Agol, E., Aihara, H., Allende Prieto, C., Anderson, S. F., Arns, J. A., Aubourg, E., Bailey, S., Balbinot, E., and et al.: 2011, arXiv:1101.1529 , ADS

Elgaroy, O. et al.: 2002, *Phys. Rev. Lett.* **89**, 061301, arXiv:astro-ph/0204152, ADS

Felder, G. N., Frolov, A. V., and Kofman, L.: 2002, *Class. Quant. Grav.* **19**, 2983, arXiv:hep-th/0112165

Firouzjahi, H. and Tye, S.-H. H.: 2005, *JCAP* **3**, 9, arXiv:hep-th/0501099, ADS

Fraisse, A. A., Ringeval, C., Spergel, D. N., and Bouchet, F. R.: 2008, *Phys. Rev. D* **78(4)**, 043535, arXiv:0708.1162, ADS

Freese, K., Frieman, J. A., and Olinto, A. V.: 1990, *Phys. Rev. Lett.* **65**, 3233, ADS

Freivogel, B., Kleban, M., Rodriguez Martinez, M., and Susskind, L.: 2006, *JHEP* **03**, 039, arXiv:hep-th/0505232

Frey, A. R., Mazumdar, A., and Myers, R. C.: 2006, *Phys. Rev. D* **73(2)**, 026003, arXiv:hep-th/0508139, ADS

Friedman, A.: 1922, *Zeitschrift fur Physik* **10**, 377, ADS

Fukuda, Y., Hayakawa, T., Ichihara, E., Inoue, K., Ishihara, K., Ishino, H., Itow, Y., Kajita, T., Kameda, J., Kasuga, S., Kobayashi, K., and Kobayashi, Y.: 1998, *Phys. Rev. Lett.* **81**, 1158, arXiv:hep-ex/9805021, ADS

Garcia-Bellido, J., Rabadan, R., and Zamora, F.: 2002, *JHEP* **0201**, 036, arXiv:hep-th/0112147

Gershtein, S. S. and Zel'dovich, Y. B.: 1966, *Soviet Journal of Experimental and Theoretical Physics Letters* **4**, 120, ADS

Gibbons, G., Hawking, S., and Stewart, J.: 1987, *Nuclear Physics B* **281**, 736

Gibbons, G. W. and Turok, N.: 2008, *Phys. Rev.* **D77**, 063516, arXiv:hep-th/0609095

Gmeiner, F. and White, C. D.: 2008, *JCAP* **0802**, 012, arXiv:0710.2009

Gnedin, N. Y. and Hamilton, A. J. S.: 2002, *MNRAS* **334**, 107, arXiv:astro-ph/0111194, ADS

Goldwirth, D. S.: 1990, *Phys. Lett. B* **243**, 41

Goldwirth, D. S.: 1991, *Phys. Rev. D* **43(10)**, 3204

Goldwirth, D. S. and Piran, T.: 1990, *Phys. Rev. Lett.* **64(24)**, 2852

Goldwirth, D. S. and Piran, T.: 1992, *Phys. Rept.* **214**, 223

Gratton, S., Lewis, A., and Efstathiou, G.: 2008, *Phys. Rev.* **D77**, 083507, arXiv:0705.3100

Green, M., Schwarz, J., and Witten, E.: 1988, *Superstring theory: Loop amplitudes, anomalies and phenomenology*, Cambridge monographs on mathematical physics, Cambridge University Press

Green, P. J. and Silverman, B. W.: 1994, *Non-parametric Regression and Generalized Linear Models*, Chapman and Hall (London)

Grishchuk, L.: 1975, *Sov. Phys. JETP* **40**, 409

Gubser, S. S., Klebanov, I. R., and Polyakov, A. M.: 1998, *Phys. Lett.* **B428**, 105, arXiv:hep-th/9802109

Gunn, J. E. and Peterson, B. A.: 1965, *Ap.J.* **142**, 1633, ADS

Guth, A. H.: 1981, *Phys. Rev.* **D23**, 347

Guth, A. H. and Pi, S.-Y.: 1982, *Phys. Rev. Lett.* **49(15)**, 1110

Haardt, F. and Madau, P.: 2001, in D. M. Neumann & J. T. V. Tran (ed.), *Clusters of Galaxies and the High Redshift Universe Observed in X-rays. CEA, Saclay*, arXiv:astro-ph/0106018, , ADS

