COSMOLOGICAL PERTURBATIONS FROM COSMIC STRINGS

JIUN-HUEI PROTY WU
Trinity Hall
University of Cambridge

Submitted in partial fulfilment of the requirements for the degree of
Doctor of Philosophy

September 1999

To my parents

Hsin-Ying Wu & Shu-Meei Chang

Declaration

I hereby declare that the work contained in this dissertation, except where explicitly stated, is my own and is to the best of my knowledge original. It is not substantially the same as any that I have submitted for a degree or diploma or other qualification at any other university. I further state that no part of it has already been or is being currently submitted for any such degree, diploma, or other qualification.

Parts of the work were done in collaboration with E. P. S. Shellard (chapters 5, 6, 7, 8 and 9), P. P. Avelino (chapters 6, 7 and 8), B. Allen (chapters 5, 6, 7 and 9), and S. Winitzki (chapter 10).

Most parts of this thesis are included in papers already published [1, 2, 3], submitted [4], or in preparation [5, 6, 7, 8, 9, 10]. Some of these papers were written with E. P. S. Shellard [1, 2, 3, 4, 5, 7, 10], P. P. Avelino [1, 2, 3, 4, 5, 10], B. Allen [1, 2, 3, 5, 7, 10], and S. Winitzki [9].

Signed: Date:

....
Jiun-Huei Proty Wu

Acknowledgements

It is a pleasure to thank my supervisor Dr. Paul Shellard for those uncountably penetrating discussions which have helped to clarify the arguments presented here, for revising the English in this dissertation, and for his invaluable advice and encouragement which often saved me from pursuing dead ends or giving up on fruitful problems.

I would also like to express my warm thanks to my parents for providing me with much of the financial support for my PhD, and to my beautiful sisters, Yea-Chen and Ya-Ling, for their encouragement. Special thanks are due to my lovely wife Mrs. Yi-Tsun Lin, for taking care of my day-to-day life throughout the period of my stay in Cambridge, and even for typing up many tedious equations and revising the dissertation. I am also in debt to my colleague and good friend, Dr. Pedro Avelino, who provided his precious advice from time to time.

I acknowledge financial support from the Cambridge Overseas Trust, the University of Cambridge, and from CVCP (UK) under the 'ORS scheme' (ORS/96009158). A large part of this work was performed on COSMOS, the Origin2000 owned by the UK Computational Cosmology Consortium, supported by SGI, HEFCE and PPARC. Additional travel grants were provided by DAMTP and Trinity Hall.

I wish to thank the faculty in the Relativity and Gravitation Group of Department of Applied Mathematics and Theoretical Physics, especially to Professor Neil Turok for his constructive criticism and advice for the work presented, and for providing invaluable advice and help in pursuing my academic career; to Dr. Carlos Martins for important discussions and spending much of his precious time in revising the dissertation; to Dr. Serge Winitzki for an enjoyable collaboration; to the computer officer Dr. Stuart Rankin who eagerly instructed me on how to use the computing facilities and resolved uncountable relevant problems; to Dr. Richard Battye, Dr. Robert Caldwell, Dr. Robert Crittenden, and Dr. Andrew Sornborger for many enlightening discussions; to Gareth Amery and Martin Landriau for creating the relaxed atmosphere in the office and for revising the dissertation in the final stage of the writing. I also thank Uros Seljak and Matias Zaldarriaga for the use of their fast Boltzmann code, CMBFAST.

I would also like to thank those who insisted I had an enjoyable time in Cambridge—in particular Maw-Kuen, Kuo-Ching, Wen-Kung, Jen-To, Ya-Fen, Jiann-Hua, Kuo-Feng, Meng-Fan, Mika, Koichiro, Daisuke, David, Paul, Tom, Andy, Bent, Dan, Julie, Kate, and Laura.

Last but not least, I would like to thank Professor Andrew Liddle and Dr. Robert Smith, who encouraged me and recommended that I come to Cambridge to pursue my PhD after I received my MSc degree from Sussex Astronomy Centre.

List of Figures

List of Tables

Contents

OVERVIEW

With the cosmological principle as the basic premise, there are currently two main competing theories for the origin of structure in the universe—inflation and cosmic defects. While the beauty and simplicity of the former appears to have enticed more adherents, this dissertation attempts to evaluate the latter, in particular cosmic strings, which have proved to be analytically and numerically more challenging due to their non-linear effects. We endeavour to make robust predictions for distinguishing observational features of string-induced perturbations. This dissertation has been divided into four main parts, each with several chapters:

I Introduction: In Chapter 1, we begin with the basics of the standard cosmology, together with its successes and problems. In this context, we stress the importance of the study of large-scale structure formation of the universe. We describe the fundamental features of cosmological perturbations and the observational tests which are commonly applied to test theories of structure formation. These observational tests will be one of the themes of the dissertation. We then review both inflationary and cosmic defect theories in Chapter 2. In particular, we introduce the groundwork for cosmic strings. Also laid out in this part are the conventions and the mathematic foundations used in the dissertation.

II Perturbation theory for causal models: In Chapter 3, we first derive general solutions for the dynamics in Friedmann-Robertson-Walker models, including the effect of a non-zero cosmological constant Λ. Around these homogeneous and isotropic background cosmologies, we then develop in the synchronous gauge the perturbation theory for causal sources, using the fluid approximation with the baryonic effect included. We numerically verify this framework for matter and radiation density perturbations in the context of the standard CDM model. In Chapter 4, we analytically solve the underlying matter perturbation equations in a flat $\Lambda = 0$ model using Green functions. A simple extrapolation scheme, whose accuracy is numerically verified, is then proposed to obtain solutions in other cosmologies. Some general features of these solutions are revealed. In particular, we analytically prove that the resulting matter density perturbations are independent of the way the causal source was compensated into the background contents of the universe when it was formed. We also use our Green function solutions to investigate the compensation mechanism of perturbations with causal seeds, and yield a mathematically and physically explicit form in interpreting it.

III Large-scale structure (LSS) with cosmic strings: Important intrinsic and phenomenological properties of cosmic string networks are first investigated and summarised in Chapter 5. In Chapter 6, the formalism developed in part II is then employed to calculate the matter perturbations induced by a cosmic string network, from which large-scale structure power spectra are inferred. Two further applications using these results are made: Chapter 7 investigating the level of non-Gaussianity with string-seeded matter perturbations, and Chapter 8 using cluster abundances to constrain the amplitude of matter density perturbations at the scale of $8h^{-1}$Mpc, σ_8. The non-Gaussian scale is found to be $\lesssim 1.5(\Omega_{\rm m0}h^2)^{-1}$Mpc, while the constraint on σ_8 from cluster abundances is shown to be consistent with the results using COBE normalisations for string models.

IV Cosmic microwave background (CMB): As an example of causal theories, we compute in Chapter 9 the string-induced CMB fluctuations in the small-field limit, relevant for comparison with experiments on small-angular scales. This includes the map making for both the primary perturbations and the integrated Sachs-Wolfe effect. We then investigate non-Gaussian tests of CMB maps as a technique for distinguishing between different models of structure formation. Two different new methods are proposed: the one in Chapter 9 being capable of extracting the non-Gaussian signal from a very strong Gaussian background, and the one in Chapter 10 being sensitive to scale-scale correlations which are generically expected in defect models. While an accurate estimate of defect-induced power spectra is difficult and computationally expensive to obtain, the non-Gaussian test could provide with another observational window through which one can easily justify or disprove the defect models. Even if defects are unimportant as the underlying mechanism for structure formation, these methods for the CMB non-Gaussian test can still be applied to constrain the existence of defects at significantly lower energies.

We conclude that although future observations with higher accuracy may favour one of the competing models for structure formation, this will not necessarily be catastrophic for either inflation or defects, which both have a much broader and generic role in cosmology and fundamental physics. Instead, combined with such observations, we hope the work presented in this dissertation will help advance our understanding of the respective roles of these theories.

Part I

INTRODUCTION

Chapter 1

Introduction

1.1 Overview

In this chapter, we first define the units and conventions of the dissertation. Section 1.3 then introduces the basics of the standard cosmology, together with its successes and problems. Within this context, we highlight the problem of structure formation of the universe. Some fundamental properties of cosmological perturbations are then briefly reviewed in section 1.4. This includes the matter perturbations resulting in the large scale structure of the universe, and the radiation perturbations giving rise to anisotropies in the cosmic microwave background radiation. In section 1.5, we review the commonly applied observational tests of theories for structure formation. We shall tailor our models to these observational tests in the latter parts of the dissertation, especially for cosmic string theories.

1.2 Units and conventions

Throughout this dissertation, we use a signature $(+---)$ for the spacetime metric $g_{\mu\nu}$. The induced metric on the string worldsheet γ_{ab} will have a signature $(+-)$. We use Greek alphabet as spacetime indices (e.g. $\mu = 0, 1, 2, 3$), mid-alphabet Latin letters as spatial indices (e.g. $i = 1, 2, 3$), and Latin letters from the beginning of the alphabet as the indices of string worldsheet (e.g. $a = 1, 2$).

Unless otherwise stated, the units of all physical quantities are normalized to

$$\hbar = c = Gm_{\rm Pl}^2 = k_B = 1. \tag{1.1}$$

The Planck time and mass are approximately

$$t_{\rm Pl} \sim 5.4 \times 10^{-44} \text{ sec}, \tag{1.2}$$
$$m_{\rm Pl} \sim 1.2 \times 10^{19} \text{ GeV}. \tag{1.3}$$

Astrophysical distances will be expressed in parsecs:

$$1\text{pc} \approx 3.2615 \text{ light-year} \approx 3.0856 \times 10^{18}\text{cm}, \tag{1.4}$$

and $1\text{Mpc} = 10^6\text{pc}$. Astrophysical masses will be expressed in solar masses:

$$1\text{M}_\odot \approx 1.989 \times 10^{33}\text{g} \approx 1.116 \times 10^{57}\text{GeV}. \tag{1.5}$$

The subscript '0' will denote quantities measured at the present epoch. For quantities relevant to the composition of the universe, we use the subscript 'm' for matter, including 'c' for cold dark matter (CDM), 'h' for hot dark matter (HDM), and 'B' for baryons; we use the subscript 'r' for radiation, including 'γ' for photons, and 'ν' for neutrinos. In the discussion of string-related quantities, the subscript '∞' refers to that of the long (infinite) string population, and the subscript 'L' denotes that of the small loops generated from the long-string network. For other abbreviations used in this dissertation, please refer to Appendix A.

We define the n-dimensional Fourier transform of a function $f(\mathbf{x})$ as:

$$\widetilde{f}(\mathbf{k}) = \frac{1}{V_n} \int d^n\mathbf{x}\; f(\mathbf{x}) e^{-i\mathbf{k}\cdot\mathbf{x}}\,, \tag{1.6}$$

so that

$$f(\mathbf{x}) = \frac{V_n}{(2\pi)^n} \int d^n\mathbf{k}\ \tilde{f}(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}}\,, \tag{1.7}$$

where $\mathbf{k}$ is the wave vector in Fourier space, $\mathbf{x}$ the coordinates in real space, and the integration in (1.6) taken over a large n-dimensional volume V_n in real space. We shall use a tilde $\tilde{}$ over a quantity to denote its Fourier transform. Finally, we shall use '$\log_{10}$' for logarithms of base ten, and 'log' or 'ln' for natural logarithms of base e.

1.3 The standard cosmology

In this section, we review the standard cosmology [11, 12, 13]. The aims are to lay the foundations for the rest of the dissertation by examining the big-bang model, highlighting both its successes and its outstanding conundrums, and also the key confrontations with observation.

1.3.1 The expanding universe

Cosmological principle and FRW metric

The hot big-bang theory is based on the 'cosmological principle', which states that the universe is spatially homogeneous and isotropic as a whole. The Friedmann-Robertson-Walker (FRW) line element possesses this property:

$$ds^2 = dt^2 - a^2(t)d\ell^2, \tag{1.8}$$

where

$$d\ell^2 = \frac{dr^2}{1-Kr^2} + r^2(d\theta^2 + \sin^2\theta d\phi^2). \tag{1.9}$$

Here t is the physical time, ℓ is the comoving distance, (r, θ, ϕ) are spherical comoving coordinates, $a(t)$ is the scale factor describing the overall time evolution of the universe, and the constant K measures the spatial topology and geometry of the universe:

$$\begin{array}{ll} K > 0: & \text{three-sphere } S^3 \text{ (closed)}, \\ K = 0: & \text{Euclidean three-space } R^3 \text{ (flat)}, \\ K < 0: & \text{three-hyperboloid } H^3 \text{ (open)}. \end{array} \tag{1.10}$$

The homogeneity of the matter content of the universe means that fluid flow lines are along constant (r, θ, ϕ) trajectories, therefore the physical distances are determined by multiplying comoving distances by the scale factor, i.e.

$$d_{\text{phy}} = a(t)\ell. \tag{1.11}$$

Analogous to comoving coordinates, it is often convenient to define an alternative time coordinate, the conformal time η:

$$d\eta = \frac{dt}{a(t)}. \tag{1.12}$$

This simplifies the FRW metric to the form $ds^2 = a^2(\eta)(d\eta^2 - d\ell^2)$. Hereafter, we shall use a dot to represent a differentiation with respect to the physical time t or the conformal time η, depending on the context. A variety of cosmological consequences follow immediately from the FRW metric (1.8). Unless otherwise indicated, we shall refer ourselves to be a comoving observer in the following discussion.

Hubble's law

When considering two comoving observers separated by a comoving distance ℓ, there exists a recession velocity between them due to the expansion of the universe. This can be seen by differentiating equation (1.11) with respect to the physical time t:

$$v_{\text{phy}} = Hd_{\text{phy}}, \tag{1.13}$$

where $v_{\text{phy}} = \dot{d}_{\text{phy}}$, and H is the Hubble parameter:

$$H = \frac{\dot{a}}{a}. \tag{1.14}$$

The relation (1.13) is the well-known Hubble law: distant comoving objects are receding from us with velocities in proportion to their distances. A commonly used dimensionless quantity reflecting the present value of the Hubble parameter is defined as

$$h = \frac{H_0}{100\text{km s}^{-1}\text{Mpc}^{-1}}. \tag{1.15}$$

As we shall see, one of the greatest observational challenges has been to measure this quantity, showing that h is of order unity but with considerable uncertainties.

Cosmological redshift

Another consequence resulting from the FRW metric is the stretching of light when the universe expands. Suppose two successive wavefronts of light are emitted at t_e and $t_\mathrm{e} + \delta t_\mathrm{e}$ from a distant comoving object, and received by us at t_0 and $t_0 + \delta t_0$ respectively. Since the comoving distances travelled by these two wavefronts are the same, it is easy to show by using equation (1.8) and taking $ds = 0$ for light rays that

$$\int_{t_\mathrm{e}}^{t_\mathrm{e}+\delta t_\mathrm{e}} \frac{dt}{a(t)} = \int_{t_0}^{t_0+\delta t_0} \frac{dt}{a(t)}, \tag{1.16}$$

which leads to $\lambda_\mathrm{e}/a(t_\mathrm{e}) = \lambda_0/a(t_0)$, where $\lambda \equiv \delta t$ is the light wavelength. This allows us to define a quantity measuring the expansion of the universe via the stretching of light, the redshift z:

$$1 + z = \frac{a(t_0)}{a(t_\mathrm{e})} = \frac{\lambda_0}{\lambda_\mathrm{e}} = \frac{\nu_\mathrm{e}}{\nu_0}, \tag{1.17}$$

where ν is the light frequency.

Einstein equations and the big bang

To investigate the time evolution of the universe, i.e. the time dependence of the scale factor $a(t)$ in a FRW model, we now need the Einstein equations:

$$G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = 8\pi G T_{\mu\nu} + \Lambda g_{\mu\nu}, \tag{1.18}$$

where $G_{\mu\nu}$ is the Einstein tensor, G is the gravitational constant, $T_{\mu\nu}$ is the energy-momentum tensor, and Λ is the cosmological constant. In the standard cosmology, the energy-momentum tensor $T_{\mu\nu}$ for the content of the universe must have the same symmetries as the FRW metric, and a key example of a perfect fluid is:

$$T_{\mu\nu} = (\rho + p)u_\mu u_\nu - p g_{\mu\nu}, \tag{1.19}$$

where the energy density ρ and the pressure p are functions of time only, and $u^\mu = (1, 0, 0, 0)$ is the four-velocity of the comoving matter. Combined with the FRW model as illustrated by (1.8), the Einstein equations (1.18) then lead to two non-trivial equations: the Friedmann equation

$$H^2 + \frac{K}{a^2} = \frac{8\pi G}{3}\rho + \frac{\Lambda}{3}, \tag{1.20}$$

and the Raychaudhuri equation

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3p) + \frac{\Lambda}{3}. \tag{1.21}$$

From the Raychaudhuri equation we see that if $\Lambda = 0$ and the strong energy condition $\rho + 3p > 0$ is not to be violated, then $\ddot{a}(t) < 0$. Adding the fact that the currently observed universe is expanding (the Hubble expansion), i.e. $\dot{a}(t) > 0$, we can conclude that the scale factor, or equivalently the size of the universe, must have equalled zero in the past. This is the well-known 'big bang'. When $K = \Lambda = 0$, it is generally referred as the Einstein-de Sitter cosmology.

From the Friedmann equation (1.20) and the Raychaudhuri equation (1.21), one can derive the 'fluid equation'

$$\dot{\rho} + 3\frac{\dot{a}}{a}(\rho + p) = 0, \tag{1.22}$$

which is also a direct result from the covariant conservation of energy $T^{\mu\nu}{}_{;\nu} = 0$. This equation tells us that the time evolution of the energy density of the universe (the first term) results from the dilution of the energy due to the expansion (the second term), and from the work done by the pressure of the fluid (the third term). With the equation of state, $p = p(\rho)$, one can fully specify the evolution of p and ρ, as functions of the scale factor $a(t)$. For example, if the universe is filled with radiation, then the equation of state is that of an ideal relativistic gas $p = \rho/3$. On the other hand, if the universe is dominated by non-relativistic matter with negligible pressure, we have $p \ll \rho$. These two cases lead to the solutions:

$$\rho \propto \begin{cases} a^{-4}, & \text{radiation-dominated } (p = \rho/3), \\ a^{-3}, & \text{matter-dominated } (p \ll \rho). \end{cases} \tag{1.23}$$

Radiation, matter, curvature, and Λ eras

In the standard model, the main composition of ρ is assumed to be from radiation (photons and neutrinos) and non-relativistic matter (baryons and dark matter), i.e. $\rho \approx \rho_{\rm r} + \rho_{\rm m}$. It follows that the evolution of H in the Friedmann equation (1.20) will be dominated chronologically by one of the terms if the relative amplitude of each term is not too small: the radiation term ($\propto a^{-4}$), the matter term ($\propto a^{-3}$), the curvature term ($\propto a^{-2}$), and the cosmological term ($\propto a^{0}$). These cases correspond to the radiation-, matter-, curvature-, and Λ-dominated eras respectively, each with a simple solution:

$$a(t) \propto \begin{cases} t^{1/2}, & \text{radiation era}, \\ t^{2/3}, & \text{matter era}, \\ t, & \text{curvature era with } K < 0, \\ 1 - \cos(At), & \text{curvature era with } K > 0, \\ \exp(\sqrt{\Lambda/3}t), & \Lambda \text{ era}. \end{cases} \tag{1.24}$$

where A is a constant. An important epoch in cosmology is the so-called 'radiation-matter equality' at which $\rho_{\rm r} = \rho_{\rm m}$. We shall use the subscript 'eq' to denote quantities measured at radiation-matter equality. We also notice that in the curvature era with $K > 0$, the solution will end with $a(t) = 0$ if the Λ term is not too large. This means that the volume of the universe will eventually shrink to zero, and we refer to this as the 'big crunch', the inverse of the big bang. The solutions to the Friedmann equation will be studied in more detail in section 3.2.

Critical density and the density parameter

The critical energy density is generically defined as

$$\rho_{\rm crit} = \frac{3H^2}{8\pi G}. \tag{1.25}$$

This is the required total energy density, including that from Λ if present, to maintain a geometrically flat universe $K = 0$. Densities are often measured as a fraction of $\rho_{\rm crit}$, which we define as the density parameter:

$$\Omega_N = \frac{\rho_N}{\rho_{\rm crit}}. \tag{1.26}$$

Here N indicates different energy composition components of the universe, including the vacuum energy resulting from Λ, i.e. $\Omega_\Lambda = \Lambda/3H^2$. Thus $\Omega = \sum_N \Omega_N < 1$, $\Omega = 1$ and $\Omega > 1$ corresponds to open, flat and closed universes respectively. We also notice that if $\Lambda = 0$, the critical energy density can be obtained from equations (1.24) and (1.25) as:

$$\rho_{\rm crit}(t) = \begin{cases} \frac{3}{32\pi G t^2}, & \text{radiation era}, \\ \frac{1}{6\pi G t^2}, & \text{matter era}. \end{cases} \tag{1.27}$$

The value of $\Omega_0 \approx \Omega_{\rm m0} + \Omega_{\Lambda 0}$ (because $\Omega_{\rm r0} \ll \Omega_{\rm m0}$) is one of the key questions in cosmology which remains to be answered, because it determines not only the geometry of the universe, but also the dynamics of its evolution and ultimately its future.

Cosmological horizons and the flatness

One vital property of the standard cosmology is that it possesses horizons: the distance light can travel since the big bang is finite. From the FRW metric (1.8), one can calculate this distance

$$d_{\rm H}(t) = a(t) \int_0^t \frac{d\hat{t}}{a(\hat{t})} = a(t)\eta, \tag{1.28}$$

where we notice that the comoving horizon size is just the conformal time η. Thus we have $d_{\rm H} = 2t = H^{-1}$ in the radiation era and $d_{\rm H} = 3t = 2H^{-1}$ in the matter era. As a consequence, we have three characteristic scales in the big-bang theory:

1. The comoving particle horizon $d_{\rm H}/a = \eta$.
2. The comoving Hubble length/radius $H^{-1}/a = \dot{a}^{-1}$.
3. The curvature scale $|K|^{-1/2}$.

The first measures the distance a photon can travel since the big bang; the second is the typical size of the currently 'visible' universe; the last gives the maximum lengthscale below which the geometry of the universe can be considered as flat (Euclidean). The first two scales, the particle horizon and the Hubble length, are equivalent in the standard cosmology, although may be different in other cases like inflation as we shall see. Thus the ratio of the comoving Hubble length and the curvature scale gives a measure of Ω according to the Friedmann equation (1.20):

$$|\Omega - 1| = \left(\frac{H^{-1}/a}{|K|^{-1/2}}\right)^2. \tag{1.29}$$

In a standard model because $\ddot{a}(t) < 0$ according to equation (1.21), we see in (1.29) that Ω departs from unity as time increases if $K \neq 0$, and that Ω equals unity forever if $K = 0$. In particular, we have

$$|\Omega - 1| \propto \begin{cases} t, & \text{radiation era,} \\ t^{2/3}, & \text{matter era.} \end{cases} \tag{1.30}$$

We also note that as long as $\Omega \approx 1$, equation (1.29) implies that the comoving Hubble length is much smaller than the curvature scale, so the observable universe can be treated as being flat.