Hamilton, A. J. S., Kumar, P., Lu, E., and Matthews, A.: 1991, *Ap. J. Letters* **374**, L1, ADS

Hawking, S. W.: 1982, *Phys. Lett. B* **115**, 295

Hernquist, L., Katz, N., Weinberg, D. H., and Miralda-Escudé, J.: 1996, *Ap. J. Letters* **457**, L51+, arXiv:astro-ph/9509105, ADS

Hertzberg, M. P., Tegmark, M., Kachru, S., Shelton, J., and Ozcan, O.: 2007, *Phys. Rev.* **D76**, 103521, arXiv:0709.0002

Hoffman, M. B. and Turner, M. S.: 2001, *Phys. Rev.* **D64**, 023506, arXiv:astro-ph/0006321

Hoi, L. and Cline, J. M.: 2009, *Phys. Rev. D* **79(8)**, 083537, arXiv:0810.1303, ADS

Hollands, S. and Wald, R. M.: 2002, arXiv:hep-th/0210001 , ADS

Holmberg, E.: 1941, *Ap.J.* **94**, 385, ADS

Hotchkiss, S., Germán, G., Ross, G. G., and Sarkar, S.: 2008, *JCAP* **10**, 15, arXiv:0804.2634, ADS

Hubble, E.: 1929, *Proceedings of the National Academy of Science* **15**, 168, ADS

Hui, L., Burles, S., Seljak, U., Rutledge, R. E., Magnier, E., and Tytler, D.: 2001, *Ap.J.* **552**, 15, arXiv:astro-ph/0005049, ADS

Iizuka, N. and Trivedi, S. P.: 2004, *Phys. Rev.* **D70**, 043519, arXiv:hep-th/0403203

Itzhaki, N.: 2008, *JHEP* **10**, 061, arXiv:0807.3216

Itzhaki, N. and Kovetz, E. D.: 2007, *JHEP* **10**, 054, arXiv:0708.2798

Jeffreys, H.: 1961, *Theory of Probability*, Oxford University Press

Jones, N., Stoica, H., and Tye, S.-H. H.: 2002, *JHEP* **7**, 51, arXiv:hep-th/0203163, ADS

Kachru, S., Kallosh, R., Linde, A., Maldacena, J., McAllister, L., and Trivedi, S. P.: 2003, *JCAP* **10**, 13, arXiv:hep-th/0308055, ADS

Kallosh, R.: 2008, in M. Lemoine, J. Martin, and P. Peter (eds.), *Lecture Notes in Physics, Berlin Springer Verlag*, Vol. 738 of *Inflationary Cosmology, Lecture Notes in Physics, Berlin Springer Verlag*, pp 119–+, ADS

Kamionkowski, M., Kosowsky, A., and Stebbins, A.: 1997, *Phys. Rev. D* **55**, 7368, arXiv:astro-ph/9611125, ADS

Katz, N., Weinberg, D. H., and Hernquist, L.: 1996, *Ap. J. Supplements* **105**, 19, arXiv:astro-ph/9509107

Kesden, M., Cooray, A., and Kamionkowski, M.: 2002, *Phys. Rev. Lett.* **89(1)**, 011304, arXiv:astro-ph/0202434, ADS

Khoury, J., Ovrut, B. A., Seiberg, N., Steinhardt, P. J., and Turok, N.: 2002, *Phys. Rev.* **D65**, 086007, arXiv:hep-th/0108187

Kim, T., Bolton, J. S., Viel, M., Haehnelt, M. G., and Carswell, R. F.: 2007, *MNRAS* **382**, 1657, arXiv:0711.1862, ADS

Kinney, W. H.: 2002, *Phys. Rev.* **D66**, 083508, arXiv:astro-ph/0206032

Kinney, W. H., Kolb, E. W., Melchiorri, A., and Riotto, A.: 2006, *Phys. Rev.* **D74**, 023502, arXiv:astro-ph/0605338

Kinney, W. H., Kolb, E. W., Melchiorri, A., and Riotto, A.: 2008, *Phys. Rev.* **D78**, 087302, arXiv:0805.2966

Klebanov, I. R. and Strassler, M. J.: 2000, *JHEP* **0008**, 052, arXiv:hep-th/0007191

Klypin, A., Holtzman, J., Primack, J., and Regos, E.: 1993, *Ap.J.* **416**, 1, arXiv:astro-ph/9305011, ADS

Knox, L. and Olinto, A.: 1993, *Phys. Rev. D* **48**, 946, ADS

Knox, L. and Song, Y.-S.: 2002, *Phys. Rev. Lett.* **89**, 011303, arXiv:astro-ph/0202286

Kobayashi, T., Mukohyama, S., and Kinoshita, S.: 2008, *JCAP* **1**, 28, arXiv:0708.4285, ADS

Kofman, L., Linde, A., and Mukhanov, V.: 2002, *JHEP* **10**, 57, arXiv:hep-th/0206088, ADS

Kofman, L. and Yi, P.: 2005, *Phys. Rev.* **D72**, 106001, arXiv:hep-th/0507257

Kolb, E. W. and Turner, M. S.: 1990, *The early universe.*, Westview Press, ADS

Komatsu, E., Dunkley, J., Nolta, M. R., Bennett, C. L., Gold, B., Hinshaw, G., Jarosik, N., Larson, D., Limon, M., Page, L., Spergel, D. N., Halpern, M., Hill, R. S., Kogut, A., Meyer, S. S., Tucker, G. S., Weiland, J. L., Wollack, E., and Wright, E. L.: 2009, *Ap. J. Supplements* **180**, 330, arXiv:0803.0547, ADS

Komatsu, E., Smith, K. M., Dunkley, J., Bennett, C. L., Gold, B., Hinshaw, G., Jarosik, N., Larson, D., and et al.: 2011, *Ap. J. Supplements* **192**, 18, arXiv:1001.4538, ADS

Krause, A. and Pajer, E.: 2008, *JCAP* **7**, 23, arXiv:0705.4682, ADS

Kung, J. H. and Brandenberger, R.: 1989, *Phys. Rev. D* **40(8)**, 2532

Kung, J. H. and Brandenberger, R. H.: 1990, *Phys. Rev. D* **42(4)**, 1008

Lemaître, A. G.: 1931, *Nature* **128**, 704, ADS

Lemaître, G.: 1927, *Annales de la Societe Scientifique de Bruxelles* **47**, 49, ADS

Lesgourgues, J., Matarrese, S., Pietroni, M., and Riotto, A.: 2009, *JCAP* **6**, 17, arXiv:0901.4550, ADS

Lesgourgues, J. and Pastor, S.: 2006, *Physics Reports* **429**, 307, arXiv:astro-ph/0603494, ADS

Lesgourgues, J. and Valkenburg, W.: 2007, *Phys. Rev.* **D75**, 123519, arXiv:astro-ph/0703625

Lewis, A. and Bridle, S.: 2002, *Phys. Rev. D* **66(10)**, 103511, arXiv:astro-ph/0205436, ADS

Lewis, A., Challinor, A., and Lasenby, A.: 2000, *Ap.J.* **538**, 473, arXiv:astro-ph/9911177, ADS

Lidz, A., Heitmann, K., Hui, L., Habib, S., Rauch, M., and Sargent, W. L. W.: 2006, *Ap.J.* **638**, 27, arXiv:astro-ph/0505138, ADS

Linde, A.: 1982a, *Physics Letters B* **116**, 335

Linde, A.: 2006, *Progress of Theoretical Physics Supplement* **163**, 295, arXiv:hep-th/0503195, ADS

Linde, A. and Westphal, A.: 2008, *JCAP* **0803**, 005, arXiv:0712.1610

Linde, A. D.: 1982b, *Phys. Lett. B* **108**, 389

Linde, A. D.: 1983, *Phys. Lett. B* **129**, 177

Linde, A. D.: 1994, *Phys. Rev.* **D49**, 748, arXiv:astro-ph/9307002

Lo, A. S. and Wright, E. L.: 2005, arXiv:astro-ph/0503120 , ADS

Lobashev, V. M.: 2003, *Nuclear Physics A* **719**, 153, ADS

Lorenz, L., Martin, J., and Ringeval, C.: 2008, *JCAP* **0804**, 001, arXiv:0709.3758

Lucchin, F. and Matarrese, S.: 1985, *Phys. Rev. D* **32**, 1316, ADS

Lyth, D. H.: 1997, *Phys. Rev. Lett.* **78**, 1861, arXiv:hep-ph/9606387

Ma, C.-P. and Bertschinger, E.: 1994, *Ap.J.* **429**, 22, arXiv:astro-ph/9308006, ADS