1.3.2 Successes vs. conundrums

The historical development of modern cosmology can be roughly divided into three stages. The first stage is the well-known competition between the steady-state theory and the hot big-bang model. Even Einstein was once susceptible to the appeal of the perfect cosmological principle. It is the foundation of the steady-state theory and states that the universe is not only spatially homogeneous and isotropic on the average, but also unchanging in time. However, the vital discovery of the Cosmic Microwave Background Radiation (CMBR) in 1965 effectively sounded the death knell for the steady-state theory, enhancing the development of our understanding in cosmology into the second stage, where we saw the successes of the big-bang model. These successes include

1. **The Hubble expansion**: The first success in testing the standard cosmology was through the observational verification of Hubble law. It is a direct consequence of the FRW model as we have seen previously.
2. **The CMBR**: In the early universe, all matter was fully ionised until the temperature dropped below the hydrogen ionisation threshold of 13.6 eV, when electrons and protons began to combine to form hydrogen atoms. This epoch is called 'recombination' or 'decoupling', and will be denoted with a subscript 'd' [13]:

$$\begin{aligned} t_{\rm d} &\approx 5.6 \times 10^{12} (\Omega_{\rm m0} h^2)^{-1/2} {\rm sec}, \\ z_{\rm d} &\approx 1100. \end{aligned} \tag{1.31}$$

   Soon after this, a dramatic fall in the free electron density resulted in the rapid drop of the Thomson scattering rate. Thus photons were free to propagate since this epoch, which is normally referred as 'last scattering', and will be denoted with a subscript 'ls'. In the standard model, the photons from the last-scattering surface arrive today with a blackbody spectrum, and a temperature of approximately $3K$. Excitingly, this is indeed what Penzias and Wilson observed in 1964 [14]. The present measure of this residual temperature is [15, 16, 17, 18, 19]

$$T_0 = 2.728 \pm 0.004K, \tag{1.32}$$

   corresponding to an energy density $\rho_\gamma \approx 4.5 \times 10^{-34}{\rm g\ cm}^{-3}$.
3. **Big Bang Nucleosynthesis (BBN)**: As mentioned earlier, the universe in the standard model is composed principally of radiation and non-relativistic matter. The latter includes baryons and dark matter. Nucleosynthesis is a process expected to take place in the first few minutes after the big bang, forming light elements (^4H, D, ^3He, ^7Li) that are the main baryonic particles we can observe today. The consistency between the predicted abundances of these light elements and the measured ones not only provides us with a strong evidence for the hot big bang [20], but also pins down the value of baryon density today [21]

$$\Omega_{\rm B0} = \frac{\rho_{\rm B0}}{\rho_{\rm crit0}} = (0.02 \pm 0.002) h^{-2}, \tag{1.33}$$

   which gives $\Omega_{\rm B0} \approx 0.05$ for $h \approx 0.65$.

Despite these major successes witnessed in the last few decades, there are still some conundrums which cannot be explained at present by the standard model. In the face of these problems, we have entered the third stage in developing cosmology. Foremost among them are:

1. **The flatness problem**—why is the currently observed universe so close to spatial flatness? As we have seen in the discussion after equation (1.29), the value of Ω will move away from one during the standard big-bang evolution. Even if we know that the current value of Ω is of order one with considerable uncertainties, it is easy to show that its value in the past would have been extremely close to unity in the standard big-bang model, e.g. $|\Omega - 1| < \mathcal{O}(10^{-12})$ for $t \approx 1$ sec. In other words, its value should have been diluted to almost zero or else grown to infinity as the universe recollapsed within the first few seconds since the creation, unless the universe is precisely flat at the beginning. It would be unnatural to have $\Omega = 1$ as the initial condition because this is like accidently choosing the origin when randomly selecting a point in an infinite dimensional parameter space. So the question is, what sort of mechanism can set up such a condition at the very early times for our universe?

2. **The horizon problem**—why is the observed CMB temperature so homogeneous and isotropic? As we have seen in equation (1.28), the distance photons can travel since the big bang is finite, so the CMB sky, which corresponds to the last-scattering surface, should not have had enough time to have established thermal equilibrium throughout the entire universe by the time of decoupling. More precisely, one can calculate on the last-scattering surface the angular size of the horizon at that time, and this leads to $\eta_{\rm d}/\eta_0 \sim 1$ degree. This implies that the CMB sky should be in theory composed of many independent regions with their own temperatures. Obviously this is not the case from CMB observations because temperatures from opposite directions are seen to be the same with an accuracy of about 10^{-5}. Hence we have to ask what can make the CMB temperatures so uniformly distributed, without embeding this into the initial conditions.

3. **The monopole problem**—where are the monopoles and other relics like domain walls or gravitinos today? In unified theories, a great number of magnetic monopoles must be formed whenever the symmetry of fundamental interactions breaks (e.g. grand unified theories (GUT's) incorporating electromagnetism, the weak interaction and the strong interaction). Because monopoles are very massive and stable particles, even if only a small amount is formed it should be still quite easy to detect them in the present universe. On the contrary, none have yet been observed, and the same is true for other relics.

4. **The singularity problem**—what was before the Planck epoch? As we have seen previously, when $t \to 0$ the size of the universe goes to zero and its energy density goes to infinity. In this situation, the framework of classical general relativity breaks down. Although this implies that quantum gravity is probably essential in understanding the origin of the universe, this sigularity problem still remains to be resolved.

5. **The dark matter problem**—what exactly is the dark matter, and does it actually exist? Accoring to observations, the value of $\Omega_{\rm m0}$ is of order one, while the observed baryon density is $\Omega_{\rm B0} \approx 0.05$; the rotation curves of galaxies have been shown to possess flat tails, while their baryonic density is seen to be exponentially decaying outwards. These have been taken as strong implications for the existence of dark matter, although none has been directly detected. However, it could be that our understanding of these phenomena is still quite limited.

6. **The structure formation problem**—how do the structures of our current universe form? According to the cosmological principle stated earlier, the universe should be homogeneous and isotropic after the big bang. Therefore, the observed structures of the universe today implies that there must be some mechanism which can produce initial irregularities that have grown through graviational instability to form the structures today. This mechanism is absent in the standard big-bang cosmology, and it is one of the main themes we shall be investigating in this dissertation.

Cosmology is at an exciting watershed. As we shall see later, progress has been made by theorists in solving the above conundrums with grander theories. Moreover, an avalanche of new cosmological observations of increasing accuracy is continuing to pin down the ranges for the values of various cosmological parameters in the standard cosmology. In Table 1.1, we summarise some recent results for estimates of the main cosmological parameters. Here we see prima facie evidence that supports the existence of the cosmological constant Λ and a low matter energy density $\Omega_{\rm m0}$, although both still have considerable uncertainties. In addition, the Hubble parameter h seems to be smaller than unity, ranging approximately from 0.2 to 1 according to different observations. The only converging result among these estimates for cosmological parameters is that of the baryon energy density $\Omega_{\rm B0}$. A consistent result with most estimates is $\Omega_{\rm B0}h^2 \approx 0.02$. In the near future with

Recent estimates of cosmological parameters

$\Omega_{\rm m0}$	$\Omega_{\Lambda 0}$	$\Omega_{\rm B0}$	h	Method
$0.25^{+0.18}_{-0.12}$	$0.63^{+0.17}_{-0.23}$			CMB,SN Ia[22]
0.25 ± 0.2				lum.-lw.[23]
-1 ± 0.5 < 1	$\{0\}$ $\{1 - \Omega_{\rm m0}\}$			SN Ia (HST)[24]
$0.8\Omega_{\rm m0} - 0.6\Omega_{\Lambda 0} \approx -0.2 \pm 0.1$ with $\Omega_{\rm m0} \lesssim 1.5$				SN Ia[25]
$\{1 - \Omega_{\Lambda 0}\}$	< 0.66			grav. len.[26]
$\{1 - \Omega_{\Lambda 0}\}$	(0.29, 0.53)		(0.39, 0.58)	CMB+LSS[27]
	0.45 ± 0.2			clu. abun.[28]
	0.2 ± 0.1			clusters[29]
		$\Omega_{\rm B} h^2 = (0.007, 0.024)$		BBN[20]
		$\Omega_{\rm B} h^2 = 0.02 \pm 0.002$		BBN[21]
		(0.007, 0.041) ~ 0.021	 $\{h = 0.7\}$	global budget[30]
			$0.73 \pm 0.06 \pm 0.08$	HST [31]
$\{1\}$	$\{0\}$ $\{0\}$		$0.3^{+0.18}_{-0.07}$ (68%) $0.4^{+0.57}_{-0.14}$ (68%)	CMB[32]
			0.82 ± 0.16	gal. dyn.[23]

Table 1.1: Summary of recent estimates of cosmological parameters. We use brackets { } to indicate the assumed value(s) in deriving the result in the same column. The form (a, b) means that the underlying quantity is estimated to be within the range between a and b. The uncertainties are within 95% confidence level unless otherwise indicated within brackets ( ).

observations of higher quality and precision, we expect these estimates to improve very significantly. Thus, we find ourselves at an important stage in the development of cosmology, with the real prospect of gaining a better understanding of the origin of our universe.

1.4 Cosmological perturbations

One of the main themes of this dissertation is to investigate mechanisms which can form the structure of the universe we observe today. While leaving to the next chapter the discussion of possible candidates for theories of structure formation, here we discuss some generic features of the cosmological perturbations that give rise to the cosmic structures. The perturbations of current interest can be divided into two types: the matter perturbations which result in the large-scale structure of the universe, and the radiation perturbations which give rise to the CMB anisotropies. To study the behaviour of these perturbations in detail, one should in theory invoke general relativity, i.e. to perturb the Einstein equation (1.18). In this section, we shall however introduce only the well-known generic results, and leave a detailed treatment to Part II. As we shall see, the amplitudes of cosmological perturbations of interest in this dissertation are weak, so we shall work only in the linear regime throughout.

1.4.1 Matter perturbations

As mentioned earlier, some mechanism in the early universe must have produced primordial density perturbations, which then grew through gravitational instability to form the observed cosmic structures. This growth can be considered within two regimes: the subhorizon scales and the superhorizon scales.

On subhorizon scales, the theory of Newtonian gravity is adequate to study linear perturbations. By simple arguments based on fluid dynamics, one can show that there exists a scale below which the growth of density perturbations is suspended. This is the well-known Jeans length $\lambda_{\rm J}$, corresponding to a Jeans mass $M_{\rm J}$:

$$\lambda_{\rm J} = c_s \left(\frac{\pi}{G\rho}\right)^{1/2}, \tag{1.34}$$

$$M_{\rm J} = \frac{4\pi}{3}\rho\left(\frac{\lambda_{\rm J}}{2}\right)^3 = \frac{c_s^3}{6}\left(\frac{\pi^5}{G^3\rho}\right)^{1/2}, \tag{1.35}$$

where c_s is the sound speed and ρ is the mean density of the universe. This is a direct result of the competition between gravitational collapse and resistance from pressure. When the sound speed is larger resulting in higher pressure and therefore larger λ_J, the lengthscale threshold allowing gravitational collapse will be longer. In cosmic structure formation scenarios, we first have the radiation-dominated era in which $c_s^2 \approx 1/3$ so that the Hubble length $H^{-1} \approx \lambda_J/3$, meaning that no subhorizon structure can be formed. Later in the matter-dominated era when pressure effects become subdominant, $c_s \approx 0$ giving $\lambda_J \approx 0$ so that structures become unstable to collapse. Thus radiation-matter equality $t_{\rm eq}$ provides a threshold for the growth of matter perturbations.

A further factor affecting the growth of matter perturbations is the nature of the matter itself. We first consider the non-baryonic matter, which can be classified by its thermal velocities when the comoving galactic scale falls within the horizon: cold dark matter (CDM) is heavy and therefore non-relativistic by this time; hot dark matter (HDM) is light and thus still relativistic at this time. In a CDM universe, the perturbation evolution inside the Jeans length after $t_{\rm eq}$ can be shown to be

$$\delta_{\rm c} \equiv \frac{\delta\rho}{\rho} \sim At^{2/3} + Bt^{-1}, \tag{1.36}$$

where A and B are constants, and $\delta_{\rm c}$ is the 'density contrast' of CDM. Consequently the perturbations at some redshift z well after $t_{\rm eq}$ can be approximated as

$$\delta_{\rm c} \approx \left(\frac{1+z_{\rm eq}}{1+z}\right) \delta_{\rm c(eq)}. \tag{1.37}$$

When $\delta_{\rm c} \gtrsim 1$, we shall need non-linear theory to study the perturbations. Fortunately, within the scales of our interest, linear theory will be sufficient. As for a universe consisting of HDM, there will be an extra free-streaming effect, smearing away perturbations on small scales. This is due to the high thermal velocities, allowing for HDM simply to flow out of over-dense regions. For the baryonic matter in the universe, because it is coupled to radiation through Thomson scattering before the recombination at $t_{\rm d}$, it can not collapse to form structures until $t_{\rm d}$ when its sound speed drops to become non-relativistic. If the universe is dominated by non-baryonic matter, then after $t_{\rm d}$ (which is normally later then $t_{\rm eq}$), the baryonic matter simply falls into the potential well created by the non-baryonic matter since $t_{\rm eq}$.

Turning to superhorizon scales, as we shall show in Chapters 3 and 4, there is always a growing mode outside the Hubble length H^{-1} in both the radiation- and matter-dominated eras. However, superhorizon primordial perturbations exist only in inflationary models. For isocurvature models like topological defects, superhorizon seeds are well compensated by a deficit in the background energy, so that the superhorizon growing mode has no effect on these scenarios.

1.4.2 Radiation perturbations

The other important type of perturbations are the radiation perturbations. As mentioned earlier, photons were coupled to baryons through Thomson scattering since deep in the radiation era. After the recombination of baryons at $t_{\rm d}$, the interaction rate of Thomson scattering suddenly drops due to the much lower number density of electrons, so that photons can propagate freely at $t_{\rm ls}$. Therefore the matter distribution at $t_{\rm ls}$ will leave an imprint in the photon distribution at the last-scattering surface, which we can observe as CMB anisotropies today. We shall call these anisotropies the 'primary' CMB perturbations, meaning that they emerge originally from the last-scattering surface.

On the way to the observer from the last-scattering surface, the photons can be further disturbed via gravitational lensing of the matter distribution or by other seeds like topological defects. This will serve as an extra contribution to the CMB anisotropies apart from the primary CMB perturbations. This extra component is normally referred as the 'Integrated Sachs-Wolfe' (ISW) effect, or simply the 'Sachs-Wolfe' (SW) effect. Thus schematically we can write the CMB anisotropies as

$$\frac{\Delta T}{T} = \left.\frac{\Delta T}{T}\right|_{\rm prim} + \left.\frac{\Delta T}{T}\right|_{\rm ISW}. \tag{1.38}$$

In Chapter 3 and Part IV, we shall investigate this in greater detail.

1.5 Observational tests of theories for structure formation

In this section we briefly discuss some observational tests on theories of structure formation, and define the relevant physical quantities. We have divided the commonly used methods into three categories: the power spectrum of large-scale structure (LSS), that of the cosmic microwave background (CMB), and the non-Gaussian (NG) tests in both LSS and CMB. These are also the main topics which we shall concentrate on in Parts III and IV.

1.5.1 Large-scale structure

The first method is to calculate the 'power spectrum' of LSS. A comparison between the theoretical result and that inferred from observations will provide us with a consistency test. The power spectrum $\mathcal{P}(k)$ for a mass density distribution $\rho(\mathbf{x})$ is defined as the spherically symmetric Fourier transform of the autocorrelation function $\xi(r)$ of the density contrast $\delta(\mathbf{x})$, i.e.

$$\mathcal{P}(k) = \frac{1}{2\pi^2} \int \xi(r) \frac{\sin kr}{k} r dr = \frac{V \langle |\widetilde{\delta}_{\mathbf{k}}|^2 \rangle}{8\pi^3}, \tag{1.39}$$

where $k = |\mathbf{k}|$, r is the correlation distance, and $\widetilde{\delta}_{\mathbf{k}}$ is the Fourier transform of the fluctuations $\delta(\mathbf{x})$. Another useful statistic is the mass fluctuation amplitude at a certain length-scale λ, i.e. the standard deviation of the mass density distribution σ_λ. We convolve the density contrast $\delta(\mathbf{x})$ with a spherical window of radius λ to obtain the variance:

$$\sigma_\lambda^2 = 4\pi \int |w(k\lambda)|^2 k^2 \mathcal{P}(k) dk \quad \text{with} \quad w(x) = \frac{3(\sin x - x \cos x)}{x^3}. \tag{1.40}$$

We note that $w(x)$ is the Fourier transform of a spherical window and $\mathcal{P}(k)$ in (1.40) can be calculated from equation (1.39). From equation (1.40), we can further define the dimensionless 'spectrum of the matter density contrast', which measures the amplitude of perturbations over a logarithmic interval $\triangle k/k \sim 1$:

$$S(k) = \left\langle \left(\frac{\delta\rho}{\rho_{\text{av}}} \right)^2 \right\rangle_k \approx \frac{V k^3}{2\pi^2} \langle |\widetilde{\delta}_k|^2 \rangle = 4\pi k^3 \mathcal{P}(k). \tag{1.41}$$

This quantity is dimensionless and gives the rms density fluctuation on a particular length-scale $l = 2\pi/k$. Different theories of structure formation will predict different power spectra of LSS, which can then be tested against observationally inferred results. The one which we shall use to test our models is that presented in Ref. [33]. It is inferred from various galaxy surveys. We shall discuss this aspect in more detail for inflation in section 2.2.4, and for cosmic strings in Part III.

1.5.2 Cosmic microwave background

Another commonly used observational test is via the calculation of CMB power spectra. Although the CMB sky as a whole is quite homogeneous and isotropic, there exist observed anisotropies of order 10^{-5} [34], which reflect the matter distribution at the last-scattering surface. The CMB anisotropies observed at a position $\mathbf{x}$ along the line of sight (θ, ϕ) in spherical coordinates can be expanded in spherical harmonics (the multipole transform):

$$\frac{\Delta T}{T}(\mathbf{x}; \theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} a_{lm}(\mathbf{x}) Y_{lm}(\theta, \phi), \tag{1.42}$$

where

$$a_{lm}(\mathbf{x}) = \int d\Omega \frac{\Delta T}{T}(\mathbf{x}; \theta, \phi) Y_{lm}^*(\theta, \phi), \tag{1.43}$$

with $d\Omega \equiv \sin\theta d\theta d\phi$. The power spectrum of the CMB is then defined as

$$C_l = \left\langle |a_{lm}(\mathbf{x})|^2 \right\rangle_l, \tag{1.44}$$

where the brackets represent an ensemble average over $\mathbf{x}$ and m for a given l. This is also referred to as the C_l curve. Since the average over $\mathbf{x}$ is impossible to carry out in practice, the power spectrum is normally estimated at a given $\mathbf{x}$ as

$$C_l \approx \frac{1}{2l+1} \sum_{m=-l}^{l} |a_{lm}(\mathbf{x})|^2. \tag{1.45}$$

To show more explicitly what the multipole transform does, Figure 1.1 plots several spherical harmonics with different indices, corresponding to different angular scales. Generally speaking, the value of l is related to the angular size $\Delta\theta$ as $l \sim 1/\Delta\theta$ (where $\Delta\theta$ is in radians), while m represents different modes at that angular scale. Once again, different models will provide different predictions for the C_l's, which can then be tested against observations. We shall investigate this in some detail for inflation in section 2.2.4, and for cosmic strings in Part IV.

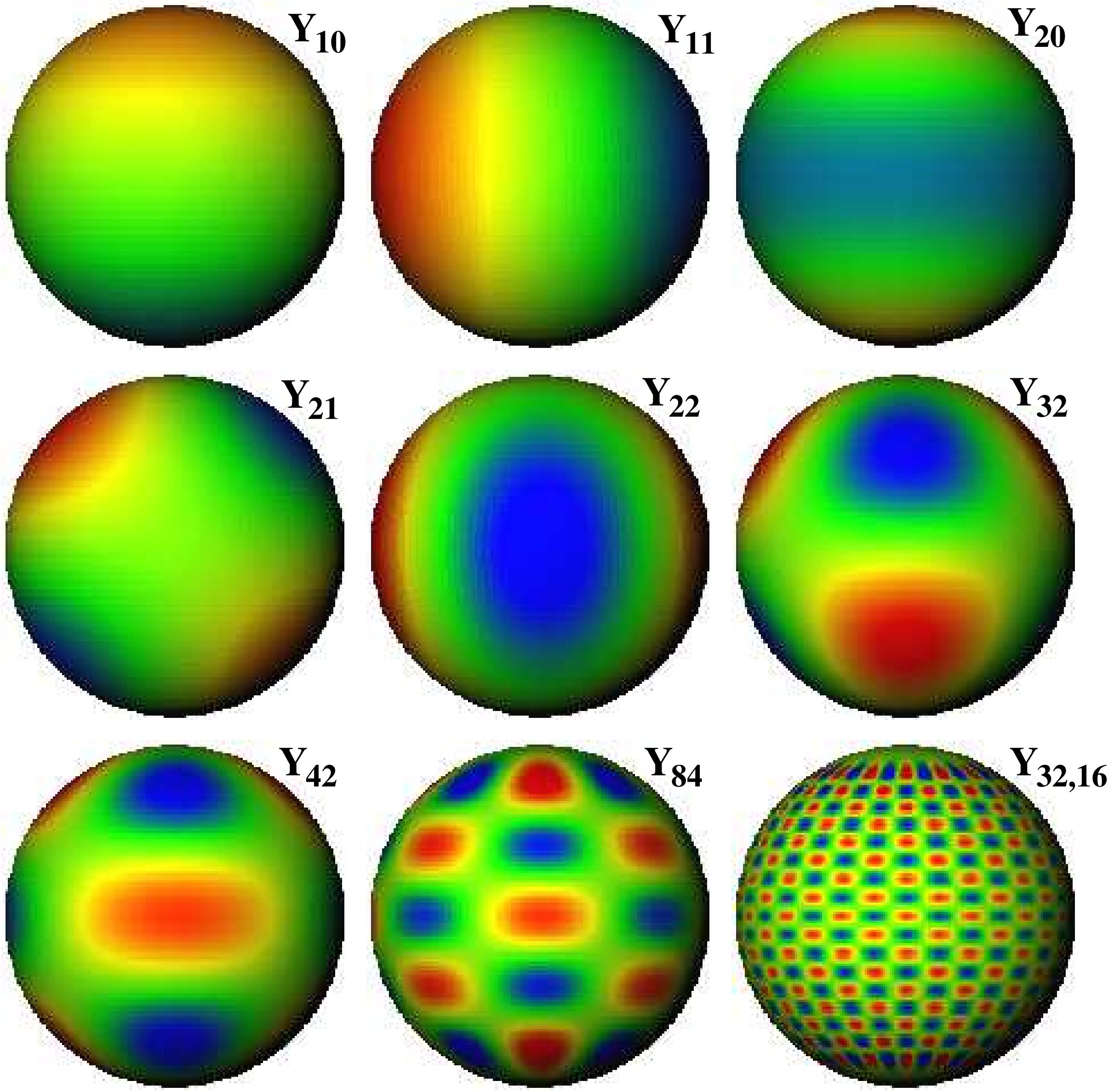

Figure 1.1: Some examples of spherical harmonics Y_{lm}. We plot only their real parts, with colour schemes interpolating between their minima (blue) and maxima (red). The value of l is related to the angular size $\Delta\theta$ as $l \sim 1/\Delta\theta$, while m represents different modes at that angular scale.

1.5.3 Non-Gaussian tests

Different models of structure formation can produce either a Gaussian or non-Gaussian distribution in the induced perturbations, depending on the details of their production mechanisms. For example, standard inflation predicts Gaussianity, while defect models supply some degree of non-Gaussianity. Thus a NG test in either the LSS or the CMB will give us a powerful tool for ruling out some of these models. However, because at late times the non-linear evolution of perturbations may operate on galaxy scales which itself creates NG, the LSS observations can hardly provide a fair sample for this purpose. On the other hand, the CMB is a snapshot of the density distribution at the last-scattering surface, an epoch well before the non-linear evolution regime begins. Thus this snapshot reflects the intrinsic statistical properties of perturbations, and we can thus use it as a discriminator of different models for structure formation. Accordingly, a discovery of non-Gaussianity in the CMB would have a strong impact on cosmology, meaning that inflation alone would not be sufficient to explain the observed structures in the universe. We shall investigate this issue further for LSS in Chapter 7, and for the CMB in Chapters 10 and 11.