Mack, K. J., Wesley, D. H., and King, L. J.: 2007, *Phys. Rev. D* **76(12)**, 123515, arXiv:astro-ph/0702648, ADS

Maldacena, J. M.: 1998, *Adv. Theor. Math. Phys.* **2**, 231, arXiv:hep-th/9711200

Mantz, A., Allen, S. W., and Rapetti, D.: 2010, *MNRAS* **406**, 1805, arXiv:0911.1788, ADS

Markevitch, M., Gonzalez, A. H., Clowe, D., Vikhlinin, A., Forman, W., Jones, C., Murray, S., and Tucker, W.: 2004, *Ap.J.* **606**, 819, arXiv:astro-ph/0309303, ADS

Martin, J. and Ringeval, C.: 2006, *JCAP* **0608**, 009, arXiv:astro-ph/0605367

Mather, J. C., Cheng, E. S., Eplee, Jr., R. E., Isaacman, R. B., Meyer, S. S., Shafer, R. A., Weiss, R., Wright, E. L., Bennett, C. L., Boggess, N. W., Dwek, E., Gulkis, S., Hauser, M. G., Janssen, M., Kelsall, T., Lubin, P. M., Moseley, Jr., S. H., Murdock, T. L., Silverberg, R. F., Smoot, G. F., and Wilkinson, D. T.: 1990, *Ap. J. Letters* **354**, L37, ADS

McAllister, L. and Silverstein, E.: 2008, *General Relativity and Gravitation* **40**, 565, arXiv:0710.2951, ADS

McDonald, P.: 2003, *Ap.J.* **585**, 34, arXiv:astro-ph/0108064, ADS

McDonald, P. and Eisenstein, D. J.: 2007, *Phys. Rev. D* **76(6)**, 063009, arXiv:astro-ph/0607122, ADS

McDonald, P. et al.: 2005a, *Ap.J.* **635**, 761, arXiv:astro-ph/0407377

McDonald, P. et al.: 2006, *Ap. J. Supplements* **163**, 80, arXiv:astro-ph/0405013

McDonald, P., Miralda-Escudé, J., Rauch, M., Sargent, W. L. W., Barlow, T. A., Cen, R., and Ostriker, J. P.: 2000, *Ap.J.* **543**, 1, arXiv:astro-ph/9911196, ADS

McDonald, P., Seljak, U., Cen, R., Bode, P., and Ostriker, J. P.: 2005b, *MNRAS* **360**, 1471, arXiv:astro-ph/0407378

McQuinn, M., Hernquist, L., Lidz, A., and Zaldarriaga, M.: 2011, *MNRAS* pp 874–+, arXiv:1010.5250, ADS

Meiksin, A. A.: 2009, *Reviews of Modern Physics* **81**, 1405, arXiv:0711.3358, ADS

Melchiorri, A., Ade, P. A. R., de Bernardis, P., Bock, J. J., Borrill, J., Boscaleri, A., Crill, B. P., De Troia, G., Farese, P., Ferreira, P. G., Ganga, K., de Gasperis, G., Giacometti, M., Hristov, V. V., Jaffe, A. H., Lange, A. E., Masi, S., Mauskopf, P. D., Miglio, L., Netterfield, C. B., Pascale, E., Piacentini, F., Romeo, G., Ruhl, J. E., and Vittorio, N.: 2000, *Ap. J. Letters* **536**, L63, arXiv:astro-ph/9911445, ADS

Mendes, L. E. and Liddle, A. R.: 2000, *Phys. Rev. D* **62(10)**, 103511, arXiv:astro-ph/0006020, ADS

Moss, I. and Sahni, V.: 1986, *Physics Letters B* **178**, 159, ADS

Mukhanov, V. and Chibisov, G.: 1981, *ZhETF Pis'ma Redaktsiiu* **33**, 549

Mukhanov, V. F., Feldman, H. A., and Brandenberger, R. H.: 1992, *Physics Reports* **215**, 203, ADS

Niemack, M. D., Ade, P. A. R., Aguirre, J., Barrientos, F., Beall, J. A., Bond, J. R., Britton, J., Cho, H. M., Das, S., Devlin, M. J., Dicker, S., Dunkley, J., Dünner, R., Fowler, J. W., Hajian, A., Halpern, M., Hasselfield, M., Hilton, G. C., Hilton, M., Hubmayr, J., Hughes, J. P., Infante, L., Irwin, K. D., Jarosik, N., Klein, J., Kosowsky, A., Marriage, T. A., McMahon, J., Menanteau, F., Moodley, K., Nibarger, J. P., Nolta, M. R., Page, L. A., Partridge, B., Reese, E. D., Sievers, J., Spergel, D. N., Staggs, S. T., Thornton, R., Tucker, C., Wollack, E., and Yoon, K. W.: 2010, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Vol. 7741 of *Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference*, arXiv:1006.5049, , ADS

O'Shea, B. W., Bryan, G., Bordner, J., Norman, M. L., Abel, T., Harkness, R., and Kritsuk, A.: 2004, arXiv:astro-ph/0403044 , ADS

Pajer, E.: 2008, *JCAP* **0804**, 031, arXiv:0802.2916

Peacock, J. A. and Dodds, S. J.: 1996, *MNRAS* **280**, L19, arXiv:astro-ph/9603031, ADS

Peebles, P. J. E.: 1974, *Ap. J. Letters* **189**, L51+, ADS

Peebles, P. J. E.: 1980, *The large-scale structure of the universe*, ADS

Peiris, H., Baumann, D., Friedman, B., and Cooray, A.: 2007, *Phys. Rev. D* **76(10)**, 103517, arXiv:0706.1240, ADS

Peiris, H. and Easther, R.: 2006a, *JCAP* **0607**, 002, arXiv:astro-ph/0603587

Peiris, H. and Easther, R.: 2006b, *JCAP* **0610**, 017, arXiv:astro-ph/0609003

Peiris, H. V. and Easther, R.: 2008, *JCAP* **0807**, 024, arXiv:0805.2154

Peiris, H. V. et al.: 2003, *Ap. J. Supplements* **148**, 213, arXiv:astro-ph/0302225

Peiris, H. V. and Verde, L.: 2010, *Phys. Rev.* **D81**, 021302, arXiv:0912.0268

Penzias, A. A. and Wilson, R. W.: 1965, *Ap.J.* **142**, 419, ADS

Percival, W. J., Sutherland, W., Peacock, J. A., Baugh, C. M., Bland-Hawthorn, J., Bridges, T., Cannon, R., Cole, S., Colless, M., Collins, C., Couch, W., Dalton, G., De Propris, R., Driver, S. P., Efstathiou, G., Ellis, R. S., Frenk, C. S., Glazebrook, K., Jackson, C., Lahav, O., Lewis, I., Lumsden, S., Maddox, S., Moody, S., Norberg, P., Peterson, B. A., and Taylor, K.: 2002, *MNRAS* **337**, 1068, arXiv:astro-ph/0206256, ADS

Piran, T.: 1986, *Physics Letters B* **181**, 238, ADS

Piran, T. and Williams, R. M.: 1985, *Physics Letters B* **163**, 331, ADS

Planck Collaboration: 2005, *ESA-SCI*, arXiv:astro-ph/0604069, ADS

Rauch, M.: 1998, *ARA&A* **36**, 267, arXiv:astro-ph/9806286, ADS

Regan, J. A., Haehnelt, M. G., and Viel, M.: 2007, *MNRAS* **374**, 196, arXiv:astro-ph/0606638, ADS

Reid, B. A., Verde, L., Jimenez, R., and Mena, O.: 2010, *JCAP* **1**, 3, arXiv:0910.0008, ADS

Riess, A. G., Filippenko, A. V., Challis, P., Clocchiatti, A., Diercks, A., Garnavich, P. M., Gilliland, R. L., Hogan, C. J., Jha, S., Kirshner, R. P., Leibundgut, B., Phillips, M. M., Reiss, D., Schmidt, B. P., Schommer, R. A., Smith, R. C., Spyromilio, J., Stubbs, C., Suntzeff, N. B., and Tonry, J.: 1998, *AJ* **116**, 1009, arXiv:astro-ph/9805201, ADS

Rubin, V. C. and Ford, Jr., W. K.: 1970, *Ap.J.* **159**, 379, ADS

Saito, S., Takada, M., and Taruya, A.: 2008, *Physical Review Letters* **100(19)**, 191301, arXiv:0801.0607, ADS