1.6 Conclusion

We introduced the basics of cosmology with regard to the standard big bang model, and reviewed its successes and problems in the face of present observations. One of the key problems is the lack of mechanism for generating cosmological perturbations, whose generic properties were briefly discussed. The solution to this problem is to attach an accessory theory with such a mechanism to the standard model, and candidates for such theories will be reviewed in the next chapter. Finally, in this chapter we have introduced the commonly used observational tests of theories for structure formation. This includes the calculation of the power spectra for LSS and CMB, and non-Gaussian tests in both. These observational tests will be the main focus of the latter parts in the dissertation.

Chapter 2

Cosmological Perturbation Theories

2.1 Introduction

In the previous chapter, we saw that one of the outstanding problems in cosmology today is developing a more precise understanding of structure formation in the universe, that is, the origin of galaxies and other large scale structures. Existing theories for the structure formation of the Universe fall into two categories, based either upon the amplification of quantum fluctuations in a scalar field during inflation, or upon a symmetry breaking phase transition in the early Universe which leads to the formation of topological defects. While techniques for computing density perturbations for the former are well established [35, 36, 37], little quantitative work exists for the latter due to calculational difficulties in modelling nonlinear effects, especially for cosmic string models [38]. In this chapter, we shall review both inflation and topological defects, among which we shall concentrate on cosmic strings. Some of the materials in discussing the inflationary models are reconstructed from Ref. [39].

2.2 Inflation

2.2.1 Solving cosmological problems

The idea of inflation was originally proposed [40] to solve some of the cosmological problems mentioned in section 1.3.2. The definition of inflation is as simple as

$$\ddot{a}(t) > 0, \tag{2.1}$$

which states that the universe undergoes a period of accelerated expansion. The physical significance of this definition can be most easily seen through

$$\ddot{a} > 0 \Longleftrightarrow \frac{d}{dt}\left(\frac{H^{-1}}{a}\right) < 0. \tag{2.2}$$

As seen in the discussion before equation (1.29), there are three characteristic cosmological scales: the comoving particle horizon η, the comoving Hubble length H^{-1}/a, and the curvature scale $|k|^{-1/2}$. Thus (2.2) implies that the comoving Hubble length and therefore the comoving size of the visible universe are decreasing with time, while the curvature scale stays the same and the particle horizon keeps increasing, or at least not decreasing. This is the key property of inflation, with which a sequence of cosmological conundrums can be resolved:

1. **Solving the flatness problem**: The definition (2.2) automatically solves the flatness problem, because the comoving Hubble length being eventually much smaller than the curvature scale means that the universe is becoming spatially flat during inflation. More specifically, according to equation (2.2) we see that the right-hand side of equation (1.29) decreases during inflation, forcing Ω towards unity. Although we know that the post-inflation regime ($\ddot{a} < 0$) will repel Ω from one, Ω can still stay very close to unity right to the present provided that sufficient inflation occurs placing Ω so accurately to one. As we shall see later that a 'small' amount of inflation will be sufficient for this purpose and thus we solve the flatness problem.

2. **Solving the horizon problem**: This can be seen in either comoving coordinates or physical coordinates. In comoving coordinates, we know that the Hubble length is shrinking while the particle horizon is, at least, not decreasing. Therefore as long as we have enough inflation, the region of the universe which we

can see after inflation will be much smaller than the region which has been in causal contact. In terms of solving the horizon problem, this means

$$\eta_{\rm dec} = \left.\frac{d_{\rm H}}{a}\right|_{\rm dec} \gg \frac{H_0^{-1}}{a_0}, \tag{2.3}$$

so that the whole CMB sky we observe today was once in thermal equilibrium before inflation. Multiplying (2.3) with the scale factor $a(t_0)$, we see an alternative interpretation in physical coordinates: a smooth thermalized region before inflation was dramatically stretched out during inflation so that its size at the decoupling is greatly in excess of our currently observable universe. Thus we solve the horizon problem.

3. **Solving the monopole problem**: As mentioned in section 1.3.2, a great amount of magnetic monopoles will be produced during the spontaneous symmetry breaking. Even allowing for monopole-anti-monopole annihilations, they are still inefficient in avoiding monopole domination today. However, if sufficient inflation is inserted between the symmetry-breaking phase and the standard big bang era, this monopole or other relic problem can be resolved via the dramatical dilution of their energy density ($\rho_{\rm rel} \propto a^{-3}$) during inflation while the energy density of the universe $\rho_{\rm univ}$ dropping more slowly during this period as we shall see. As long as $\rho_{\rm univ}$ can be converted into the conventional matter at the end of inflation without recreating the unwanted relics, we solve the monopole problem.

2.2.2 Modelling inflation

As we have seen qualitatively, a period of inflation inserted before the standard big bang era can help solve several crucial cosmological problems. To concretise these concepts and therefore make proper calculations, we need to model the inflationary expansion. This is done through a scalar field, the inflaton $\phi(t)$, which is a function of time alone. For simplicity, we shall set $\Lambda = 0$ in the following discussion.

The use of scalar fields is well motivated by both condensed matter physics and particle physics. In condensed matter systems, scalar fields are well observed during phase transitions, and are normally referred as 'order parameters'; in particle physics, scalar fields play a crucial role in spontaneous symmetry breaking as we shall see later, although they are yet to be observed. Here, as a starting point, we write down the expressions for the energy density and the pressure of the universe satisfying the Lagrangian mechanism of an inflaton $\phi(t)$

$$\rho_\phi = \frac{1}{2}\dot{\phi}^2 + V(\phi), \tag{2.4}$$

$$p_\phi = \frac{1}{2}\dot{\phi}^2 - V(\phi). \tag{2.5}$$

These equations supplant the equation of state $p = (\gamma - 1)\rho$ for a barotropic fluid, and imply that inflation is so effective at diluting the density of other matter and the curvature that the energy density of the universe is dominated by the inflaton ϕ alone. Although this scalar field is yet to be found, it can be thought of as a new type of particle with spin zero and coupled with other 'ordinary' relativistic matter, the radiation. We also note that the potential energy here is not diminished by the expansion as it is simply a function of ϕ. Thus we can rewrite the Friedmann equation (1.20) and the fluid equation (1.22) as

$$H^2 + \frac{k}{a^2} = \frac{8\pi G}{3}\left[\frac{1}{2}\dot{\phi}^2 + V(\phi)\right], \tag{2.6}$$

$$\ddot{\phi} + 3H\dot{\phi} = -V', \tag{2.7}$$

respectively, where a prime $' \equiv d/d\phi$ denotes the derivative with respect to ϕ (although in this section only), and equation (2.7) is normally referred as the inflaton equation. This inflaton equation is an analogue to the equation of motion in classical dynamics, where ϕ is like the coordinate of a particle moving in a potential $V(\phi)$, subject to the frictional damping by the Hubble expansion. Moreover, by substituting equations (2.4) and (2.5) into (1.21) with $\Lambda = 0$, the definition of inflation (2.1) can be rewritten as

$$\ddot{a} > 0 \Longleftrightarrow p_\phi < -\frac{\rho_\phi}{3} \Longleftrightarrow V(\phi) > 2\left(\frac{\dot{\phi}^2}{2}\right), \tag{2.8}$$

indicating that the strong energy condition is violated by inflation, and that during inflation the energy density ρ_ϕ is dominated by the potential energy $V(\phi)$ with a portion of at least two thirds of the total energy.

To incorporate inflation into the standard cosmology, we need an appropriate scenario to bring the inflationary regime smoothly into the conventional big bang evolution. This includes specifying the form the potential $V(\phi)$, which generically defines various types of inflation. In literature, many different forms of the potential and therefore numerous types of inflation have been proposed. Here we show some examples in Table 2.1, and will present a schematic one that quotes most of the generic features of inflation.

Some inflationary models	
Seiberg-Witten inflation	multi-scalar inflation
topological inflation	chaotic inflation
hybrid inflation	quasi-open inflation
two-field inflation	two-stage inflation
M-theory inflation	D-term inflation
double D-term inflation	double inflation
supersymmetric inflation	new inflation
inverted hybrid inflation	dual inflation
pre-Big-Bang inflation	dynamical SUSY inflation
nonminimal chaotic inflation	multiple inflation
supergravity inflation	open inflation
open hybrid inflation	titled hybrid inflation
induced gravity open inflation	natural supergravity open inflation
natural supergravity inflation	late-time mild inflation
Brans-Dicke chaotic inflation	extended inflation
thermal inflation	supernatural inflation
warm inflation	false vacuum inflation
scalar-tensor gravity inflation	dilaton-driven inflation
gravity driven inflation	induced gravity inflation
natural inflation	moduli inflation
scalar-tensor stochastic inflation	Brans-Dicke inflation
superstring inflation	extra dimension inflation
false vacuum chaotic inflation	weak scale inflation
kinetic inflation	first order inflation
stochastic inflation	polynomial inflation
external inflation	intermediate inflation
hyperextended inflation	polynomial chaotic inflation
old inflation	power-law inflation
cosmic inflation	oscillating inflation
assisted inflation	quintessential inflation
monopole inflation	R-invariant topological inflation
running-mass inflation	string-forming open inflation
topological inflation	multiple-stage inflation
single-bubble open inflation	supersymmetric D-term inflation
chaotic new inflation	assisted chaotic inflation
isocurvature inflation	thermal brane inflation
supersymmetric hybrid inflation	k-inflation
F-term inflation	

Table 2.1: Some inflationary models proposed in the literature between the years 1992 and 1999. [41]

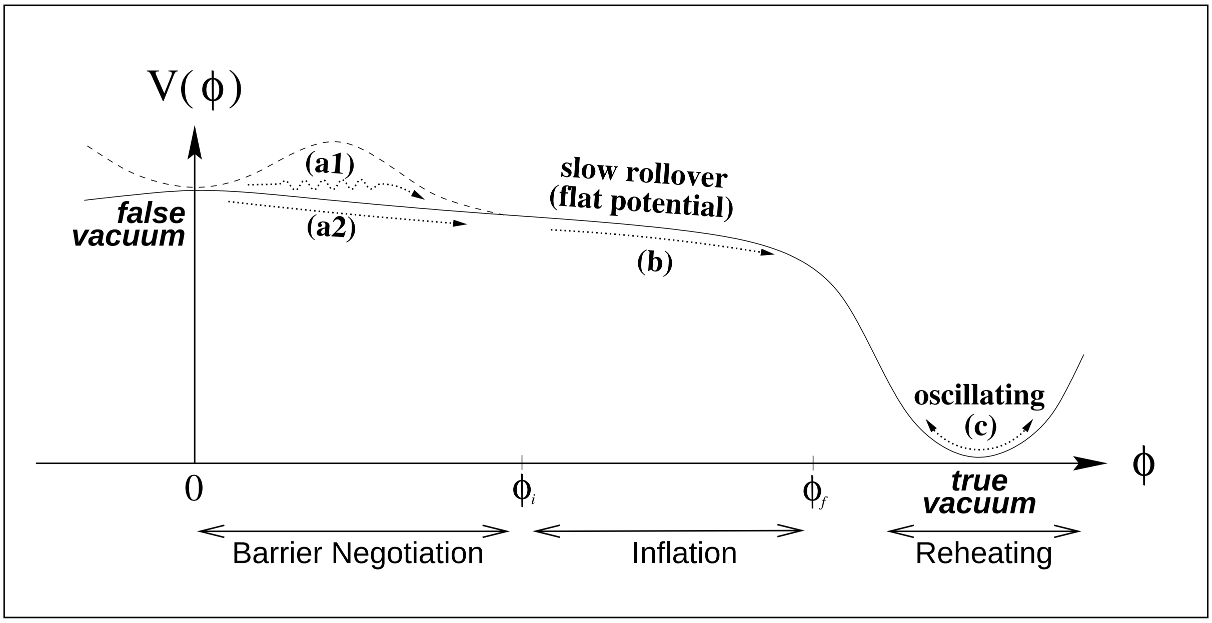

Figure 2.1: A schematic form of inflationary potential (adapted from [39]). The whole process for ϕ to roll from the false vacuum to the true vacuum can be divided into three stages: (a) barrier negotiation (if any), (b) inflation, and (c) reheating.

A scenario

Let us consider the inflaton potential as a function of temperature in the early universe. When the temperature is still well above the GUT scale $T_{\rm GUT} \sim 10^{16}$GeV, the inflaton potential possesses a global minimum, say at $\phi = 0$. As the temperature falls below a critical temperature $T_{\rm crit} < T_{\rm GUT}$, it develops a second minimum other than the original minimum at $\phi = 0$. Thus there are two vacuum states: the 'false vacuum' $\phi = 0$, which is the original minimum storing the energy of the universe above the GUT scale, i.e. $\rho_{\rm vac} = V(0) \equiv M^4 \geq (10^{16}{\rm GeV})^4$; and the 'true vacuum', which is the global minimum at $T < T_{\rm crit}$. The 'false' here means metastable and therefore temporary, while the 'vacuum' means the state of the lowest possible energy density. Thus when $T < T_{\rm crit}$, the inflaton will depart from the false vacuum as it is no longer the global minimum, moving towards the true vacuum. Figure 2.1 illustrates this situation.

In the due process, the inflaton will, if necessary, negotiate a barrier and then enter the inflationary regime governed by equation (2.7). At this stage, the potential $V(\phi)$ has two key features:

1. Flat enough: For inflation to occur, the potential must be flat enough so that the 'kinetic energy' $\dot{\phi}^2/2$ is at least two times smaller than the 'potential energy' $V(\phi)$ as required by (2.8).

2. Long enough: To solve the cosmological problems in the standard big bang model, the flat part of the potential mentioned above must be long enough so that the condition (2.8) can be held for a sufficiently long period to generate enough inflation.

As indicated by the phrases 'flat enough' and 'long enough', this period is normally referred as 'slow rollover'.

While ϕ continues to roll down the potential, a violation of (2.1) will bring inflation to the end. Once the inflaton ϕ reaches the true vacuum, it starts oscillating around it. As ϕ is coupled to other fields, in theory a damping term $\Gamma_\phi \dot{\phi}$ for the 'decay' of ϕ on a time scale Γ_ϕ^{-1} should be added into the left-hand side of equation (2.7) [42, 43, 44]. During inflation, $H^{-1} \ll \Gamma_\phi^{-1}$ so that the decay mechanism is negligible; during the oscillating phase, $H^{-1} \gg \Gamma_\phi^{-1}$ so that the the oscillation converts the scalar field into ordinary relativistic matter. This is the so-called 'reheating' process and can be regarded as a phase transition converting the energy stored in the false vacuum into the ordinary matter. When the matter thermalizes resulting in a temperature $T_{\rm reh}$ which is almost of the same order as the GUT scale, the scenario merges into the standard cosmology, the hot big bang evolution.

How much inflation

A characteristic quantity that measures the amount of inflation is the so-called 'number of e-foldings'. It is defined as

$$N(t_{\rm i}, t_{\rm f}) \equiv \ln\left[\frac{a(t_{\rm f})}{a(t_{\rm i})}\right] = \int_{t_{\rm i}}^{t_{\rm f}} H dt, \tag{2.9}$$

where the subscripts 'i' and 'f' indicate the beginning and end of inflation respectively. We notice that $t_{\rm f}$ is determined by the violation of equation (2.1), while $t_{\rm i}$ is model dependent and has a required minimum associated with the required minimum N for solving the cosmological problems. Thus one can show that in order to solve all the flatness, horizon, and monopole problems, the required number of e-foldings needs to be at least about 60, i.e. $N(t_{\rm i}, t_{\rm f}) \gtrsim 60$ [37]. This number can then be used to inspect various inflationary models for 'cosmological sufficient' inflation.

2.2.3 Analysis Techniques—PSRA and HSRA

Having modelled the inflationary mechanism, we now review the tools for the analysis and investigation of various inflationary models. Two main techniques are the potential-slow-roll approximation (PSRA) and the Hubble-slow-roll approximation (HSRA).

Potential-slow-roll approximation

In the schematic inflationary scenario presented in the previous section, we saw that the whole process of inflation is a slow rollover of the inflaton on the potential. Therefore the basic assumption of PSRA is just that both the 'kinetic term' $\dot{\phi}^2/2$ in equation (2.6) and the 'acceleration term' $\ddot{\phi}$ in equation (2.7) are negligible. This is an analogue to an object moving with a terminal velocity in the classical dynamics. Thus both equations become

$$H^2 \simeq \frac{8\pi G}{3} V(\phi), \tag{2.10}$$

$$3H\dot{\phi} \simeq -V'(\phi), \tag{2.11}$$

where the curvature term has also been ignored due to the flatness imposed by inflation.

Based on the basic assumption of PSRA, one can derive the following:

- We first square (2.11) and use (2.10) to eliminate H. Then the requirement $\dot{\phi}^2/2 \ll V(\phi)$ yields

$$\epsilon_V(\phi) = \frac{1}{16\pi G}\left(\frac{V'(\phi)}{V(\phi)}\right)^2 \ll 1. \tag{2.12}$$

- We use (2.10) and the differentiation of (2.11) with respect to t so as to eliminate both H and $\ddot{\phi}$ respectively in $|\ddot{\phi}| \ll |3H\dot{\phi}|$. A further use of $\dot{\phi}^2/2 \ll V(\phi)$ then gives

$$\eta_V(\phi) = \frac{1}{8\pi G}\left|\frac{V''(\phi)}{V(\phi)}\right| \ll 1. \tag{2.13}$$

Here ϵ_V and η_V are the so-called potential-slow-roll (PSR) parameters, and the equations (2.12) and (2.13) are called the PSR conditions. Although the basic assumption of PSRA results in the PSR conditions, the reverse does not hold. In other words, conditions (2.12) and (2.13) are just necessary but not sufficient to drop the kinetic and acceleration terms. Therefore, to complete the PSRA, an extra condition is required

$$\dot{\phi} \simeq -\frac{V'(\phi)}{3H(\phi)}, \tag{2.14}$$

since the information of $\dot{\phi}^2/2$ being small in (2.6) is lost in our derivation for the PSR conditions. This extra condition is called the 'attractor solution', which should be also checked during the slow rollover for the appropriateness of PSRA.

Now let us compare the PSR conditions with the generic condition for inflation. We find

$$\epsilon_V \overset{PSRA}{\simeq} \frac{-\dot{H}}{H^2} \ll 1 \Longrightarrow \frac{-\dot{H}}{H^2} < 1 \Longleftrightarrow \ddot{a} > 0. \tag{2.15}$$

This means that the PSR conditions $\epsilon_V \ll 1$ and $\eta_V \ll 1$ guarantee inflation to happen although they are not necessary. Therefore, the violation of PSR conditions provides a rough but analytically simple indicator to the end of inflation.

Armed with the PSRA, we are now in a position to derive the analytic solutions to the Friedmann and fluid equations. With their simplified forms (2.10) and (2.11), it is straightforward to show that

$$t(\phi) \simeq -\sqrt{24\pi G}\int \frac{[V(\phi)]^{1/2}}{V'(\phi)}d\phi, \tag{2.16}$$

$$\ln a(t) \simeq \sqrt{\frac{8\pi G}{3}}\int V(t)^{1/2}dt, \tag{2.17}$$

$$N(t_{\rm i}, t_{\rm f}) \overset{\phi(t)}{\equiv} N(\phi_{\rm i}, \phi_{\rm f}) \simeq -8\pi G\int_{\phi_{\rm i}}^{\phi_{\rm f}} \frac{V(\phi)}{V'(\phi)}d\phi = \left|\sqrt{4\pi G}\int_{\phi_{\rm i}}^{\phi_{\rm f}} \frac{d\phi}{\sqrt{\epsilon_V}}\right|. \tag{2.18}$$

Once $V(\phi)$ is specified, we automatically obtain these solutions. Here we can see qualitatively from equation (2.18) that even if only a small part of the potential is flat (i.e. $V(\phi)/V'(\phi) \gg 1$ or alternatively $\epsilon_V \ll 1$ within only certain range of ϕ) the number of e-foldings will obtain a significant boost. This illustrates not only the importance of the flatness in $V(\phi)$ for inducing sufficient inflation, but also the ease with which inflation solves those cosmological problems.

Hubble-slow-roll approximation

To invoke PSRA, we need to know the exact form of the potential $V(\phi)$. If the form of the potential is unknown, then PSRA is not applicable. Moreover, in spite of the simplicity of PSRA as shown above, the PSR conditions are just sufficient but not necessary for inflation to occur. Thus in situations where the potential form is invalid, or in studies like the perturbation theory where high precision in locating the inflationary regime is required, we need an alternative analysis technique—the Hubble-slow-roll approximation (HSRA) [45], in which $H(\phi)$ is the primary quantity rather than $V(\phi)$. To begin with, we introduce the Hamilton-Jacobi formalism. The trick here is to use ϕ instead of t as the time variable in Friedmann equation and the inflaton equation [46, 47, 48]. This requires a monotonic relation between ϕ and t, and we will choose $\dot{\phi} > 0$ throughout. We first differentiate (2.6) with respect to t (taking $k = 0$) and use (2.7) to yield

$$\dot{H} = -4\pi G\dot{\phi}^2, \tag{2.19}$$

which implies $\dot{H} < 0$. Then dividing each side by $\dot{\phi}$ to eliminate the time dependence gives

$$\dot{\phi} = -\frac{1}{4\pi G}H'(\phi). \tag{2.20}$$

Taking its square and using (2.6) to eliminate $\dot{\phi}$ further yield

$$[H'(\phi)]^2 - 12\pi G H^2(\phi) = -32\pi^2 G^2 V(\phi). \tag{2.21}$$

Equations (2.20) and (2.21) are called the Hamilton-Jacobi formalism, which is an alternative to equations (2.6) and (2.7).

Analogously to the PSRA, but using $H(\phi)$ as the primary quantity, we now define the conditions under which we may drop both the acceleration term in (2.7) and the kinetic term in (2.6):

$$\epsilon_H(\phi) = \frac{1}{4\pi G}\left(\frac{H'(\phi)}{H(\phi)}\right)^2 = 3\frac{\dot{\phi}^2/2}{V + \dot{\phi}^2/2} \ll 1, \tag{2.22}$$

$$\eta_H(\phi) = \frac{1}{4\pi G}\frac{H''(\phi)}{H(\phi)} = 3\frac{\ddot{\phi}}{3H\dot{\phi}} \ll 1, \tag{2.23}$$

where (2.20) and (2.21) have been used to eliminate H, H' and H''. These are called the Hubble-slow-roll (HSR) conditions, and ϵ_H and η_H are the HSR parameters. The use of the HSR conditions is called the HSRA. Since the smallness of both the kinetic term and the acceleration terms is apparent in HSR conditions, we shall not need any extra condition, like the attractor solution (2.14) in PSRA, to complete HSRA. However, it is worth noticing that a linear perturbation in $H(\phi)$ will possess the form resulting from equation (2.21)

$$\delta H(\phi) = \delta H(\phi_{\rm i})\exp\left(6\sqrt{\pi G}\int_{\phi_{\rm i}}^{\phi}\frac{-d\phi}{\sqrt{\epsilon_H(\phi)}}\right), \tag{2.24}$$

Thus we see that this linear perturbation is always decaying, no matter if inflation is occurring. In the inflationary regime, this decay is even more dramatic, because $\epsilon_H < 1$ implies that the decay is at least exponential. This illustrates the importance of the 'inflationary attractor', which assures at least exponential convergence among inflationary solutions with different initial conditions. This is a crucial feature of inflation, particularly useful in the application of the slow-roll approximation.