Sato, K.: 1981, *MNRAS* **195**, 467

Schaye, J., Theuns, T., Rauch, M., Efstathiou, G., and Sargent, W. L. W.: 2000, *MNRAS* **318**, 817, arXiv:astro-ph/9912432

Schlegel, D., White, M., and Eisenstein, D.: 2009, in *Astro2010, A&A Decadal Survey.*, p. 314, arXiv:0902.4680, , ADS

Scoccimarro, R.: 1998, *MNRAS* **299**, 1097, arXiv:astro-ph/9711187, ADS

Sealfon, C., Verde, L., and Jimenez, R.: 2005, *Phys. Rev.* **D72**, 103520, arXiv:astro-ph/0506707

Seljak, U., Makarov, A., McDonald, P., Anderson, S. F., Bahcall, N. A., et al.: 2005, *Phys. Rev. D* **71(10)**, 103515, arXiv:astro-ph/0407372, ADS

Seljak, U. and Slosar, A.: 2006, *Phys. Rev. D* **74(6)**, 063523, arXiv:astro-ph/0604143, ADS

Seljak, U., Slosar, A., and McDonald, P.: 2006, *JCAP* **10**, 14, arXiv:astro-ph/0604335, ADS

Seljak, U. and Zaldarriaga, M.: 1996, *Ap.J.* **469**, 437, arXiv:astro-ph/9603033, ADS

Shandera, S. E. and Tye, S. H. H.: 2006, *JCAP* **0605**, 007, arXiv:hep-th/0601099

Shiu, G. and Tye, S. H. H.: 2001, *Phys. Lett.* **B516**, 421, arXiv:hep-th/0106274

Silverstein, E. and Westphal, A.: 2008, *Phys. Rev. D* **78(10)**, 106003, arXiv:0803.3085, ADS

Simha, V. and Steigman, G.: 2008, *JCAP* **6**, 16, arXiv:0803.3465, ADS

Slosar, A., Font-Ribera, A., Pieri, M. M., Rich, J., Le Goff, J.-M., Aubourg, É., Brinkmann, J., Busca, N., Carithers, B., Charlassier, R., Cortês, M., Croft, R., Dawson, K. S., Eisenstein, D., Hamilton, J.-C., Ho, S., Lee, K.-G., Lupton, R., McDonald, P., Medolin, B., Miralda-Escudé, J., Myers, A. D., Nichol, R. C., Palanque-Delabrouille, N., Pâris, I., Petitjean, P., Piškur, Y., Rollinde, E., Ross, N. P., Schlegel, D. J., Schneider, D. P., Sheldon, E., Weaver, B. A., Weinberg, D. H., Yeche, C., and York, D. G.: 2011, arXiv:1104.5244 , ADS

Slosar, A., Ho, S., White, M., and Louis, T.: 2009, *Journal of Cosmology and Astro-Particle Physics* **10**, 19, arXiv:0906.2414, ADS

Smith, R. E., Peacock, J. A., Jenkins, A., White, S. D. M., Frenk, C. S., Pearce, F. R., Thomas, P. A., Efstathiou, G., and Couchman, H. M. P.: 2003, *MNRAS* **341**, 1311, arXiv:astro-ph/0207664, ADS

Smith, T. L., Kamionkowski, M., and Cooray, A.: 2008, *Phys. Rev. D* **78(8)**, 083525, arXiv:0802.1530, ADS

Smoot, G. F., Bennett, C. L., Kogut, A., Wright, E. L., Aymon, J., Boggess, N. W., Cheng, E. S., de Amici, G., Gulkis, S., Hauser, M. G., Hinshaw, G., Jackson, P. D., Janssen, M., Kaita, E., Kelsall, T., Keegstra, P., Lineweaver, C., Loewenstein, K., Lubin, P., Mather, J., Meyer, S. S., Moseley, S. H., Murdock, T., Rokke, L., Silverberg, R. F., Tenorio, L., Weiss, R., and Wilkinson, D. T.: 1992, *Ap. J. Letters* **396**, L1, ADS

Spaliński, M.: 2009, *Phys. Rev. D* **80(6)**, 063529, arXiv:0903.4999, ADS

Spergel, D. N., Verde, L., Peiris, H. V., Komatsu, E., Nolta, M. R., Bennett, C. L., Halpern, M., Hinshaw, G., Jarosik, N., Kogut, A., Limon, M., Meyer, S. S., Page, L., Tucker, G. S., Weiland, J. L., Wollack, E., and Wright, E. L.: 2003, *Ap. J. Supplements* **148**, 175, arXiv:astro-ph/0302209, ADS