Let us compare the HSR conditions with the generic condition for inflation:

$$\ddot{a} > 0 \Longleftrightarrow \epsilon_H \overset{(2.20)}{=} \frac{\dot{H}}{H^2} < 1. \tag{2.25}$$

This means that a conservative HSR condition $\epsilon_H < 1$ is actually equivalent to the generic condition for inflation. Thus the violation of the HSR condition $\epsilon_H < 1$ provides a precise indicator for the end of inflation. This is a vital strength of HSRA over PSRA.

Within the regime of HSRA, it is straightforward to show that

$$t(\phi) = -4\pi G \int \frac{d\phi}{H'(\phi)}, \tag{2.26}$$

$$\ln a(t) = \int H(t)dt, \tag{2.27}$$

$$N(\phi_{\rm i}, \phi_{\rm f}) \overset{(2.20)}{=} -4\pi G \int_{\phi_{\rm i}}^{\phi_{\rm f}} \frac{H(\phi)}{H'(\phi)} d\phi \overset{(2.22)}{=} \sqrt{4\pi G} \int_{\phi_{\rm i}}^{\phi_{\rm f}} \frac{d\phi}{\sqrt{\epsilon_H(\phi)}}. \tag{2.28}$$

These solutions are exact, and this shows again the power of HSRA over PSRA. We also note that the HSR condition $\epsilon_H \ll 1$ automatically assures a large number of e-foldings.

Finally, we briefly investigate the relations between PSRA and HSRA. By using the definitions (2.12), (2.13), (2.22), (2.23), and the differentiation of (2.21) with respect to ϕ, one can easily obtain

$$\epsilon_V = \epsilon_H \left(\frac{1 - \eta_H/3}{1 - \epsilon_H/3} \right)^2, \tag{2.29}$$

$$\eta_V = (\epsilon_H + \eta_H) \left(\frac{1 - \eta_H/3}{1 - \epsilon_H/3} \right) + \frac{1}{\sqrt{4\pi G}} \frac{\sqrt{\epsilon_H}\eta_H'}{3 - \epsilon_H}. \tag{2.30}$$

To the lowest order, we have

$$\epsilon_V \simeq \epsilon_H; \quad \eta_V \simeq \eta_H + \epsilon_H. \tag{2.31}$$

This shows that the smallness of the HSR parameters implies the smallness of the PSR parameters, although this argument holds only for the lowest order. By recasting equation (2.21), we also have a relation linking $V(\phi)$ and $H(\phi)$:

$$[H(\phi)]^2 = \frac{8\pi G}{3} V(\phi) \left[1 - \frac{1}{3}\epsilon_H(\phi) \right]^{-1}. \tag{2.32}$$

Thus once $V(\phi)$ is known, $H(\phi)$ can be obtained, and vice versa, although the detailed treatment requires some subtleties.

2.2.4 Density perturbations and gravitational waves

A striking feature of inflation is that while providing sufficiently large inflation so as to solve the homogeneity problem, it can in addition leave enough irregularities behind to attempt to explain the observed structures in the universe today. This is done through a combination of a quantum mechanism and the fact that the comoving Hubble length is decreasing during inflation. Although the detailed process of producing inflationary perturbations is far too much for us to review here, the following should provide a reasonable understanding of the whole process.

The mechanism

Consider first the behaviour of linear perturbations in the scalar field, $\delta\phi$. If we decompose the perturbations into their Fourier modes $\widetilde{\delta\phi}_k$ and perturb the inflaton equation (2.7), it is easy to show that $\widetilde{\delta\phi}_k \propto \exp[\pm i(k/a)t]$ for $k^{-1} \ll H^{-1}/a$, and that $\widetilde{\delta\phi}_k \to$ *constant* for $k^{-1} \gg H^{-1}/a$. This means that perturbations on scales smaller than the Hubble length will oscillate stably, while those on scales outside the Hubble radius will soon approach a constant, or be 'frozen in'. This is largely due to the loss of causality which is confined within the Hubble radius. In addition, we know that during inflation the comoving Hubble length H^{-1}/a is decreasing, so some

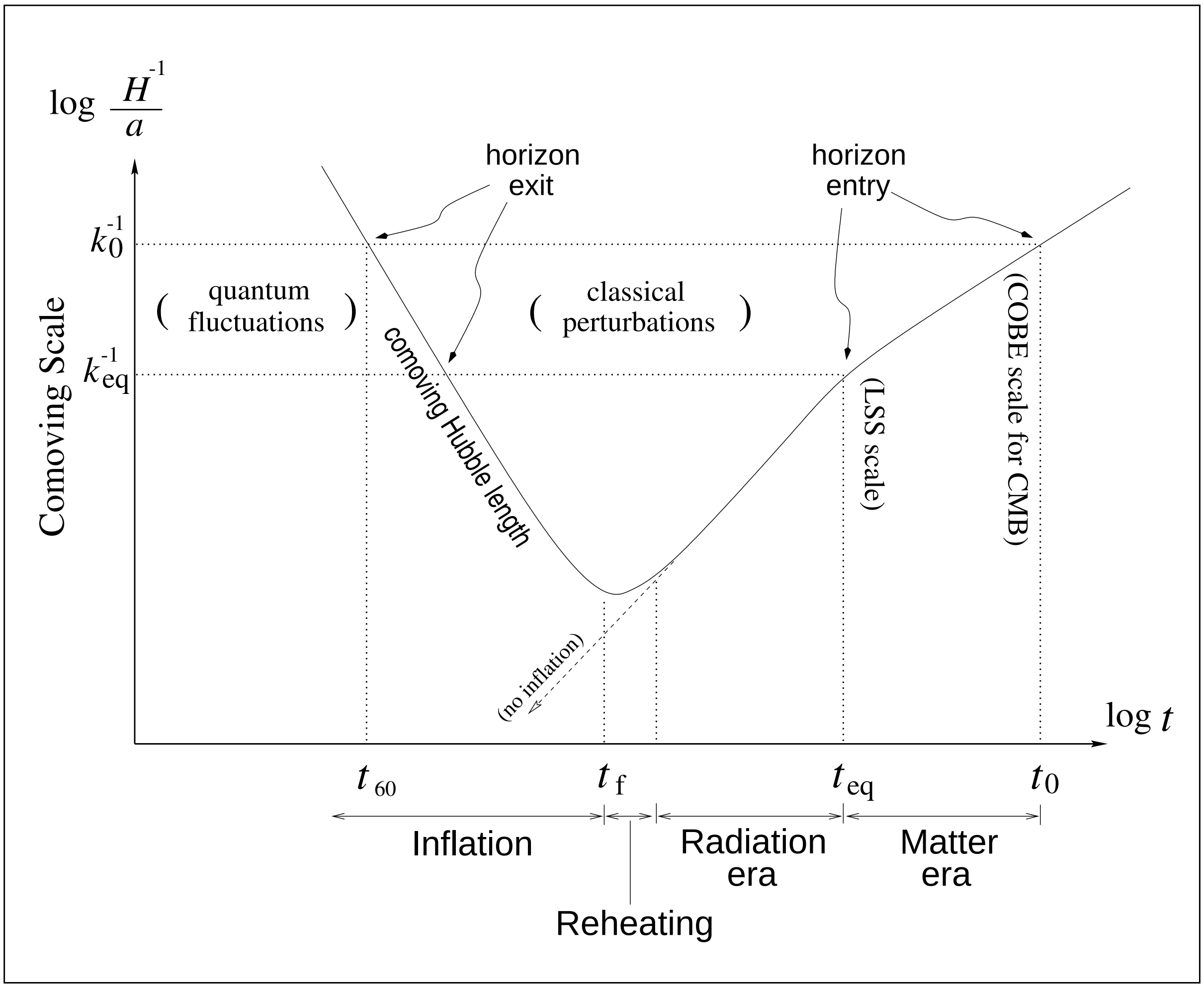

Figure 2.2: The evolution of comoving Hubble length H^{-1}/a in both inflationary and post-inflationary regimes, and its relation with different lengthscales (adapted from [39]). For a given scale, two horizon crossing epochs exist: the 'horizon exit', when the comoving Hubble length decreases crossing through the underlying scale during inflation, and the 'horizon entry', when the comoving Hubble length increases crossing through the same scale for a second time in the post-inflationary regime.

wave modes which are initially inside the Hubble radius will eventually cross out. We call this horizon crossing epoch the 'horizon exit', and will denote it by a subscript '*'. Thus for a given scale k^{-1} we have $k = H_* a_*$ at the horizon exit. Adding the fact that the Uncertainty Principle will produce inevitable quantum fluctuations, we see that quantum fluctuations which are initially oscillating on scales inside the Hubble radius will suddenly become causally unconnected after the horizon exit, and therefore the perturbations they induce in the scalar field will be 'frozen in'. When inflation ends and rejoins the standard big bang evolution, the Hubble length will increase again, so the fluctuations 'frozen' at the horizon exit $k^{-1} = H_*^{-1}/a_*$ during inflation will finally re-enter the horizon. We refer this second horizon crossing epoch as the 'horizon entry'. These returned fluctuations at the horizon entry will then serve as the initial perturbations which will evolve via causal physics. Thus we have a mechanism to generate primordial perturbations during inflation, and Figure 2.2 shows the process. When considering perturbations on a specific scale k^{-1}, we know that they were seeded at the horizon exit $H_*^{-1}/a_* = k^{-1}$ during inflation. Hence there exists a relationship between a given scale k^{-1}, and an epoch during inflation at which $a_* H_* = k$. A normalized expression for this relationship is through the number of e-foldings [37]:

$$N_* \equiv N(t_*, t_{\rm f}) = 62 - \ln \frac{k}{a_0 H_0} - \text{small corrections}, \tag{2.33}$$

where $t_* \equiv t(\phi_*)$ is the epoch corresponding to the scale k, and $t_{\rm f}$ indicates the end of inflation. Since $t_{\rm f}$ is fixed by the underlying inflationary model, we see that a larger scale (smaller k) corresponds to a larger N_* and thus an earlier epoch t_*, which can be determined by equation (2.9). The largest scale on which we can observe the inflationary perturbations today will be at the present Hubble length, i.e. at $k = a_0 H_0$, which leads to $N_* \approx 60$.

We shall use a subscript '60' to denote this scale and the corresponded horizon exit epoch. As we shall see, this is an important number via which we can normalise inflationary perturbations to the COBE result.

Modelling inflationary structure formation

Returning back to the theme of the section, there are two independent components in the inflationary perturbations, depending on their transformation properties: the scalar perturbations (density perturbations), and the tensor perturbations (gravitational waves). Invoking PSRA to the first order, their amplitudes can be formulated respectively as [49, 50]:

$$A_S(k) \equiv \delta_H(k) \simeq \sqrt{\frac{512\pi G^3}{75}\frac{V^{3/2}}{|V'|}}\Bigg|_* = \sqrt{\frac{32}{75}\frac{GV^{1/2}}{\epsilon_V^{1/2}}}\Bigg|_* , \tag{2.34}$$

$$A_T(k) \simeq \sqrt{\frac{32}{75}}GV_*^{1/2}, \tag{2.35}$$

where again the subscript denotes the horizon exit epoch at which $k = a_* H_*$. For the convenience of analytic work, the scale dependence of these amplitudes are normally approximated by power laws:

$$\delta_H^2(k) \propto k^{n-1}, \tag{2.36}$$

$$A_T^2(k) \propto k^{n_T}. \tag{2.37}$$

Thus the amplitudes δ_H and A_T can be related to the usual observables as follows:

1. LSS: The spectrum $S(k)$ defined in equation (1.41) is in fact

$$S(k) = \delta_H^2(k)T^2(k)\left(\frac{k}{H_0}\right)^4, \tag{2.38}$$

where $T(k)$ is the 'transfer function'. This transfer function determines the relative difference of perturbation evolution on different scales, or say, it 'transfers' the initial spectrum shape to its present form. In technical terms, we first write the Fourier transform of the density contrast as a function of both scales and time, i.e. $\tilde{\delta}(k,t)$, then the transfer function can be defined as

$$T(k) = \frac{\tilde{\delta}(k,t_0)\tilde{\delta}(0,0)}{\tilde{\delta}(k,0)\tilde{\delta}(0,t_0)} \tag{2.39}$$

This function is model dependent, i.e. depends on the choice of cosmological parameters. We also notice that on the present horizon scale $k_{H_0} = a_0 H_0$, the transfer function is approximately unity in inflationary cosmology, therefore equation (1.41) gives $S(k_{H_0}) = \delta_H^2(k)$.

2. CMB: Consider contributions to the CMB spectrum from both the scalar and tensor modes, i.e. $C_l = C_l^S + C_L^T$. Then in a purely matter-dominated universe, the scalar component in the CMB on the horizon scale is [51]

$$C_l^{\mathrm{S}} = \frac{\pi}{2}\left[\frac{\sqrt{\pi}}{2}l(l+1)\frac{\Gamma((3-n)/2)}{\Gamma((4-n)/2)}\frac{\Gamma(l+(n-1)/2)}{\Gamma(l+(5-n)/2)}\right]\frac{\delta_H^2(H_0/2)}{l(l+1)}, \tag{2.40}$$

where we notice that the square bracket is unity if $n = 1$. On the other hand, if $A_T(k)$ is weakly dependent on scales, i.e. $n_T \simeq 0$ in equation (2.37), then the tensor component on the horizon scale is (modified from [52])

$$C_l^{\mathrm{T}} = \frac{25\pi}{9}\left(1+\frac{48\pi^2}{385}\right)\frac{A_T^2 D_l}{l(l+1)} \tag{2.41}$$

where D_l's are constants close to one.

With the help of PSRA, one can show to the lowest order that

$$n(k) - 1 \simeq 2\eta_{V*} - 6\epsilon_{V*}, \tag{2.42}$$

$$n_T(k) \simeq -2\epsilon_{V*}. \tag{2.43}$$

To see the relative importance of the scalar and tensor modes in the CMB, we can divide equation (2.41) by equation (2.40) to yield

$$R_0 = \frac{C_l^{\mathrm{T}}}{C_l^{\mathrm{S}}} \simeq 12.4\frac{A_T^2(k)}{A_S^2(k)} \simeq 12.4\epsilon_{V*}, \tag{2.44}$$

where we have assumed $n \simeq 1$, $n_T \simeq 0$ and matter domination. Since C_l is measured from the last scattering surface at which the universe is not purely matter-dominated, a correction to the above result is required [53]:

$$R = 0.75 R_0 \simeq 9.3 \epsilon_{V*}. \tag{2.45}$$

Thus we see that to get the scalar component from the CMB, the observed spectrum needs to be tightened by a factor of $(1+R)$, i.e. $C_l^S = C_l/(1+R)$ or $\delta_H = \delta_{H(\mathrm{obs})}/\sqrt{1+R}$. Furthermore, if we compare equations (2.43) and (2.45), we find

$$R \simeq -4.7 n_T, \tag{2.46}$$

which is called the 'consistency equation'. It implies a certain relationship between the spectra of density perturbations and gravitational waves. This is a direct result from the fact that both spectra are derived from a single assumption—the underlying inflationary potential $V(\phi)$. Although different inflationary models with different $V(\phi)$ will predict different values for n_T and R, the consistency equation (2.46) should be always obeyed to the lowest order.

So far, we have introduced five main parameters in modelling inflationary structure formation: δ_H, A_T, n, n_T, and R. With equations (2.45) and (2.46), we can reduce them to three parameters only: δ_H, n, and R, all determined from the underlying inflationary potential $V(\phi)$.

Observational tests

As mentioned in section 1.5, the commonly used observational tests for theories of structure formation are via the spectra of LSS and CMB, and the non-Gaussianity tests in both. Here we briefly review the current status of inflationary cosmology in the face of observational tests using these techniques.

1. LSS spectrum: For a given inflationary model, there is always at least one free parameter in the potential $V(\phi)$, specifying the energy level of the inflaton and subsequently determining the overall normalization of perturbations (δ_H and A_T), as well as the scale dependence (n and n_T). In principle a comparison of the predicted spectrum with the observation should constrain this parameter(s). However, with the current accuracy in the observations of LSS, a comparison between (2.38) and galaxy surveys (e.g. [33]) can not give us a robust constraint on δ_H and n. This is also partly because, in the calculation of $S(k)$, there is a cosmological model dependence, arising from different choices of Ω_0, Λ, Ω_B, h, etc., affecting the calculation of the transfer function $T(k)$. Thus all these lead to only a weak constraint $n \approx 1$ with considerable uncertainties.

2. CMB spectrum: In spite of the difficulty in applying the LSS observations to constrain inflationary models, another more direct method is via the CMB spectrum. By using the four-year COBE result [54], we can obtain for inflationary models the CMB normalization on the COBE scale $k_{\mathrm{COBE}} = 7 a_0 H_0$ (modified from [53]):

$$\delta_{H(\mathrm{COBE})} = 7^{(n-1)/2} \times 1.91 \times 10^{-5} \; \frac{\exp[1.01(1-n)]}{\sqrt{1+0.75R_0}}, \tag{2.47}$$

where

$$R_0 = -6.2 n_T \left[1 - \frac{n_T}{2} + (n-1)\right] \tag{2.48}$$

is as defined in equation (2.44) but approximated here as a function of n_T and n. Within the regions

$$|n-1| \leq 0.3, \quad -0.3 \leq n_T \leq 0, \tag{2.49}$$

this fitting function is good to 1.5%. Thus for a given model with a certain $V(\phi)$, the PSRA further yields

$$\delta_{H(\mathrm{COBE})} \simeq 1.91 \times 10^{-5} \, (1 - 4.43 \epsilon_{V_{60}} - 0.07 \eta_{V_{60}}). \tag{2.50}$$

A comparison of this result with (2.34) at the horizon exit scale $k = a_{60} H_{60}$ will provide a constraint on the parameter(s) in the potential $V(\phi)$, i.e. we can constrain the parameter via

$$1.91 \times 10^{-5} \, (1 - 4.43 \epsilon_{V_{60}} - 0.07 \eta_{V_{60}}) \simeq \sqrt{\frac{32}{75} \frac{G V^{1/2}}{\epsilon_V^{1/2}}}\Bigg|_{60}. \tag{2.51}$$

If the underlying $V(\phi)$ has only one parameter, then equation (2.51) determines its value. However, most inflationary models have more than one parameter in $V(\phi)$, allowing for different values of δ_H, A_T, n, n_T and R in a single model without violating equation (2.51). In particular, the flexibility in R and n could enable the consistency between the observational normalisations of δ_H from LSS and CMB. Therefore, the

COBE normalization alone, which is currently the most robust information about CMB, is not sufficient to 'test' inflationary models of structure formation, let alone to distinguish them. Nevertheless, there is still one intrinsic feature of inflationary models that may be tested in the CMB—the Doppler peaks. These are oscillatory peaks in the C_l curve, which are evidence of phase coherence in the evolution of perturbations. Such phase coherence would indicate that perturbations are entirely in the growing mode, meaning that they were already in place while well before the last scattering [55]. This is the characteristic of inflationary perturbations, but unlikely to be present in isocurvature models like topological defects [56]. Although current observations have not robustly confirmed this feature, a future evidence in the Doppler peaks would provide strong support for inflationary paradigm. We also notice that there is a dependence of C_l on cosmological models, i.e. the choice of Ω_0, Λ, $\Omega_{\rm B}$, h, etc. Although this dependence will enable one to calibrate cosmological parameters assuming inflation is right (e.g. [22]), it will again inevitably weaken the constraint on inflationary parameters.

3. NG test: Since the origin of inflationary perturbations is from quantum mechanism, their production process must be random, and therefore results in a spatially Gaussian distribution in the final perturbations. This Gaussian property is a crucial feature of inflation, contrary to the topological defect models, which predict non-Gaussian perturbations. However, because the late-time non-linear evolution of perturbations may operate on galaxy scales, the original Gaussianity in inflationary models can hardly be preserved in the LSS. Nevertheless, we still have the CMB, which can provide a good sample reflecting the intrinsic distribution of perturbations. This is because the CMB is like a snapshot of the density distribution at the last-scattering surface, which is long before the non-linear evolution regime begins. Thus if we found significant non-Gaussian signal in the CMB, inflation alone would not be sufficient in explaining the structure of the universe. So far no robust conclusion about the non-Gaussianity in the observed CMB has been drawn, due very much to the observational limit in the resolution and accuracy. We shall concentrate on this issue again in Part IV of the dissertation.

With the COBE result, one can also derive the time scale and energy scale at which inflation has to occur in order to be consistent with the observation. Combining equations (2.31), (2.32), (2.34), and (2.50) gives to the lowest order

$$H_{60}^{-1} \simeq (1.2\times 10^4 t_{\rm Pl}^{-1})\, \epsilon_{V_{60}}^{-1/2} \tag{2.52}$$

$$\simeq (6.4\times 10^{-40}{\rm sec})\, \epsilon_{V_{60}}^{-1/2}. \tag{2.53}$$

With the help of equation (2.32), one also has

$$V_{60}^{1/4} \simeq (5.4\times 10^{-3} m_{\rm Pl})\, \epsilon_{V_{60}}^{1/4} \tag{2.54}$$

$$\simeq (6.5\times 10^{16}{\rm GeV})\, \epsilon_{V_{60}}^{1/4}, \tag{2.55}$$

which will be well below the GUT scale 10^{16} GeV if $\epsilon_{V_{60}} \ll 1$.

Finally, we comment on the future of observational tests on inflationary models of structure formation. As we have seen, because of the flexibility of inflationary models in the face of observational tests, it is very difficult to rule them out. One hope to test the inflationary scenario as a whole will be to verify the Doppler peaks in the CMB spectrum. This test will soon be carried out as an avalanche of future CMB observations lays down a spectrum with higher accuracy and higher resolution. At the same time, the non-Gaussian test in the CMB will also provide another independent test of the paradigm. Having passed both of the tests, one will still need to verify the consistency equation (2.46) (though the result here is valid only to the lowest order), which holds for inflationary models but is unlikely for others. However, the difficulty of this verification will be to quantify the contribution of gravitational waves in the CMB, i.e. to measure R and n_T, and this requires much further progress in detecting gravitational waves. Even if inflation has passed all these tests, there will still remain the challenging task of distinguishing among various inflationary models. This task will be very challenging, due to the flexibility of each model. Nevertheless, instead of using observations to constrain the inflationary parameters in existing models so as to rule out some of them, Liddle has suggested an alternative—use the observations to reconstruct the inflationary potential [57]. Therefore, we see that to verify the inflationary scenario or even to reveal the underlying potential form will require much further progress in the future observations. We expect this to be one of the themes for cosmology in the next few decades of the new century. Moreover, while the mathematical form and the cosmological consequences of inflation have undergone an explosive investigation in the past two decades, its physical origin still needs to be further explored (e.g. [58, 59]).

2.3 Cosmic strings

As previously mentioned, alternative competitors to inflation as theories of structure formation are topological defect models. Among these we shall concentrate on cosmic strings. In this section, we shall begin with a review of cosmological phase transitions, which can produce topological defects. Considering the aim of this dissertation, particular emphasis is then given to cosmic strings, including their modelling, dynamics, gravitational effect, and some basics regarding their network evolution [38].

2.3.1 Phase transitions and topological defects

Elementary particle physics is described at low energies by a gauge theory with three independent coupling constants, all with logarithmic dependencies on energy. These constants correspond to three groups of low-energy symmetry $SU(3)_{\rm QCD} \times SU(2)_{\rm W} \times U(1)_{\rm Y}$. Because above a large energy scale $\eta_{\rm GUT} \sim 10^{16}$GeV these three couplings are approximately the same, there is the expectation that some grand unified theory (GUT) above this scale includes all these three interactions. Typically in particle physics, as the temperature increases, so does the degree of symmetry characterising particle interactions. Adding the fact that in the standard big bang model, the temperature becomes higher when tracing backwards in time, there should exist a high energy regime in which the universe was in a state of maximal symmetry. This maximal symmetry would then be broken as the temperature dropped at later times, resulting in 'spontaneous symmetry breaking' in a chain such as:

$$G \to H \to \ldots \to SU(3)_{\rm QCD} \times SU(2)_{\rm W} \times U(1)_{\rm Y} \to SU(3)_{\rm QCD} \times U(1)_{\rm EM} \,. \tag{2.56}$$

The essential features of spontaneous symmetry breaking can be illustrated in a simple model—the Goldstone model [60]. It has the classical Lagrangian density

$$\mathcal{L} = \partial_\mu \bar{\phi}\, \partial^\mu \phi - V(\phi), \tag{2.57}$$

where ϕ is a complex scalar field and the potential is

$$V(\phi) = \frac{1}{4}\lambda(\bar{\phi}\phi - \eta^2)^2, \tag{2.58}$$

with positive constants λ and η. This model is invariant under the group U(1) of global phase transitions

$$\phi(x) \longrightarrow e^{i\alpha}\phi(x). \tag{2.59}$$

It is called 'global' because α is independent of the spacetime location x. Thus we see that the true vacua lie on a circle $|\phi| = \eta$, implying a non zero expectation value

$$\langle 0|\phi|0\rangle = \eta e^{i\theta}, \tag{2.60}$$

with an arbitrary phase θ. It is apparent that a given vacuum state is not invariant under (2.60), so the symmetry is said to be spontaneously broken.

As in condensed matter systems, cosmological phase transitions can be of first or second order. For simplicity, we consider a second-order phase transition. For the Goldstone model (2.57)–(2.58) thermal corrections can be summed to yield the high-temperature effective potential:

$$V_{\rm eff}(\phi, T) = m^2(T)|\phi|^2 + \frac{\lambda}{4}|\phi|^4, \tag{2.61}$$

where

$$m^2(T) = \frac{\lambda}{12}\left(T^2 - 6\eta^2\right), \tag{2.62}$$

which is the effective mass of the field in the symmetric state, $\langle\phi\rangle = 0$, and we have omitted the ϕ-independent terms in $V_{\rm eff}$. We thus see that there exists a critical temperature

$$T_{\rm c} = \sqrt{6}\eta, \tag{2.63}$$

above which the symmetry is maintained with a global minimum at $\langle|\phi|\rangle^2 = 0$, the false vacuum, because $m^2(T)$ is positive in equation (2.61). On the other hand, when $T < T_{\rm c}$, the effective mass-squared term $m^2(T)$ is negative, so the symmetric state becomes unstable and ϕ develops a non-zero expectation value:

$$\langle|\phi|\rangle^2 = \frac{1}{6}\left(T_{\rm c}^2 - T^2\right), \tag{2.64}$$

which are called the true vacua.

A consequence of spontaneous symmetry breaking and symmetry restoration at high temperature is the inevitable formation of topological defects. Consider again the Goldstone model (2.57) with the Mexican hat potential and a degenerate circle of minima $|\phi| = \eta$. As one goes around a large enough region in space, causality will allow the phase θ to vary. Once it has changed by a multiple of 2π along a closed path, we know the scalar field ϕ must climb up to the top of the potential somewhere inside the closed path; there must be a point where the phase twist is resolved. This is similar to what happens with vortices in water where there must be a vortex-line of zero velocity. The line in our case is called the global string. Its width δ_ϕ is inversely proportional to the boson mass [38]

$$\delta_\phi \sim m_\phi^{-1} = (\sqrt{\lambda}\eta)^{-1}. \tag{2.65}$$

Thus the energy per unit length in the core of size δ_ϕ is given by

$$\mu \sim \delta_\phi^2 V(0) \sim \eta^2. \tag{2.66}$$

For GUT strings with $\eta \sim 10^{16}$ GeV this corresponds to an enormous mass density

$$\mu \sim 10^{-6} m_{\rm Pl}^2 \Longleftrightarrow G\mu \sim 10^{-6}, \tag{2.67}$$

which suggests the significance of GUT strings in cosmology.

During a phase transition when $T < T_{\rm c}$, the scalar field ϕ will begin to fall into different true vacua. The choice of this vacuum state will depend on random fluctuations in ϕ which can be expected to differ in different regions of space. Therefore, we see that causality will enforce the formation of false vacuum remnants, topological defects such as strings in the model (2.57). In the context of cosmology, the existence of a horizon guarantees that the transition is fast on a scale comparable to the size of the universe so that topological defects can be formed on a cosmological scale. This is called the Kibble Mechanism [61].

Depending on the topology of the vacuum manifold, topological defects can be classified into several categories [61]:

1. **Domain walls**: These are sheet-like defects, which are formed when a discrete symmetry is broken. More precisely, this means that the zeroth homotopy group is non-trivial, i.e. the vacuum manifold has

$$\pi_0(G/H) \neq I. \tag{2.68}$$

2. **Cosmic strings**: these are line-like objects, which are formed when an axial symmetry is broken. This means that the vacuum manifold contains non-contractible one-spheres, i.e. the first homotopy group is non-trivial:

$$\pi_1(G/H) \neq I. \tag{2.69}$$

3. **Monopoles**: These are point-like defects, which are formed when a spherical symmetry is broken. In other words, the vacuum manifold has non-contractible two-spheres, i.e.

$$\pi_2(G/H) \neq I. \tag{2.70}$$

There are also textures, which are formed for example when a global non-abelian symmetry is spontaneously and completely broken. They do not have localised cores, but exist in the form of field gradients. On the other hand, although gauge textures can also exist, they are not interesting because they can be gauged away by cancelling the scalar gradients with the gauge fields [62]. Finally, there are also hybrid types of defects such as domain walls attached to strings, whose details are well reviewed in Ref. [38].

Different types of defects will lead to different cosmological consequences. For example, if cosmic domain walls [63] or local monopoles [64, 65] survived up to the present, they would have dominated the evolution of the universe. Thus the potential candidates left for theories of structure formation are cosmic strings and global defects. In this dissertation, we shall be concentrating on the former only. Moreover, depending on the energy level of formation, strings can have different roles in a cosmological context. The cosmic strings we shall be discussing are those formed at the GUT scale, thereby holding the name 'GUT strings'.

2.3.2 Modelling cosmic strings

Although global symmetry breaking is simple as seen in the previous section, we shall be more interested in the more complicated gauge theories that play a central role in elementary particle physics. The simplest model is the abelian-Higgs model [66] with a Lagrangian given by

$$\mathcal{L} = \bar{\mathcal{D}}_\mu \bar{\phi} \mathcal{D}^\mu \phi - V(\phi) - \frac{1}{4} F_{\mu\nu} F^{\mu\nu}, \tag{2.71}$$

where $\mathcal{D}_\mu = \partial_\mu - ieA_\mu$ is the covariant derivative with A_μ a gauge vector field and e the gauge coupling, and $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is an antisymmetric tensor. This model is invariant under the group $U(1)$ of local gauge transformations

$$\phi(x) \longrightarrow e^{i\alpha(x)}\phi(x), \quad A_\mu(x) \longrightarrow A_\mu(x) + e^{-1}\partial_\mu\alpha(x). \tag{2.72}$$

In analogy to the global symmetry breaking, the true vacua here are at $|\phi| = \eta$, and thus the symmetry is spontaneously broken. In this case the gauge field compensates away from the string core for the variation of the phase of ϕ at large distances, cancelling the logarithmic divergence expected in the case of global strings. These are therefore called 'gauge strings' or 'local strings', and it is the cosmological consequences of these which are to be investigated in this dissertation.

The Nambu action

When the radius of curvature of a local string is much greater than its thickness, one can treat the string as a one-dimensional object. Its world history can then be represented in spacetime by a two-dimensional surface, the worldsheet:

$$x_{\rm s}^\mu = x_{\rm s}^\mu(\zeta^a), \quad a = 0,\ 1, \tag{2.73}$$

where $x_{\rm s}^\mu$ are the spacetime coordinates of the string trajectory and ζ^a are the worldsheet coordinates. Thereby the line element in spacetime is

$$ds^2 = \gamma_{ab} d\zeta^a d\zeta^b, \tag{2.74}$$

where

$$\gamma_{ab} = g_{\mu\nu} x_{{\rm s},a}^{\ \mu} x_{{\rm s},b}^{\ \nu}\ . \tag{2.75}$$

Here $g_{\mu\nu}$ and γ_{ab} are the spacetime and worldsheet metrics respectively. Note that for strings in the abelian-Higgs model, the Nambu action can be derived rigorously on the assumption that the scale of perturbations along the string is much larger than its width. The action of the string is then [38]

$$S = -\mu \int \sqrt{-\gamma} d^2\zeta, \tag{2.76}$$

where γ is the determinant of γ_{ab}. This is proportional to the area swept out by the string, and hence analogous to the action for a relativistic particle

$$S = -m \int ds, \tag{2.77}$$

in which the action is proportional to the length of the worldline of the particle.

2.3.3 String dynamics

To obtain the string energy-momentum tensor and the equations of motion, one varies the Nambu action with respect to $g_{\mu\nu}$ and $x_{\rm s}^\mu$ respectively. For simplicity, we first consider strings in flat spacetime. To fix the freedom in parametrising the worldsheet, one can choose the following gauge conditions:

$$\dot{\mathbf{x}}_{\rm s}^2 + \mathbf{x}_{\rm s}'^2 = 1, \tag{2.78}$$

$$\dot{\mathbf{x}}_{\rm s} \cdot \mathbf{x}_{\rm s}' = 0, \tag{2.79}$$

$$\zeta^0 = t, \tag{2.80}$$

which imply respectively that the energy per unit parameter length ζ^1 is a constant, that the physical velocity is perpendicular to the string, and that the timelike parameter is equal to the physical time t. This allows us to write the string trajectory as a three vector $\mathbf{x}_{\rm s}(\zeta^0 = t, \zeta^1 = \sigma)$. Thus the string energy-momentum tensor and the equation of motion are respectively

$$\Theta^{\mu\nu}(\mathbf{r}, t) = \mu \int d\sigma\, (\dot{x}_{\rm s}^\mu \dot{x}_{\rm s}^\nu - x_{\rm s}'^{\mu} x_{\rm s}'^{\nu})\, \delta^{(3)}(\mathbf{r} - \mathbf{x}_{\rm s}(\sigma, t))\,, \tag{2.81}$$

$$\ddot{\mathbf{x}}_{\rm s} - \mathbf{x}_{\rm s}'' = 0, \tag{2.82}$$

where dots and primes denote derivatives with respect to t and σ respectively. Except for the constraints (2.78), (2.79) and (2.80), solutions to equation (2.82) are trivial, simply given as a linear sum of the left- and right-moving modes

$$\mathbf{x}_{\rm s}(\sigma, t) = \frac{1}{2}[\mathbf{a}(\sigma - t) + \mathbf{b}(\sigma + t)], \tag{2.83}$$

where the gauge constraints imply

$$\mathbf{a}'^2 = \mathbf{b}'^2 = 1. \tag{2.84}$$

For a closed loop, σ varies in the range $0 \leq \sigma < L$ where L is the invariant length of the loop. It is then easy to show in the centre-of-mass frame of the loop that

$$\mathbf{x}_s(\sigma + L/2, t + L/2) = \mathbf{x}_s(\sigma, t). \tag{2.85}$$

Consequently, the period of the loop oscillation is $T = L/2$. Thus the mean square velocity of segments on the loop is

$$\langle v^2 \rangle = \int_0^T \frac{dt}{T} \int_0^L \frac{d\sigma}{L} \dot{\mathbf{x}}_s^2. \tag{2.86}$$

This can be calculated by integrating by parts and using the gauge conditions (2.78) and (2.79) and yields

$$\langle v^2 \rangle = 0.5\ . \tag{2.87}$$

We now consider strings in an expanding universe described by the FRW metric (1.8). In this case, the gauge conditions (2.78, 2.79, 2.80) can not be satisfied at the same time, so here we choose the gauge conditions to be

$$\dot{\mathbf{x}}_s \cdot \mathbf{x}_s' = 0, \tag{2.88}$$

$$\zeta^0 = \eta, \tag{2.89}$$

where dots and primes are derivatives with respect to the conformal time η and the spacelike parameter σ, respectively. The remaining reparametrisation freedom can then be represented in the quantity

$$\epsilon = \left(\frac{\mathbf{x}_s'^2}{1 - \dot{\mathbf{x}}_s^2} \right)^{\frac{1}{2}}. \tag{2.90}$$

This quantity is the energy density per unit σ along the string. We notice that ϵ will be unity with the extra gauge condition (2.78) which we use in flat spacetime, but it will in general decay due to the expansion of the universe. By varying the Nambu action with respect to x_s^μ, we have the equations of motion as [67]

$$\ddot{\mathbf{x}}_s + 2\frac{\dot{a}}{a}(1 - \dot{\mathbf{x}}_s^2)\dot{\mathbf{x}}_s = \epsilon^{-1}(\epsilon^{-1}\mathbf{x}'_s)'\,, \tag{2.91}$$

$$\dot{\epsilon} = -2\frac{\dot{a}}{a}\epsilon\dot{\mathbf{x}}_s^2\,, \tag{2.92}$$

and the string energy-momentum tensor is

$$\Theta^{\mu\nu}(\mathbf{x}, \eta) = \mu \int d\sigma \left(\epsilon \dot{x}_s^\mu \dot{x}_s^\nu - \epsilon^{-1} x_s'^\mu x_s'^\nu \right) \delta^{(3)}(\mathbf{x} - \mathbf{x}_s(\sigma, \eta))\,. \tag{2.93}$$

We note that if a loop is much smaller than the horizon size, then the period of the loop motion will be much smaller than one Hubble time ($T = L/2 \ll H^{-1}$), resulting in

$$\ddot{\mathbf{x}}_s \gg H\dot{\mathbf{x}}_s, \qquad \dot{\epsilon} \ll \epsilon/T. \tag{2.94}$$

This property enables us to treat small loops as if they were in flat space. For long strings, the same could also apply to small-scale oscillations or wiggles.

Finally, it should be emphasised that although the Nambu action provides an accurate description of string dynamics, it breaks down when string elements intersect each other, violating the zero width approximation near the crossing point. Therefore proper non-linear field theory is needed in order to investigate these sort of events. For the abelian-Higgs model, numerical simulations show that strings 'exchange partners' or 'reconnect' [68]. Although there are some complicated models in which strings do not reconnect [38], in this dissertation we shall only consider strings that do reconnect.

2.3.4 String gravity

Although GUT strings are extremely massive, their gravitational effects are governed by the parameter $G\mu$, which is very small as seen in (2.67). For $G\mu \ll 1$, it is sufficient to use linearised gravity to investigate string-induced metric perturbations. In a weak gravitational field, the spacetime metric is very close to the Minkowski metric $\eta_{\mu\nu}$, i.e.

$$g^{\mu\nu} = \eta^{\mu\nu} + h^{\mu\nu}, \qquad |h^{\mu\nu}| \ll 1. \tag{2.95}$$

With the choice of the harmonic gauge

$$\partial_\nu(h^\nu{}_\mu - \frac{1}{2}\delta^\nu_\mu h^\lambda{}_\lambda) = 0\,, \tag{2.96}$$

the linearised Einstein equations become

$$\partial_\lambda\partial^\lambda h^{\mu\nu} = -16\pi G S^{\mu\nu}, \tag{2.97}$$

where

$$S^{\mu\nu} = \Theta^{\mu\nu} - \frac{1}{2}\eta^{\mu\nu}\Theta^\sigma{}_\sigma\,. \tag{2.98}$$

Here $\Theta^{\mu\nu}$ is given by (2.81), and indices are raised and lowered using $\eta_{\mu\nu}$. The retarded solution of the wave equation (2.97) can then be solved as [38]

$$h^{\mu\nu}(\mathbf{r},t) = -4G\mu\int\frac{F^{\mu\nu}(\sigma,\tau)d\sigma}{|\mathbf{r}-\mathbf{x}_\mathrm{s}(\sigma,\tau)|\,(1-\mathbf{n}\cdot\dot{\mathbf{x}}_\mathrm{s}(\sigma,\tau))}, \tag{2.99}$$

where

$$\mathbf{n} = \frac{\mathbf{r}-\mathbf{x}_\mathrm{s}(\sigma,\tau)}{|\mathbf{r}-\mathbf{x}_\mathrm{s}(\sigma,\tau)|}\,, \tag{2.100}$$

$$\tau = t - |\mathbf{r}-\mathbf{x}_\mathrm{s}(\sigma,\tau)|, \tag{2.101}$$

$$F^{\mu\nu} = \dot{x}^\mu_\mathrm{s}\dot{x}^\nu_\mathrm{s} - x'^{\,\mu}_\mathrm{s}x'^{\,\nu}_\mathrm{s} + \eta^{\mu\nu}x'^{\,\sigma}_\mathrm{s}x'_{\mathrm{s}\,\sigma}. \tag{2.102}$$

We can solve this in a variety of simple physical contexts including a straight string, a small loop, and a wiggly string.

A straight string

Consider first a static straight string lying along the z-axis with the energy-momentum tensor

$$T^\mu{}_\nu = \mu\delta(x)\delta(y)\mathrm{diag}(1,0,0,1). \tag{2.103}$$

The solution (2.99) can be easily computed and leads to a metric in cylindrical coordinates as

$$ds^2 = dt^2 - dz^2 - (1-8G\mu\ln(r/r_0))\left(dr^2 + r^2d\theta^2\right). \tag{2.104}$$

By the change of coordinates

$$[1-8G\mu]r'^2 = [1-8G\mu\ln(r/r_0)]r^2, \quad \theta' = (1-4G\mu)\theta\,, \tag{2.105}$$

the metric (2.104) becomes

$$ds^2 = dt^2 - dz^2 - dr'^2 - r'^2d\theta'^2, \tag{2.106}$$

which appears to be the same as Minkowski space. However, here we notice from (2.105) that when θ changes by 2π, θ' only changes by $2\pi(1-4G\mu)$. Thus this azimuthal 'deficit angle' $\Delta = 8\pi G\mu$ implies that a surface of constant t and z has the geometry of a cone rather than that of a plane. This conical spacetime around a cosmic string is illustrated in Figure 2.3.

Moreover, by changing the coordinates back to the original cartesian coordinates and performing a Lorentz transformation from the string rest frame to the rest frame of a test particle, it can be shown that when a test particle passes the string, there will be a velocity gain perpendicular to the string plane of motion and towards that plane. This velocity gain is given by $\Delta v_\mathrm{s} = 4\pi G\mu\gamma_\mathrm{s}v_\mathrm{s}$, where v_s is the string velocity with respect to the particle and $\gamma_\mathrm{s} = (1-{v_\mathrm{s}}^2)^{-1/2}$ [69].

A small loop

Next we consider the gravitational effect of a small loop on non-relativistic test particles. To see this effect, we take the time-averaged field induced by the loop

$$\langle h^{\mu\nu}(\mathbf{r})\rangle = \frac{1}{T}\int_0^T dt\, h^{\mu\nu}(\mathbf{r},t), \tag{2.107}$$

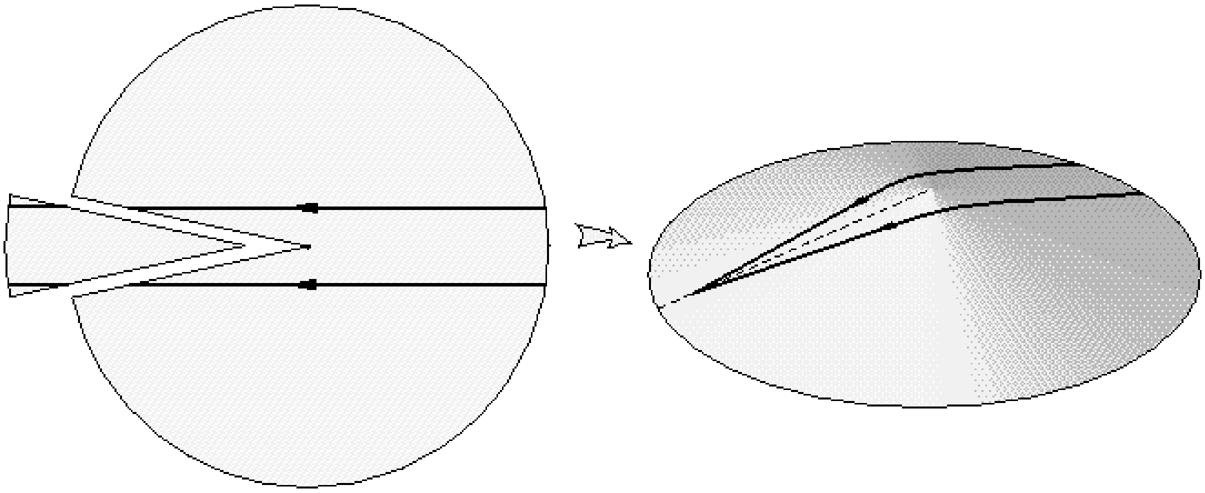

Figure 2.3: Conical spacetime around a cosmic string. We can see the geodesic paths of light curved towards a string when light passing by it.

where T is the oscillation period of the loop. If we change the integration variable to retarded time $dt = d\tau[1 - \mathbf{n}\cdot\dot{\mathbf{x}}_{\rm s}]$ substitute $T = L/2$, and invoke the solution (2.99), then the average field at a large distance from the loop (i.e. $|\mathbf{r} - \mathbf{x}_{\rm s}| \gg L$) becomes

$$\langle h^{\mu\nu}(\mathbf{r})\rangle = -\frac{8G\mu}{Lr}\int_0^{L/2} d\tau \int_0^L d\sigma F^{\mu\nu}(\sigma,\tau). \tag{2.108}$$

Using the definition (2.102) and the result (2.87), this integration results in

$$\langle h^{\mu\nu}(\mathbf{r})\rangle = -\frac{2G\mu L}{r}\delta^{\mu\nu}. \tag{2.109}$$

Thus we see that the average field of a small loop from a distance coincides with that of a Schwarzschild field of mass $M = \mu L$ [70].

A wiggly string

From a distance, the kinks and wiggles on a long string can not be resolved, so the string will appear smooth with an effective mass per unit length $\tilde{\mu}$ and tension $\tilde{T}$, which are different from those of long straight strings. The energy-momentum tensor of a wiggly string lying along the z-axis can be approximated in the form [71, 72]

$$T^{\mu}{}_{\nu} = \delta(x)\delta(y)\mathrm{diag}(\tilde{\mu}, 0, 0, \tilde{T}), \tag{2.110}$$

with $\tilde{\mu}\tilde{T} = \mu^2$. Then the solution (2.99) leads to

$$ds^2 = (1 + h_{00})\left(dt^2 - dz^2 - dr^2 - (1 - 8G\tilde{\mu})r^2 d\theta^2\right), \tag{2.111}$$

with $h_{00} = 4G(\tilde{\mu} - \tilde{T})\ln(r/r_0)$. Moreover, when a particle passes the wiggly string, the velocity gain in the direction perpendicular to the initial motion can be shown to be [73]

$$\Delta v_{\rm s} = 4\pi G\tilde{\mu}\gamma_{\rm s}v_{\rm s} + \frac{2\pi G(\tilde{\mu} - \tilde{T})}{v_{\rm s}\gamma_{\rm s}}. \tag{2.112}$$

The first term is again due to the deficit angle as in the case of a straight string, while the second term is due to the attractive gravitational force from the wiggles.