Springel, V.: 2005, *MNRAS* **364**, 1105, arXiv:astro-ph/0505010

Springel, V.: 2010, *MNRAS* **401**, 791, arXiv:0901.4107, ADS

Springel, V. and Hernquist, L.: 2002, *MNRAS* **333**, 649, arXiv:astro-ph/0111016, ADS

Starobinsky, A. A.: 1979, *Pis'ma Zh. Eksp. Teor. Fiz.* **30**, 719

Starobinsky, A. A.: 1982, *Phys. Lett. B* **117**, 175

't Hooft, G.: 1980, *NATO Adv. Study Inst. Ser. B Phys.* **59**, 135

Tegmark, M. et al.: 2006, *Phys. Rev.* **D74**, 123507, arXiv:astro-ph/0608632, ADS

Teyssier, R.: 2002, *A&A* **385**, 337, arXiv:astro-ph/0111367, ADS

Theuns, T., Leonard, A., Efstathiou, G., Pearce, F. R., and Thomas, P. A.: 1998, *MNRAS* **301**, 478, arXiv:astro-ph/9805119, ADS

Thomas, S. A., Abdalla, F. B., and Lahav, O.: 2010, *Phys. Rev. Lett.* **105(3)**, 031301, arXiv:0911.5291, ADS

Tye, S.-H. H.: 2008, in M. Gasperini and J. Maharana (eds.), *String Theory and Fundamental Interactions, Lecture Notes in Physics, Berlin Springer Verlag*, Vol. 737 of *Lecture Notes in Physics, Berlin Springer Verlag*, pp 949–+, ADS

Tzirakis, K. and Kinney, W. H.: 2009, *JCAP* **0901**, 028, arXiv:0810.0270

Underwood, B.: 2008, *Phys. Rev.* **D78**, 023509, arXiv:0802.2117

Valkenburg, W., Krauss, L. M., and Hamann, J.: 2008, *Phys. Rev. D* **78(6)**, 063521, arXiv:0804.3390, ADS

van Daalen, M. P., Schaye, J., Booth, C. M., and Dalla Vecchia, C.: 2011, arXiv:1104.1174 , ADS

Verde, L., Peiris, H., and Jimenez, R.: 2006, *JCAP* **0601**, 019, arXiv:astro-ph/0506036

Verde, L. and Peiris, H. V.: 2008, *JCAP* **0807**, 009, arXiv:0802.1219

Viel, M., Bolton, J. S., and Haehnelt, M. G.: 2009, *MNRAS* **399**, L39, arXiv:0907.2927, ADS

Viel, M. and Haehnelt, M. G.: 2006, *MNRAS* **365**, 231, arXiv:astro-ph/0508177

Viel, M., Haehnelt, M. G., and Springel, V.: 2004, *MNRAS* **354**, 684, arXiv:astro-ph/0404600, ADS

Viel, M., Haehnelt, M. G., and Springel, V.: 2010, *JCAP* **6**, 15, arXiv:1003.2422, ADS

Viel, M., Matarrese, S., Mo, H. J., Haehnelt, M. G., and Theuns, T.: 2002, *MNRAS* **329**, 848, arXiv:astro-ph/0105233, ADS

Wang, J. and White, S. D. M.: 2007, *MNRAS* **380**, 93, arXiv:astro-ph/0702575, ADS

Wang, S., Haiman, Z., Hu, W., Khoury, J., and May, M.: 2005, *Phys. Rev. Lett.* **95**, 011302, arXiv:astro-ph/0505390

White, M., Pope, A., Carlson, J., Heitmann, K., Habib, S., Fasel, P., Daniel, D., and Lukic, Z.: 2010, *Ap.J.* **713**, 383, arXiv:0911.5341, ADS

Witten, E.: 1998, *Adv. Theor. Math. Phys.* **2**, 253, arXiv:hep-th/9802150

Wolf, J. and KATRIN collaboration: 2010, *Nuclear Instruments and Methods in Physics Research A* **623**, 442, arXiv:0810.3281, ADS

Zel'dovich, Y. B.: 1970, *A&A* **5**, 84, ADS

Zwicky, F.: 1933, *Helvetica Physica Acta* **6**, 110, ADS