2.3.5 String evolution

Cosmic strings are formed through the Kibble mechanism during a symmetry breaking phase transition in the early universe. This formation of strings can be simulated approximately by randomly assigning values to the scalar field ϕ on a lattice with grid spacing of order of the string correlation length ξ. For a cubic lattice, it was found that about 80% of the total string length is in the form of infinite Brownian strings and only 20% is in the form of closed loops [73]. To characterise the shapes of individual strings in a network, one can express the average length of string L between two points separated by a distance R on a string as

$$L = A\xi(R/\xi)^d, \tag{2.113}$$

where d is the fractal dimension. Using boxes of side-length 40ξ, it was found that $d = 2.00 \pm 0.07$ and $A = 0.97 \pm 0.03$ for infinite strings.

After formation, strings experience a significant damping force from the relatively high radiation background density. However, for GUT strings this frictional force resulting from particle scattering will become negligible at later times when $t \gtrsim (G\mu)^{-2} t_{\mathrm{Pl}}$, and it should not affect the properties of the string network which we shall use for structure formation [74]. In this dissertation, we shall assume that the strings oscillate freely.

In an expanding universe, there are two factors which dominate the evolution of a string network. The first is the breaking of the network into smaller and smaller pieces, in order to achieve thermodynamic equilibrium. The second is the expansion of the universe and causality, which tends to delay and limit this process. As a result, the evolution reaches a quasi-static equilibrium state after some time, in which the two factors are balanced. In this regime, the long-string network properties are essentially the same at all times when scaled to the horizon size, so we refer to this as the 'scaling regime'. This scaling behaviour is observed in simulations with reasonable initial conditions, and we now investigate it in more detail.

At first the string energy within a large physical volume V can be written as

$$E = a(t)\mu \int_V \epsilon d\sigma. \tag{2.114}$$

Differentiating both sides of this with respect to physical time t yields

$$\dot{E} + H(2\langle v^2\rangle - 1)E = 0, \tag{2.115}$$

where $\langle v^2\rangle = \int \epsilon \dot{\mathbf{x}}_s^2 d\sigma / \int \epsilon d\sigma$. Thus we see how the expansion of the universe can affect the energy of strings. To see the evolution of long strings, the dominant part of the network, we further consider its contribution in E. For long strings, the network is Brownian on scales much larger than the correlation length ξ, so there is approximately one string segment of length ξ per correlation volume ξ^3. This allows us to write the energy density of long strings as

$$\rho_\infty = \frac{\mu}{\xi^2}. \tag{2.116}$$

A string segment of size ξ can then be expected to encounter another segment after travelling a distance ξ, so we expect the rate of loop formation to be proportional to ξ^{-1}/ξ^3 per unit volume. As a consequence, equation (2.115) can be modified for the long-string energy E_∞ as

$$\dot{E}_\infty + H(2\langle v^2\rangle - 1)E_\infty = -c_l(\mu\xi)\xi^{-4}V, \tag{2.117}$$

where c_l is a constant. Dividing both sides by V and using equation (2.116) gives

$$\frac{d\rho_\infty}{dt} + 2H(1 + \langle v^2\rangle)\rho_\infty = -c_l\frac{\rho_\infty}{\xi}. \tag{2.118}$$

If we define $\zeta = \xi/t$, then we have from the above equation

$$\frac{1}{\zeta}\frac{d\zeta}{dt} = -\frac{1}{2t}\left(1 - \langle v^2\rangle - \frac{c_l}{\zeta}\right), \tag{2.119}$$

where we have used $a(t) \propto t^{1/2}$ in the radiation era. If $\zeta > c_l/(1 - \langle v^2\rangle)$ then the right hand side is negative so that ζ decreases. If $\zeta < c_l/(1 - \langle v^2\rangle)$ then the right hand side is positive so that ζ will grow. Consequently, ζ converges to

$$\zeta_{\mathrm{av}} = \frac{c_l}{1 - \langle v^2\rangle} \tag{2.120}$$

over time, meaning that the long string correlation length $\xi = \zeta t$ will approach to a fixed fraction of the horizon size. This is called the 'scaling regime'. Substituting ζ_{av} back into equation (2.118), we obtain in the scaling regime an evolution equation of long-string energy:

$$\frac{d\rho_\infty}{dt} = -2\frac{\rho_\infty}{t}. \tag{2.121}$$

We note that although we have assumed radiation domination in obtaining this result, one would find this result unchanged if a matter-dominated regime was assumed. It is then straightforward to show that

$$\rho_\infty \propto t^{-2} \propto \begin{cases} a^{-4} & \text{in radiation era,} \\ a^{-3} & \text{in matter era.} \end{cases} \tag{2.122}$$

Compared with equation (1.23), we see that in the scaling regime the long-string energy density evolves in the same way as the mean energy density of the universe. Although we have assumed $\langle v^2 \rangle$ to be a constant, more elaborate analysis has also shown the same scaling behaviour [75].

In the above analysis, we have only one fundamental scale which characterises the long-string network properties—the correlation length ξ. Thus it is called the one-scale model. This one-scale model has ignored the information on small scales like the 'kinks' or 'wiggles', and is valid only from a statistical point of view. Therefore we shall need simulations to study more details of the string network. So far, there are three independent groups which have successfully simulated string evolution using different numerical techniques. They are Albrecht & Turok (AT) [76, 77], Bennett & Bouchet [78, 79] (BB) and Allen & Shellard [80] (AS). They all agree qualitatively, and in particular quantitatively in the last two groups. In this dissertation we shall use the AS simulations [80] to study the cosmological perturbations induced by cosmic strings, some of whose consequences can then be generalised to understand cosmological implications from other defects.

Part II

PERTURBATION THEORY FOR CAUSAL MODELS

Chapter 3

Background Cosmologies and Causal Perturbations

3.1 Introduction

In Chapter 1 we introduced the basics of the standard cosmology and the problems it faces, among which we highlighted the lack of a mechanism to produce cosmological perturbations. In order to compensate for this flaw in the standard model, we described two paradigms for structure formation in Chapter 2—inflation and topological defects. Among the cosmic defect theories, we shall concentrate on cosmic strings in this dissertation, although many of the results and methods we shall present can still be applied to the other cases. To lay down some foundations for studying cosmological perturbations, we investigate linear perturbation theory in the presence of causal seeds in this chapter. We shall study solutions for matter perturbations in the next chapter, and those for radiation perturbations in Chapter 9. The formalism developed here is parallel to those presented in Ref. [81] and Ref. [82], but we give a more detailed treatment which includes the effects of baryons and photon decoupling. We also note that although in the literature, there have been detailed treatments of theories with causal seeds [83, 84], the formalism and its analytic solutions which we shall present will provide not only a physically transparent way for understanding the evolution of density perturbations, but also a computationally economical method which is particularly pertinent when one needs to investigate the phase information of the resulting cosmological perturbations.

The structure of this chapter is as follows. In section 3.2, we investigate solutions for the dynamics of various background cosmologies, including the consideration of non-zero curvature and a cosmological constant. In section 3.3, with both the fluid approximation and the baryon-photon tight-coupling approximation, we derive perturbation equations in the synchronous gauge sourced by causal seeds. Then, in section 3.4, we verify the accuracy of these methods in the standard CDM model, by comparing directly with the results from CMBFAST [85], a fast Einstein-Boltzmann solver. We conclude in section 3.5 that for reasonable parameter ranges, this approach provides satisfactory precision at greatly reduced numerical cost.

3.2 Background cosmologies

With the discovery of the CMBR in 1964 [14], the universe is believed to be mainly composed of not only matter but also radiation. After the discovery, several authors worked out the solutions in some FRW models with both radiation and matter [86, 87, 88, 89]. In this section, we aim to derive the general solution of FRW models, in the presence of both curvature and a cosmological constant.

We assume that the universe is homogeneous and isotropic, and is filled with two fluids, radiation and dark matter, whose stress-energy tensors are also homogeneous and isotropic on average. We shall ignore the overall contribution of the stress energy from causal seeds like defect fields, which is always much smaller than the total energy density of radiation and matter. In a FRW universe with only radiation and matter components which evolve independently and adiabatically, the scale factor $a(\eta)$ is then determined by the unperturbed Einstein equation, which can be rewritten from the Friedmann equation (1.20) as

$$\dot{a}^2 + Ka^2 = \frac{8\pi G\rho_{\mathrm{m}0}a_0^3}{3}(1+a) + \frac{\Lambda}{3}a^4, \tag{3.1}$$

where a dot represents a derivative with respect to the conformal time η, K is the curvature, ρ_{m} is the matter energy density, Λ is the cosmological constant, and we have normalized $a_{\rm eq} = 1$. If we define

$$\Omega_{\rm m} = \frac{8\pi G\rho_{\rm m}}{3H^2}, \tag{3.2}$$
$$\Omega_{\rm r} = \frac{8\pi G\rho_{\rm r}}{3H^2} = \frac{8\pi G\rho_{\rm m}}{3aH^2}, \tag{3.3}$$
$$\Omega_\Lambda = \frac{\Lambda}{3H^2}, \tag{3.4}$$
$$\Omega_K = -\frac{K}{a^2H^2}, \tag{3.5}$$

where $H = \dot{a}/a^2$ is the Hubble parameter, then we have from (3.1) that $\Omega_{\rm m} + \Omega_{\rm r} + \Omega_\Lambda + \Omega_K = 1$ and

$$\frac{\Omega_{\Lambda 0}}{\Omega_{\rm m0}} = \frac{\Lambda}{8\pi G\rho_{\rm m0}}, \quad \frac{\Omega_{K0}}{\Omega_{\rm m0}} = \frac{-3K}{8\pi G\rho_{\rm m0}a_0^2}. \tag{3.6}$$

We notice also that $\Omega_{\rm r0}/\Omega_{\rm m0} = a_0^{-1} \ll 1$. We define

$$A = \frac{2(\sqrt{2}-1)}{\eta_{\rm eq}}, \quad B = \frac{\Omega_{K0}}{\Omega_{\rm m0}a_0}, \quad C = \frac{\Omega_{\Lambda 0}}{\Omega_{\rm m0}a_0^3}, \tag{3.7}$$

where we note that $B, C \ll 1$ due to $a_0 \gg 1$ and $\Omega_{\rm m0} \sim \Omega_{K0} \sim \Omega_{\Lambda 0}$ according to the current observational results. Thus we can rewrite equation (3.1) as

$$\left(\frac{da}{d\eta}\right)^2 = \bar{A}^2(1 + a + Ba^2 + Ca^4), \tag{3.8}$$

where

$$\bar{A} = \frac{1}{\eta_{\rm eq}}\int_0^1 \frac{da}{(1 + a + Ba^2 + Ca^4)^{1/2}} \approx A. \tag{3.9}$$

Equation (3.8) can then be numerically evaluated with certain choices of $\Omega_{\rm m0}$, $\Omega_{\Lambda 0}$ and Ω_{K0}. Assuming three species of neutrinos and using $\rho_{\gamma 0} = 2.0747 \times 10^{-51}\text{GeV}^4$ [13] and the fact that at $\eta_{\rm eq}$ both the curvature and the cosmological constant terms are negligible in (3.1), we obtain

$$a_0 = 23219\ \Omega_{\rm m0}h^2, \tag{3.10}$$
$$\eta_{\rm eq} = 16.310\ (\Omega_{\rm m0}h^2)^{-1}\text{Mpc}, \tag{3.11}$$
$$t_{\rm eq} = 3.4058 \times 10^{10}(\Omega_{\rm m0}h^2)^{-2}\text{sec}, \tag{3.12}$$

where $\eta_{\rm eq}$ is in the units measured today. In certain cases, (3.8) can be exactly solved:

1. $K = \Lambda = 0$ (i.e. $\Omega_{\rm m0} = 1, \Omega_{\Lambda 0} = 0$):

$$a(\eta) = A^2\eta^2/4 + A\eta, \tag{3.13}$$
$$t(\eta) = A^2\eta^3/12 + A\eta^2/2, \tag{3.14}$$

   which give $\eta_{\rm eq} = 3t_{\rm eq}/\sqrt{2}$.

2. $K < 0,\ \Lambda = 0$ (i.e. $\Omega_{\rm m0} < 1, \Omega_{\Lambda 0} = 0$):

$$a(\eta) = \frac{1}{2B}\left[\cosh(\bar{A}\sqrt{B}\eta) + 2\sqrt{B}\sinh(\bar{A}\sqrt{B}\eta) - 1\right], \tag{3.15}$$
$$t(\eta) = \frac{1}{\bar{A}B}\left[\cosh(\bar{A}\sqrt{B}\eta) + \frac{1}{2\sqrt{B}}\sinh(\bar{A}\sqrt{B}\eta) - \frac{\bar{A}\eta}{2} - 1\right]. \tag{3.16}$$

3. $K > 0,\ \Lambda = 0$ (i.e. $\Omega_{\rm m0} > 1, \Omega_{\Lambda 0} = 0$):

$$a(\eta) = \frac{1}{2B}\left[\cos(\bar{A}\sqrt{-B}\eta) - 2\sqrt{-B}\sin(\bar{A}\sqrt{-B}\eta) - 1\right], \tag{3.17}$$
$$t(\eta) = \frac{1}{\bar{A}B}\left[\cos(\bar{A}\sqrt{-B}\eta) + \frac{1}{2\sqrt{-B}}\sin(\bar{A}\sqrt{-B}\eta) - \frac{\bar{A}\eta}{2} - 1\right]. \tag{3.18}$$

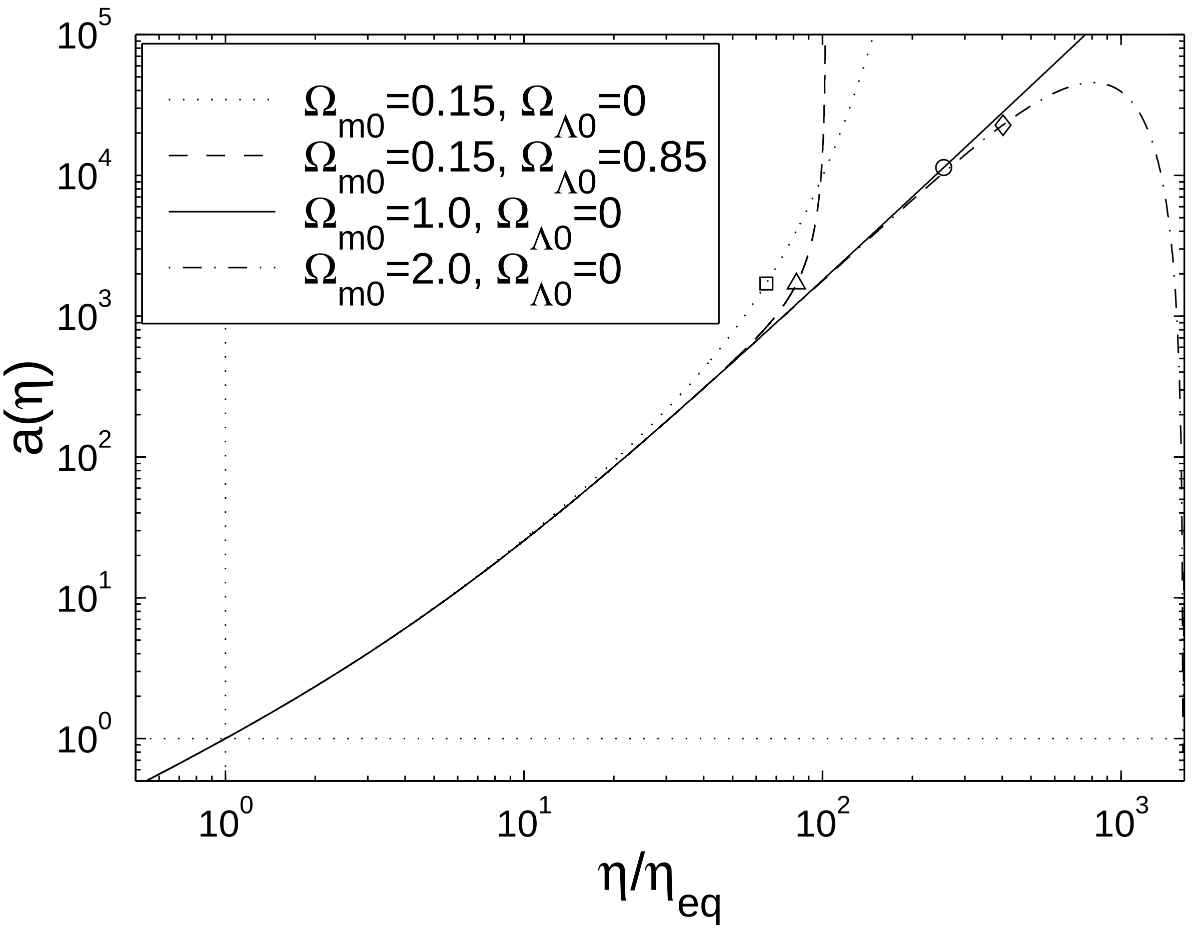

Figure 3.1: The evolution of background dynamics in various cosmologies. Plotted are exact solutions of the scale factor $a(\eta)$. The square, triangle, circle and diamond mark the universe today for different models, each with $H_0 = 70$ km s^{-1}Mpc^{-1}.

We notice that at early times equations (3.15–3.16) and (3.17–3.18) reduce to equations (3.13–3.14). At late times equations (3.13–3.14), (3.15–3.16) and (3.17–3.18) give the asymptotic forms

$$a(\eta) \propto \begin{cases} \eta^2, & K = \Lambda = 0, \\ \exp(\bar{A}\sqrt{B}\eta), & K < 0,\ \Lambda = 0, \\ 1 - \cos(\bar{A}\sqrt{-B}\eta), & K > 0,\ \Lambda = 0, \end{cases} \tag{3.19}$$

or

$$a(t) \propto \begin{cases} t^{2/3}, & K = \Lambda = 0, \\ t, & K < 0,\ \Lambda = 0, \\ 1 - \cos[2\bar{A}(-B)^{3/2}t], & K > 0,\ \Lambda = 0. \end{cases} \tag{3.20}$$

Figure 3.1 shows some examples of these solutions. As one can see, the destiny of the universe diverges in the different cosmologies, although all have identical features around or before the radiation-matter equality $t_{\rm eq}$. This converging behaviour at early times will help simplify the calculation of cosmological perturbations, which we know are seeded mainly from the radiation-matter transition era. We shall further discuss this point later.

3.3 Synchronous gauge perturbation theory

In this section, we derive the evolution equations for cosmological perturbations. To calculate the density and metric perturbations, we model the contents of the universe as perfect fluids: radiation (photons and neutrinos) and pressureless matter (CDM and baryons). We shall use the photon-baryon tight-coupling approximation until the epoch of last scattering, at which we assume instantaneous decoupling, also taking into account the effect of Silk damping due to the photon diffusion at that time. After decoupling, the baryonic perturbations originating from the radiation perturbations are then merged linearly into the CDM content. In a scenario that includes the presence of causal seeds, the radiation and matter fields are assumed to be initially uniform, and then perturbed by the causal seeds which induce density fluctuations inside the horizon. The radiation, matter, and causal seeds are assumed to interact only through gravity, meaning that their stress-energy tensors are separately covariantly conserved.

We shall work in the synchronous gauge, in which the perturbations $h_{\mu\nu}$ to the spacetime metric obey the constraint $h_{0\mu} = 0$. The perturbed flat FRW metric is then given by

$$g_{00} = a^2(\eta), \quad g_{ij} = -a^2(\eta)[\delta_{ij} + h_{ij}(\eta, \mathbf{x})], \tag{3.21}$$

where $|h_{ij}| \ll 1$ in the linear theory, which we shall assume. Although the synchronous gauge is sometimes criticised in the literature due to its residual gauge freedom, this gauge choice is well suited to the defect models discussed here. In these models, the universe is assumed to be perfectly homogeneous and isotropic before the symmetry breaking phase transition. Hence all perturbation variables can be set to zero, and this is normally referred to as the 'initially unperturbed synchronous gauge' (IUSG) [81], which possesses no residual gauge freedom. Thus the Einstein equations are completely causal in IUSG, with the value of all perturbation variables at a given spacetime point being completely determined by initial conditions within the past light cone of the point.

In section 3.3.1, we derive in the IUSG the conservation equations of radiation and matter fields. In section 3.3.2, we consider the conservation of source stress energy. The linearly perturbed Einstein equations are derived in section 3.3.3. We then turn to investigate the internal energy transfer among various fields by considering the stress-energy pseudotensor in section 3.3.4. The treatment of the instantaneous decoupling approximation will be given in section 3.3.6. The formalism derived in this section will then be verified for the standard CDM model in section 3.4, in comparison with the results from CMBFAST [85], a fast Einstein-Boltzmann solver.

3.3.1 Stress-energy conservation of radiation and matter fields

The radiation and matter fields are considered as perfect fluids, whose energy-momentum tensors have the form

$$T^{\mu}_{N\nu} = (\rho_N + p_N)u^{\mu}_N u_{N\nu} + p_N \delta^{\mu}_{\nu}, \quad u^{\mu}_N u_{N\nu} = -1, \tag{3.22}$$

where ρ_N, p_N, and u^{μ}_N are respectively the density, pressure, and four-velocity of the Nth fluid. In the homogeneous background, we have $u^{\mu}_N = (a^{-1}, \vec{0})$, which implies that $\delta u^0_N = 0$ to first order. We thus define the velocity perturbation as $v^i_N = a\delta u^i_N$. The equation of state and the sound speed are defined respectively as

$$\kappa_N = \frac{p_N}{\rho_N}, \quad c^2_N = \frac{\delta p_N}{\delta \rho_N}. \tag{3.23}$$

Consequently, the covariant conservation of stress energy for each fluid $T^{\mu\nu}_{N;\nu} = 0$ gives [81]

$$\dot{\delta}_N + (1+\kappa_N)(\nabla\cdot\mathbf{v}_N + \frac{1}{2}\dot{h}) + 3\frac{\dot{a}}{a}(c^2_N - \kappa_N)\delta_N = 0, \tag{3.24}$$

$$\dot{\mathbf{v}}_N + \frac{\dot{a}}{a}(1-3c^2_N)\mathbf{v}_N + \frac{c^2_N}{1+\kappa_N}\nabla\delta_N = 0, \tag{3.25}$$

$$\dot{\mathbf{v}}^{\perp}_N + \frac{\dot{a}}{a}(1-3c^2_N)\mathbf{v}^{\perp}_N = 0, \tag{3.26}$$

where $\delta_N = \delta\rho_N/\rho_N$, and we have decomposed the velocities as $\mathbf{v}_N = \mathbf{v}^{\parallel}_N + \mathbf{v}^{\perp}_N$, with $\nabla\times\mathbf{v}^{\parallel}_N = 0$ and $\nabla\cdot\mathbf{v}^{\perp}_N = 0$.

In the regime of tight coupling between photons and baryons, we have two main fluids: the CDM component and the tightly-coupled photon-baryon fluid. They will be denoted as $N = \mathrm{c}, \gamma\mathrm{B}$ respectively, and discussed separately as follows. Note that we have ignored the neutrinos in the radiation.

CDM fluid

We first consider the CDM fluid, i.e. $N = \mathrm{c}$. With $\kappa_\mathrm{c} = c^2_\mathrm{c} = 0$ for pressureless matter, the equations of stress-energy conservation (3.24)–(3.26) become

$$\dot{\delta}_\mathrm{c} + \nabla\cdot\mathbf{v}_\mathrm{c} = -\frac{1}{2}\dot{h}, \quad \dot{\mathbf{v}}_\mathrm{c} + \frac{\dot{a}}{a}\mathbf{v}_\mathrm{c} = 0. \tag{3.27}$$

As we can see, any perturbations in the CDM velocity will decay as a^{-1}. Thus we can simply choose $\mathbf{v}_\mathrm{c} = \vec{0}$ in the IUSG. Once $\mathbf{v}_\mathrm{c} = \vec{0}$, it will remain so as there is no linear gravitational source. As a consequence, the CDM obeys a single nontrivial conservation law resulting from equation (3.27)

$$\dot{h} + 2\dot{\delta}_\mathrm{c} = 0 \Longrightarrow h = -2\delta_\mathrm{c}, \tag{3.28}$$

where the second equation results from the initial condition $h = \delta_c = 0$.

Photon-baryon tightly coupled fluid and its photon component

For the tightly-coupled photon-baryon (γB) fluid, we have

$$\mathbf{v}_{\gamma B} = \mathbf{v}_{\gamma} = \mathbf{v}_{B}, \quad p_{\gamma B} = p_{\gamma}, \quad \rho_{\gamma B} = \rho_{\gamma} + \rho_{B}\,. \tag{3.29}$$

Thus we can define

$$R = \frac{\delta\rho_B}{\delta\rho_\gamma} = \frac{3\rho_B}{4\rho_\gamma}, \tag{3.30}$$

where the second result comes from the fact that $\rho_\gamma \propto a^{-4}$ and $\rho_B \propto a^{-3}$. Using the definitions (3.23), we then have

$$\kappa_{\gamma B} = \frac{1}{3+4R}, \quad c^2_{\gamma B} = \frac{1}{3(1+R)}. \tag{3.31}$$

With these results, the equations of stress-energy conservation for the γB fluid can be obtained from equations (3.24)–(3.26):

$$\dot{\delta}_{\gamma B} + \frac{4+4R}{3+4R}(\nabla\cdot\mathbf{v}_{\gamma B} - \dot{\delta}_c) + \frac{\dot{a}}{a}\frac{R}{(1+R)(3+4R)}\delta_{\gamma B} = 0, \tag{3.32}$$

$$\dot{\mathbf{v}}_{\gamma B} + \frac{\dot{a}}{a}\frac{R}{1+R}\mathbf{v}_{\gamma B} + \frac{3+4R}{12(1+R)^2}\nabla\delta_{\gamma B} = 0, \tag{3.33}$$

$$\dot{\mathbf{v}}^{\perp}_{\gamma B} + \frac{\dot{a}}{a}\frac{R}{1+R}\mathbf{v}^{\perp}_{\gamma B} = 0. \tag{3.34}$$

In cosmological applications, such as CMB anisotropies, we are more interested in the photon perturbations. Therefore by using equations (3.29) and (3.30), we can extract the photon component from the above equations to yield [90]

$$\dot{\delta}_r + \frac{4}{3}\nabla\cdot\mathbf{v}_r - \frac{4}{3}\dot{\delta}_c = 0, \tag{3.35}$$

$$\dot{\mathbf{v}}_r + \frac{\dot{a}}{a}\frac{R}{1+R}\mathbf{v}_r + \frac{1}{4+4R}\nabla\delta_r = 0, \tag{3.36}$$

where we have ignored neutrinos in radiation so as to replace the subscript γ with r. The velocity can then be eliminated to yield a single second-order equation for the radiation perturbations:

$$\ddot{\delta}_r - \frac{4}{3}\ddot{\delta}_c + \frac{\dot{R}}{1+R}(\dot{\delta}_r - \frac{4}{3}\dot{\delta}_c) - \frac{1}{3+3R}\nabla^2\delta_r = 0. \tag{3.37}$$

We note that at any given moment the photon velocities, which are missing from this equation, can be recovered from equation (3.35).

An alternative presentation of equations (3.35) and (3.37) is via the entropy perturbation s. It is defined as the fluctuation in the number of photons per dark matter particle

$$s = \frac{3}{4}\delta_r - \delta_c. \tag{3.38}$$

Thus equations (3.35) and (3.37) can be rewritten as

$$\dot{s} = -\nabla\cdot\mathbf{v}_r, \tag{3.39}$$

$$\ddot{s} = -\frac{\dot{R}}{1+R}\dot{s} + \frac{1}{3+3R}\nabla^2(s+\delta_c)\,. \tag{3.40}$$

As we shall see, δ_r can only have a white noise power spectrum on superhorizon scales. From equation (3.36), this implies a k^2 power spectrum in $\mathbf{v}_r$ on these scales. Adding the fact that the entropy fluctuation s starts from zero on superhorizon scales due to the fixed number of dark matter particles per photon, it then follows from equation (3.39) that both s and $\dot{s}$ have a k^4 fall off outside the horizon. Consequently in numerical simulations, as long as the initial horizon size is smaller than scales of our interest, we can simply set $s = \dot{s} = 0$.

3.3.2 Stress-energy conservation of the source

The source we shall consider is weak, so it will appear only as first-order terms in the perturbed stress-energy tensor. Thus in the linear theory we are considering here, they can be considered as stiff, meaning that their evolution depends only on their own self-interactions and the background dynamics of the universe, but not on their self-gravity or the weak gravitational fields of the inhomogeneities they produce. This assumption will enable us to separate the calculation of their dynamics from that of the associated inhomogeneities, evolving the source as if it is in a completely homogeneous background. Since the source is stiff, its energy-momentum tensor $\Theta_{\mu\nu}$ need only be locally covariantly conserved with respect to the background:

$$\Theta_{00,0} + \frac{\dot{a}}{a}\Theta_{+} = \Theta_{0i,i} \ , \tag{3.41}$$

$$\Theta_{0i,0} + 2\frac{\dot{a}}{a}\Theta_{0i} = \Theta_{ij,j} \ , \tag{3.42}$$

where $\Theta_{+} = \Theta_{00} + \Theta_{ii}$. Another important aspect of structure formation with defect models is the fact that the defects, formed at very early times, will ultimately create under-densities in the initially homogeneous background, out of which they are carved. This is a direct result of energy conservation in the universe, and is normally termed 'compensation'. We shall discuss this issue in more detail later.

3.3.3 Perturbed Einstein equations

If we perturb the ten Einstein equations (1.18) within the photon-baryon tight-coupling approximation, we obtain ten linearised perturbation equations:

$$2\delta R_{00} = -\ddot{h} - \frac{\dot{a}}{a}\dot{h} = +3\left(\frac{\dot{a}}{a}\right)^2 [(2+R)\Omega_{\rm r}\delta_{\rm r} + \Omega_{\rm c}\delta_{\rm c}] + 8\pi G\Theta_{+} \ , \tag{3.43}$$

$$\begin{aligned} 2\left\{\delta\tilde{R}_{ij} - \left[\frac{\ddot{a}}{a} + \left(\frac{\dot{a}}{a}\right)^2\right]\tilde{h}_{ij}\right\} &= \ddot{\tilde{h}}_{ij} + 2\frac{\dot{a}}{a}\dot{\tilde{h}}_{ij} - \nabla^2\tilde{h}_{ij} - \frac{1}{3}h_{,ij} + \frac{1}{9}\delta_{ij}\nabla^2 h \\ &\quad + \tilde{h}_{ik,kj} + \tilde{h}_{jk,ki} - \frac{2}{3}\delta_{ij}\tilde{h}_{kl,kl} = 16\pi G\tilde{\Theta}_{ij} \ , \end{aligned} \tag{3.44}$$

$$2\delta G_{00} = \tilde{h}_{ij,ij} - \frac{2}{3}\nabla^2 h + 2\frac{\dot{a}}{a}\dot{h} = 6\left(\frac{\dot{a}}{a}\right)^2 [\Omega_{\rm c}\delta_{\rm c} + (1+R)\Omega_{\rm r}\delta_{\rm r}] + 16\pi G\Theta_{00} \ , \tag{3.45}$$

$$2\delta G_{0i} = \dot{\tilde{h}}_{ij,j} - \frac{2}{3}\dot{h}_{,i} = -8\left(\frac{\dot{a}}{a}\right)^2 (1+R)\Omega_{\rm r}v_{\rm r}^i + 16\pi G\Theta_{0i} \ , \tag{3.46}$$

where the traceless parts are defined by $\tilde{R}_{ij} = R_{ij} - \delta_{ij}R^k{}_k/3$, and similarly for $\tilde{h}_{ij}$ and $\tilde{\Theta}_{ij}$. We note that if the source obeys the covariant conservation equations (3.41) and (3.42), then equations (3.45) and (3.46) are preserved by equations (3.43).

In the standard CDM model where the source is absent, equation (3.43) can be greatly simplified on super-horizon scales $k\eta \ll 1$ in the radiation or matter era:

$$\ddot{\delta}_{\rm c} + \frac{1}{\eta}\dot{\delta}_{\rm c} - \frac{2(2+R)}{\eta^2}\delta_{\rm c} = 0, \quad \text{in radiation era,} \tag{3.47}$$

$$\ddot{\delta}_{\rm c} + \frac{2}{\eta}\dot{\delta}_{\rm c} - \frac{6}{\eta^2}\delta_{\rm c} = 0, \quad \text{in matter era.} \tag{3.48}$$

Since $R = 3\Omega_{\rm B0}a/4\Omega_{\rm c0}a_{\rm eq}$ by definition, we know $R \ll 1$ deep in the radiation era. Thus the above equations both have a growing mode $\delta_{\rm c} \propto \eta^2$. This result can be also used to investigate the numerical errors in simulations of structure formation with defects. In this case, numerical errors might appear as white noise on superhorizon modes $k \lesssim 1/\eta$, so they will have a growing behaviour $S(k) = 4\pi k^3\mathcal{P}(k) \propto k^3\eta^4$. For the horizon crossing mode $k \sim 1/\eta$, this becomes $S(k) \propto \eta$ [82]. Thus we see that although energy conservation together with causality should forbid the growth of density perturbations on superhorizon scales, any numerical errors seeded from early times would induce a spurious growing mode. To overcome this problem, one needs to perfectly compensate the source energy in the initially homogeneous background. The following is one of the methods which can achieve this.

3.3.4 Stress-energy conservation of the pseudotensor

The concept of the stress-energy pseudotensor in an expanding universe was first discussed in this context by Veeraraghavan and Stebbins [81], and further investigated by Pen, Spergel and Turok [82]. To introduce this concept, we start from the Minkowski space, where the Bianchi identity $\nabla^{\mu}G_{\mu\nu}=0$ leads to an ordinary conservation law $\partial^{\mu}G_{\mu\nu(1)}=0$ at linear order. Adding the fact that the Einstein equations give $G_{\mu\nu(1)}=8\pi GT_{\mu\nu}-G_{\mu\nu(\mathrm{nl})}$ where $G_{\mu\nu(\mathrm{nl})}$ is the sum of non-linear terms, we see that the right-hand side of this equation provides an ordinarily conserved tensor, the stress-energy pseudotensor.

The generalization of this result to an FRW model is straightforward, with only corrections due to the expansion of the universe. Moving all these corrections to the right-hand side of the Einstein equations, we obtain a pseudo-stress-energy tensor $\tau_{\mu\nu}\equiv G_{\mu\nu(1)}/8\pi G$:

$$\tau_{00} = \frac{3}{8\pi G}\left(\frac{\dot{a}}{a}\right)^2[\Omega_{\mathrm{c}}\delta_{\mathrm{c}}+(1+R)\Omega_{\mathrm{r}}\delta_{\mathrm{r}}]-\frac{1}{8\pi G}\frac{\dot{a}}{a}\dot{h}+\Theta_{00}\ , \tag{3.49}$$

$$\tau_{0i} = -\frac{1}{2\pi G}\left(\frac{\dot{a}}{a}\right)^2(1+R)\Omega_{\mathrm{r}}v_{\mathrm{r}}^i+\Theta_{0i}\ , \tag{3.50}$$

$$\tau_{ij} = \delta_{ij}\frac{1}{8\pi G}\left(\frac{\dot{a}}{a}\right)^2\Omega_{\mathrm{r}}\delta_{\mathrm{r}}-\frac{1}{8\pi G}\frac{\dot{a}}{a}(\dot{\tilde{h}}_{ij}-\frac{2}{3}\dot{h}\delta_{ij})+\Theta_{ij}\ . \tag{3.51}$$

This tensor obeys an ordinary conservation law $\tau^{\mu\nu}{}_{,\nu}=0$ according to the Einstein equations, or equivalently

$$\tau^{00}_{,0} = \tau^{0i}_{,i}\ , \tag{3.52}$$

$$\tau^{i0}_{,0} = \tau^{ij}_{,j}\ . \tag{3.53}$$

This is not a fundamentally new conservation law, but it describes the interchange of energy and momentum among the different components in the universe, i.e. the radiation, matter, and the source in our case. This description appears to be physically more transparent than the original Einstein equations.

The advantage of invoking this formalism is that it is easier for numerical simulations to specify the initial conditions and to maintain proper compensation on superhorizon scales. As we shall explain later, τ_{ij} can only have a white-noise power spectrum on superhorizon scales. Thus integrating equations (3.52) and (3.53) over time will show that τ_{00} has a k^4 power spectrum and that τ_{0i} has a k^2 power spectrum. Therefore, as long as the horizon size at the beginning of the simulation is smaller than scales of our interest, we can set $\tau_{00}=\tau_{0i}=0$ as the initial condition, allowing for perturbations to grow inside the horizon and fall off as k^4 in τ_{00} outside the horizon. For simulations of structure formation with defects, a check of $\tau_{00}\propto k^4$ on superhorizon modes will tell us whether or not the compensation is well obeyed.

One example of using the pseudo-stress-energy tensor formalism is in the study of density perturbations. In this case, a convenient closed set of equations will be to combine the conservation equation for radiation (3.40), the definition of pseudoenergy (3.49), and one of the alternative Einstein equations using the pseudo-stress-energy tensor (3.52), thus yielding

$$\ddot{s} = -\frac{\dot{R}}{1+R}\dot{s}+\frac{1}{3+3R}\nabla^2(s+\delta_{\mathrm{c}})\ , \tag{3.54}$$

$$\dot{\delta}_{\mathrm{c}} = 4\pi G\frac{\dot{a}}{a}(\tau_{00}-\Theta_{00})-\frac{\dot{a}}{a}\left\{\left[\frac{3}{2}\Omega_{\mathrm{c}}+2(1+R)\Omega_{\mathrm{r}}\right]\delta_{\mathrm{c}}+2(1+R)\Omega_{\mathrm{r}}s\right\}\ , \tag{3.55}$$

$$\dot{\tau}_{00} = \Theta_{0i,i}+\frac{1}{2\pi G}\left(\frac{\dot{a}}{a}\right)^2(1+R)\Omega_{\mathrm{r}}\dot{s}\ , \tag{3.56}$$

where we have used equations (3.28), (3.38), (3.39) and (3.50) to eliminate $\dot{h}$, δ_{r}, v_{r}^i and τ_{0i} respectively. By analogy to the results in Ref. [82], here we see that the pseudoenergy τ_{00} and the entropy fluctuation s have both been built into the equations.

3.3.5 Initial conditions

Causality means that local physical processes moving matter around in an initially smooth universe can never induce perturbations on scales much larger than the horizon. Therefore the spatial part of $\Theta_{\mu\nu}$, $\tau_{\mu\nu}$ and $h_{\mu\nu}$ can only have white noise power spectra on superhorizon modes [81]. The same applies to δ_N and therefore h. It then follows from equations (3.36), (3.42) and (3.53) respectively that the power spectra of $\mathbf{v}_{\mathrm{r}}$, Θ_{0i} and τ_{0i} all fall off as k^2 outside the horizon. From equations (3.39) and (3.52), we also have the spectra of s, $\dot{s}$ and τ_{00} proportional to k^4 on these scales, as previously discussed. As a summary, we have for $k\ll 1/\eta$ that

$$\delta_{\mathrm{c}},\delta_{\mathrm{r}},h,h_{ij},\Theta_{ij},\tau_{ij} \propto k^0\ , \tag{3.57}$$

$$\mathbf{v}_{\mathrm{r}},\Theta_{0i},\tau_{0i} \propto k^1\ , \tag{3.58}$$

$$s,\dot{s},\tau_{00} \propto k^2\ , \tag{3.59}$$

where '$\propto k^n$' means the power spectrum is proportional to k^{2n}. Thus a convenient set of initial conditions for simulations will be

$$v_{\rm r}^i = \Theta_{0i} = \tau_{0i} = s = \dot{s} = \tau_{00} = 0. \tag{3.60}$$

With such a choice, we can see from equation (3.49) that there is still freedom in choosing among $\delta_{\rm c}$, $\delta_{\rm r}$ and $\dot{h}$ in which to compensate Θ_{00}. Nevertheless, as we shall prove analytically in Chapter 4, no matter how Θ_{00} was compensated into the background contents of the universe when the source was initially formed, the resulting matter density perturbations today would be the same, i.e. the resulting matter perturbations today are independent of the way the source was initially compensated into the background. This was first observed numerically in Ref. [82], and we shall provide an analytic proof later.

3.3.6 Instantaneous decoupling

One thing we have not included in our formalism is the treatment after the decoupling epoch $\eta_{\rm d}$. Before this epoch, photons and baryons are assumed to be tightly coupled, forming a single γB fluid. At the decoupling epoch $\eta_{\rm d}$, baryons and photons are assumed to be instantaneously decoupled from each other, so that δ_γ and $\delta_{\rm B}$ evolve separately afterwards. A numerical fit to the redshift of the decoupling epoch is [91]

$$\begin{aligned}
z_{\rm d} &= 1291\frac{(\Omega_{\rm m0}h^2)^{0.251}}{1+0.659(\Omega_{\rm m0}h^2)^{0.828}}\left[1+b_1(\Omega_{\rm B0}h^2)^{b_2}\right], && (3.61)\\
b_1 &= 0.313(\Omega_{\rm m0}h^2)^{-0.419}\left[1+0.607(\Omega_{\rm m0}h^2)^{0.674}\right], && (3.62)\\
b_2 &= 0.238(\Omega_{\rm m0}h^2)^{0.223}. && (3.63)
\end{aligned}$$

Although this result was obtained for the decoupling epoch of baryons and there was another fit for that of photons, these two epochs—the recombination of baryons and the last scattering of photons—coincide approximately in the absence of subsequent reionization [92, 93].

In addition, the photons and baryons are not in fact perfectly coupled, and this leads to the diffusion damping of photons and Silk damping of baryons [94] during the decoupling epoch. To model these effects, we apply damping envelopes to both δ_γ and $\delta_{\rm B}$ at the decoupling epoch $z_{\rm d}$, i.e.

$$\hat{\delta}_{N(\rm d)} = \widetilde{\delta}_{N(\rm d)}\mathcal{D}_N(k), \quad N = \gamma,\ \rm B, \tag{3.64}$$

where the tilde indicates the Fourier transform and k is the wave number. The photon diffusion damping envelope can be approximated by the form [93]

$$\mathcal{D}_\gamma(k) \simeq e^{-(k/k_\gamma)^{m_\gamma}}, \tag{3.65}$$

where

$$\begin{aligned}
\frac{k_\gamma}{\rm Mpc^{-1}} &= \left\{\frac{2}{\pi}\arctan\left[\frac{\pi}{2}\left(\frac{F_2}{F_1}\right)^{p2/p1}(\Omega_{\rm B0}h^2)^{p2}\right]\right\}^{p1/p2} F_1\,, && (3.66)\\
m_\gamma &= 1.46(\Omega_{\rm m0}h^2)^{0.0303}\left(1+0.128\arctan\left\{\ln\left[(32.8\Omega_{\rm B0}h^2)^{-0.643}\right]\right\}\right), && (3.67)\\
p_1 &= 0.29, && (3.68)\\
p_2 &= 2.38(\Omega_{\rm m0}h^2)^{0.184}, && (3.69)\\
F_1 &= 0.293(\Omega_{\rm m0}h^2)^{0.545}\left[1+(25.1\Omega_{\rm m0}h^2)^{-0.648}\right], && (3.70)\\
F_2 &= 0.524(\Omega_{\rm m0}h^2)^{0.505}\left[1+(10.5\Omega_{\rm m0}h^2)^{-0.564}\right]. && (3.71)
\end{aligned}$$

Silk damping for the baryons can likewise be approximated as [93]

$$\mathcal{D}_{\rm B}(k) \simeq e^{-(k/k_{\rm S})^{m_{\rm S}}}, \tag{3.72}$$

where

$$\begin{aligned}
\frac{k_{\rm S}}{\rm Mpc^{-1}} &= 1.38(\Omega_{\rm m0}h^2)^{0.398}(\Omega_{\rm B0}h^2)^{0.487}\frac{1+(96.2\Omega_{\rm m0}h^2)^{-0.684}}{1+(346\Omega_{\rm B0}h^2)^{-0.842}}, && (3.73)\\
m_{\rm S} &= 1.40\frac{(\Omega_{\rm B0}h^2)^{-0.0297}(\Omega_{\rm m0}h^2)^{0.0282}}{1+(781\Omega_{\rm B0}h^2)^{-0.926}}. && (3.74)
\end{aligned}$$

In the context with active sources, the damping envelopes (3.65) and (3.72) will depart from their form of exponential fall-off to power-law decay towards smaller angular scales [95]. This is due to the survival of perturbations which are actively seeded just before the last scattering. However, this departure from the exponential fall-off was observed to be on scales of order a few arc-minutes (e.g. the multipole index $l \gtrsim 3000$ for cosmic strings) [95], and this is certainly beyond the scales of our interest. Moreover, because we expect the post-decoupling contribution in the perturbations seeded by defects to have a power-law fall-off on small-angular scales due to a certain topology of the source, the small-angular power in the final perturbations is likely to be dominated by this post-decoupling contribution, rather than the primary perturbations (those seeded before the decoupling), which first exponentially decay and then turn to a power-law fall-off when l is greater than a few thousands. Therefore, the damping approximation employed here should be appropriate for defect models on the scales of our interest.

Now we consider the evolution of δ_γ and $\delta_{\rm B}$ after the decoupling epoch $z_{\rm d}$. From the energy conservation law (3.24)–(3.26), we have for the baryon perturbations

$$\ddot{\delta}_{\rm B} - \ddot{\delta}_{\rm c} + \frac{\dot{a}}{a}(\dot{\delta}_{\rm B} - \dot{\delta}_{\rm c}) = 0. \tag{3.75}$$

This implies $(\dot{\delta}_{\rm B} - \dot{\delta}_{\rm c}) \propto a^{-1}$, meaning that the evolution of $\delta_{\rm B}$ and $\delta_{\rm c}$ will soon approach the same behaviour. Adding the fact that matter perturbations grow as η^2 in the matter era so that $[\delta_{\rm B(d)} - \delta_{\rm c(d)}]$ is relatively small when compared with either $\delta_{\rm B0}$ or $\delta_{\rm c0}$, in calculating $\delta_{\rm B0}$ and $\delta_{\rm c0}$ to linear order it is appropriate to combine $\delta_{\rm B}$ and $\delta_{\rm c}$ at the decoupling epoch $z_{\rm d}$ as

$$\widetilde{\delta}_{\rm m(d)} = \frac{\Omega_{\rm B0}\hat{\widetilde{\delta}}_{\rm B(d)} + \Omega_{\rm c0}\widetilde{\delta}_{\rm c(d)}}{\Omega_{\rm B0} + \Omega_{\rm c0}} = \frac{3\Omega_{\rm B0}\widetilde{\delta}_{\gamma({\rm d})}\mathcal{D}_{\rm B}/4 + \Omega_{\rm c0}\widetilde{\delta}_{\rm c(d)}}{\Omega_{\rm B0} + \Omega_{\rm c0}}, \tag{3.76}$$

and the same for their time derivatives. We then have only two fluids after the decoupling: the photon fluid (Ω_γ) and the matter fluid linearly combined from the CDM and baryon fluids $(\Omega_{\rm m} = \Omega_{\rm c} + \Omega_{\rm B})$. Eventually we can take the matter perturbations at the present epoch to be $\widetilde{\delta}_{\rm c0} \equiv \widetilde{\delta}_{\rm B0} \equiv \widetilde{\delta}_{\rm m0}$.

To sum up, we first evolve the CDM and γB perturbations up to the decoupling epoch $z_{\rm d}$ given by (3.61), noting that our formalism extracts the photon component from the γB fluid. We then apply damping envelopes to $\widetilde{\delta}_{\gamma({\rm d})}$ and $\widetilde{\delta}_{\rm c(d)}$ as illustrated in equation (3.64) to account for the photon diffusion and Silk damping. $\widetilde{\delta}_{\rm m(d)}$ is then obtained by equation (3.76). Finally we carry on the evolution of $\widetilde{\delta}_{\rm m}$ and $\widetilde{\delta}_{\rm r}$ from the epoch $z_{\rm d}$ to the present, using our previous results with $R = 0$ and the subscript 'c' replaced by 'm'.

3.4 Accuracy for the standard CDM model

To verify our scheme for evolving cosmological perturbations, we first calculate the CDM transfer function in the context of the standard CDM model (c.f. equation (2.39)):

$$T_{\rm c}(k, \eta_0) = \frac{\widetilde{\delta}_{\rm c}(k, \eta_0)\widetilde{\delta}_{\rm c}(0, 0)}{\widetilde{\delta}_{\rm c}(k, 0)\widetilde{\delta}_{\rm c}(0, \eta_0)}, \tag{3.77}$$

where η_0 is the present conformal time. To this end, we use equations (3.54), (3.55) and (3.56) in the absence of the source terms, and the instantaneous decoupling approximation described above. We start the evolution in the deep radiation era when $\Omega_{\rm m} \ll \Omega_{\rm r} \approx 1$, $R \ll 1$, and $\eta_{\rm i} \ll 1/k$ for a given mode k. In this case, one choice of the initial conditions is

$$s = \dot{s} = 0, \quad \delta_{\rm c} = \eta_{\rm i}^2, \quad \tau_{00} = \frac{1}{\pi G}. \tag{3.78}$$

Figure 3.2 shows our results for the CDM transfer functions at the present epoch $T_{\rm c}(k, \eta_0)$ in different cosmologies, together with the results obtained from CMBFAST [85]. We see that our results agree very well with those from CMBFAST. The discrepancy reaches its maximum of about 5% at the scale $k \sim 1h{\rm Mpc}^{-1}$ in the open model with $\Omega_{\rm c0} = 0.15$ and $\Omega_{\rm B0} = 0.05$. Our results have also been checked against the results in Ref. [35], and they are in agreement again within a 5% error. In addition, we notice the oscillations resulting from the photon-baryon coupling before $\eta_{\rm d}$ in cosmologies with a high baryon fraction $\Omega_{\rm B0}/\Omega_{\rm m0}$ (see the bottom curves in Figure 3.2). Next, we calculate the radiation transfer function at the decoupling epoch, since the radiation perturbations at this epoch will appear as the intrinsic CMB anisotropies (see section 1.4.2). We define this transfer function as:

$$T_{\rm r}(k, \eta_{\rm d}) = \frac{\widetilde{\delta}_{\rm r}(k, \eta_{\rm d})\widetilde{\delta}_{\rm c}(0, 0)}{\widetilde{\delta}_{\rm c}(k, 0)\widetilde{\delta}_{\rm c}(0, \eta_0)}, \tag{3.79}$$

where we normalise the radiation perturbations at $\eta_{\rm d}$ to both the amplitude of CDM perturbations today and the initial CDM power spectrum. This definition will enable us to verify not only the scale dependence of

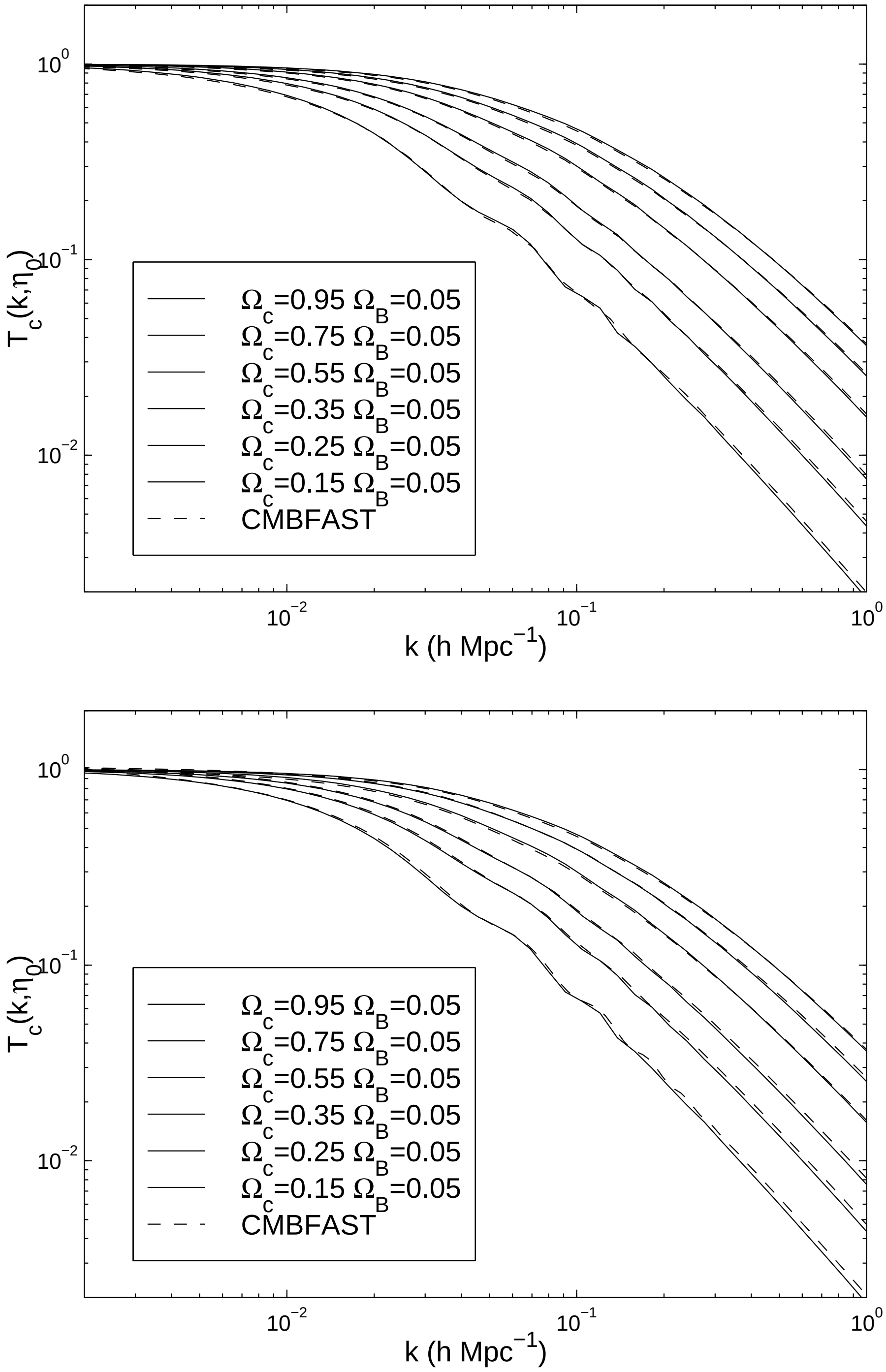

Figure 3.2: Comparison of our CDM transfer functions at the present epoch $T_{\rm c}(k, \eta_0)$ with results obtained from CMBFAST. On the top are results in flat models with a cosmological constant (i.e. $\Omega_{\Lambda 0} + \Omega_{\rm c0} + \Omega_{\rm B0} = 1$). At the bottom are results in open models without a cosmological constant. Results using our formalism are plotted as solid lines, while the results from CMBFAST are plotted as dashed lines. We have used $h = 0.7$ throughout. The mass fraction of Helium-4 $Y_{\rm He} = 0.24$ and the number of neutrino species $N_\nu = 3.04$ have been used in obtaining the results from CMBFAST.

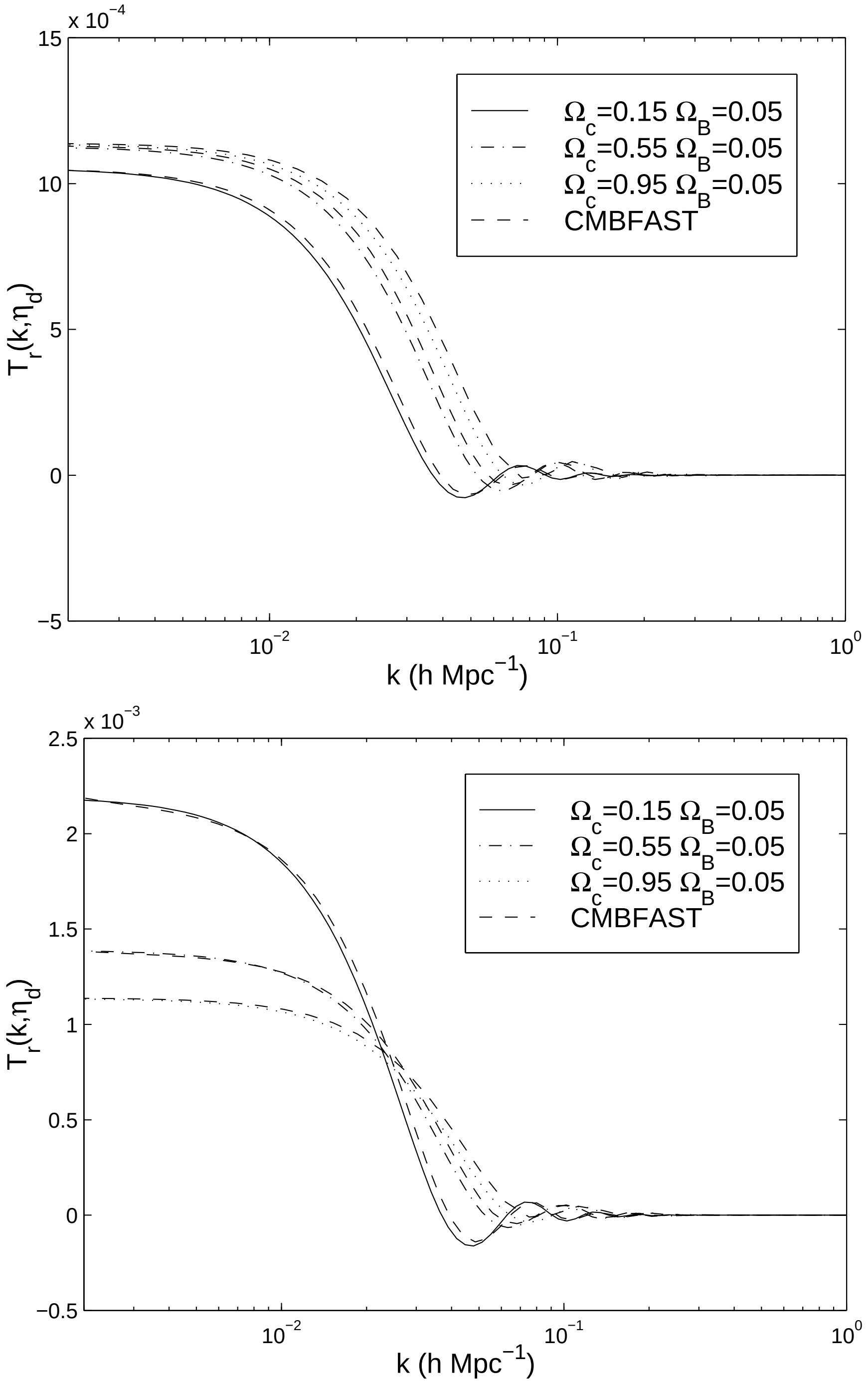

Figure 3.3: Comparison of our radiation transfer functions at the decoupling epoch $T_{\rm r}(k, \eta_{\rm d})$ with results obtained from CMBFAST. On the top are results in flat models with a cosmological constant (i.e. $\Omega_{\Lambda 0} + \Omega_{\rm c0} + \Omega_{\rm B0} = 1$). At the bottom are results in open models without a cosmological constant. Our results are plotted as solid, dot-dashed, and dotted lines, while the CMBFAST results are plotted as dashed lines. Note that this transfer function has been normalised to both the amplitude of CDM perturbations today and the initial CDM power spectrum (see equation (3.79)).

the evolution of radiation perturbations, but also their overall normalisation. Figure 3.3 shows our results, again as a comparison with the results obtained from CMBFAST. We see that although the scale dependence of our results is slightly different from that of the CMBFAST results, the overall normalisation appears to be quite accurate. The sideway shift of the oscillatory peaks in our results when compared with the peaks from CMBFAST has a maximum of about 5% in the flat model with $\Omega_{c0} = 0.95$ and $\Omega_{B0} = 0.05$. This discrepancy results naturally from the instantaneous decoupling approximation which we have employed. Thereby, despite the small departure from results obtained using the CMBFAST Boltzmann scheme, our formalism provides a much more numerically efficient way of calculating density perturbations and scanning a large parameter range of background cosmologies. We consequently have confidence in applying this formalism to causal models with cosmic defect seeds.

3.5 Conclusion

In this chapter we first solved the background dynamics of the universe in different cosmologies, and then set up a formalism which can be used to study the evolution of cosmological perturbations in the presence of causal seeds. In this formalism we invoked the fluid approximation to model the contents of the universe, and assumed photon-baryon tight coupling before the drag epoch to account for the baryonic effect. The approximation of instantaneous decoupling of photons and baryons was also invoked. In particular, by comparing our results for density perturbations in the context of the standard CDM model to the results calculated from CMBFAST, we have demonstrated the good accuracy of our formalism. Having established this formalism, we can now proceed to investigate its solutions and applications in the context of topological defect models, especially the cosmic strings.

Chapter 4

Causal Compensated Matter Perturbations

4.1 Introduction

In the previous chapter, the linearly perturbed Einstein equations with source terms were derived in the synchronous gauge. In this chapter, we further investigate their solutions for density perturbations, in particular the matter density perturbations. We first derive solutions in a flat $\Lambda = 0$ universe. A simple yet accurate extrapolation scheme is then employed to obtain solutions in $K \neq 0$ or $\Lambda \neq 0$ models. With these solutions, we reveal some important properties of the evolution of density perturbations. In particular, we prove that the resulting matter perturbations today are independent of the way the source energy is initially compensated in the background contents of the universe. The compensation mechanism and the scales on which it operates are also studied using our solutions. We find that this compensation scale is determined not only by the dynamics of the universe, but also by the properties of the source near the horizon scale. This scale can be accurately located using the solutions obtained here once the detailed features of the source are known.

Although the framework presented here has been proposed already by Veeraraghavan and Stebbins in the literature [81], their solutions are wrong due to incorrect initial conditions and incorrect assumptions about the form of the subsequent perturbations induced by the defect source (see text later). We correct these mistakes in this chapter, and further provide an explicit set of analytic solutions which are valid for studying matter perturbations in general causal models with any reasonably chosen background cosmology.

The structure of the chapter is as follows. In section 4.2 we introduce the evolution equations of linear density perturbations. The perturbations are divided into two parts: the initial and the subsequent perturbations. With this decomposition for the perturbations, in section 4.3 we find the exact solutions on scales much larger or much smaller than the horizon size, namely the superhorizon or the subhorizon solutions respectively. Solutions on intermediate scales are then obtained by an accurate interpolation scheme using the well-known standard CDM transfer function. In section 4.4, a simple extrapolation scheme is introduced to obtain solutions in the $K \neq 0$ or $\Lambda \neq 0$ cosmologies. Section 4.5 investigates some important properties of these causally-sourced density perturbations. This includes the relation between our solutions and the standard CDM transfer function, the independence of initial condition as mentioned above, and finally the compensation mechanism.

4.2 Evolution equations and decomposition of perturbations

4.2.1 Perturbations equations

We first consider density perturbations about a flat FRW model with a cosmological constant Λ, which are causally sourced by an evolving source network with the energy-momentum tensor $\Theta_{\mu\nu}(\mathbf{x}, \eta)$. As seen in the previous chapter, with the photon-baryon tight coupling approximation in the synchronous gauge, the linear evolution equations of the radiation and CDM perturbations can be given by equations (3.54), (3.55) and (3.56), which are derived from equations (3.40), (3.49) and (3.52). For analytic simplicity, however, one can drop the use of τ_{00}, which we shall consider later, and use equations (3.37) and (3.43) to form an alternative set of evolution equations for density perturbations:

$$\ddot{\delta}_{\rm r} - \frac{4}{3}\ddot{\delta}_{\rm c} + \frac{\dot{R}}{1+R}(\dot{\delta}_{\rm r} - \frac{4}{3}\dot{\delta}_{\rm c}) - \frac{1}{3(1+R)}\nabla^2\delta_{\rm r} = 0, \tag{4.1}$$

$$\ddot{\delta}_{\rm c} + \frac{\dot{a}}{a}\dot{\delta}_{\rm c} - \frac{3}{2}\left(\frac{\dot{a}}{a}\right)^2 [\Omega_{\rm c}\delta_{\rm c} + (2+R)\Omega_{\rm r}\delta_{\rm r}] = 4\pi G\Theta_+, \tag{4.2}$$

After the decoupling epoch $\eta_{\rm d}$, the treatment is essentially the same as that introduced in section 3.3.6. The set of equations (4.1) and (4.2) is in fact equivalent to the set of equations (3.54), (3.55) and (3.56). We have verified in the context of the standard CDM model that both sets of equations give identical transfer functions for the density perturbations, with a numerical discrepancy of less than 0.1%.

Equations (4.1) and (4.2) are exact when $K=0$ (i.e. $\Omega_{K0}=0$). When $K \neq 0$, equations (4.1) and (4.2) are still approximately accurate for models of causal seeds for two reasons. First, the power of density perturbations on the scales of our interest ($k \sim 0.01$–$1h{\rm Mpc}^{-1}$) is mainly seeded around $\eta_{\rm eq}$, when curvature effects are negligible (see Figure 3.1 in Chapter 3). Second, at late times when the curvature effects become important, these scales of interest are already well inside the horizon so that curvature can be neglected. Therefore, the required change to equations (4.1) and (4.2) which account for the effects of curvature only influences the background dynamics, and this involves only $a(\eta)$, $\Omega_{\rm c}(\eta)$ and $\Omega_{\rm r}(\eta)$, whose solutions are given in section 3.2. As can be seen in Figure 3.1, the presence of curvature or a cosmological constant affects the background dynamics only at late times. More precisely, we notice that the largest observable scale for matter perturbations $k \sim 0.01h{\rm Mpc}^{-1}$ corresponds to the horizon size at $\eta \sim 5, 5, 27, 54\eta_{\rm eq}$ for $(\Omega_{\rm m0}, \Omega_{\Lambda 0}) = (0.2, 0), (0.2, 0.8), (1, 0), (2.0, 0)$ respectively, whereas for these models curvature or cosmological constant domination occurs much later at $\eta > \eta_0$. At these moments, the scale factor in the $K \neq 0$ or $\Lambda \neq 0$ models departs from that in the flat $\Lambda = 0$ model only by less than one percent. This also implies that the curvature effect is negligible for the evolution of the source network for structure formation, again because when the curvature becomes significant the source has become unimportant in seeding power on LSS scales.

4.2.2 Decomposing the perturbations

Assuming the source was formed at some initial time $\eta_{\rm i}$ and then evolved to the current time η, it proves useful to split the source-seeded linear perturbations into initial (I) and subsequent (S) parts [81],

$$\delta_N(\mathbf{x},\eta) = \delta_N^{\rm I}(\mathbf{x},\eta) + \delta_N^{\rm S}(\mathbf{x},\eta)\,, \quad N = {\rm c},{\rm r}. \tag{4.3}$$

The initial perturbations $\delta_N^{\rm I}(\mathbf{x},\eta)$ originate from the source configuration at $\eta_{\rm i}$, while the subsequent perturbations $\delta_N^{\rm S}(\mathbf{x},\eta)$ are actively and cumulatively seeded by the later evolution of the source at each $\hat{\eta}$, where $\eta_{\rm i} < \hat{\eta} < \eta$. This is equivalent to having the initial conditions

$$\delta_N^{\rm I}(\eta_{\rm i}) = \delta_N(\eta_{\rm i}), \quad \dot{\delta}_N^{\rm I}(\eta_{\rm i}) = \dot{\delta}_N(\eta_{\rm i}), \tag{4.4}$$

$$\delta_N^{\rm S}(\hat{\eta}) = \dot{\delta}_N^{\rm S}(\hat{\eta}) = 0, \quad \text{with } \eta_{\rm i} < \hat{\eta} < \eta. \tag{4.5}$$

Because the source induces isocurvature perturbations, $\delta^{\rm I}(\mathbf{x},\eta)$ must compensate $\delta^{\rm S}(\mathbf{x},\eta)$ on comoving scales $|\mathbf{x}-\mathbf{x}'| > \eta$ to prevent acausal fluctuation growth on superhorizon scales. One of the aims of this chapter is to show analytically and accurately how this compensation mechanism can be achieved at different scales.

To solve the system of equations (4.1) and (4.2), we use the integral equation with Green functions:

$$\delta_N^{\rm I}(\mathbf{x},\eta) = \sum_{N'} \int d^3x'\, \mathcal{G}_1^{NN'}(X;\eta,\eta_{\rm i})\delta_{N'}(\mathbf{x}',\eta_{\rm i}) + \int d^3x'\, \mathcal{G}_2^{NN'}(X;\eta,\eta_{\rm i})\dot{\delta}_{N'}(\mathbf{x}',\eta_{\rm i}), \tag{4.6}$$

$$\delta_N^{\rm S}(\mathbf{x},\eta) = 4\pi G \int_{\eta_{\rm i}}^{\eta} d\hat{\eta} \int d^3x'\, \mathcal{G}^{N{\rm s}}(X;\eta,\hat{\eta})\Theta_+(\mathbf{x}',\hat{\eta}), \tag{4.7}$$

where $X = |\mathbf{x}-\mathbf{x}'|$. The easiest method of obtaining the Green function solutions is to go to Fourier space and solve the resulting homogeneous system of ordinary differential equations with appropriate initial conditions. Since the Green functions depend only on the modulus of $X = |\mathbf{x}-\mathbf{x}'|$, it follows that their Fourier amplitudes must depend only on the modulus of $\mathbf{k}$. Thus we have

$$\widetilde{\delta}_N^{\rm I}(\mathbf{k},\eta) = \sum_{N'} \left[\widetilde{\mathcal{G}}_1^{NN'}(k;\eta,\eta_{\rm i})\widetilde{\delta}_{N'}(\mathbf{k},\eta_{\rm i}) + \widetilde{\mathcal{G}}_2^{NN'}(k;\eta,\eta_{\rm i})\dot{\widetilde{\delta}}_{N'}(\mathbf{k},\eta_{\rm i})\right], \tag{4.8}$$

$$\widetilde{\delta}_N^{\rm S}(\mathbf{k},\eta) = 4\pi G \int_{\eta_{\rm i}}^{\eta} \widetilde{\mathcal{G}}^{N{\rm s}}(k;\eta,\hat{\eta})\widetilde{\Theta}_+(\mathbf{k},\hat{\eta})\, d\hat{\eta}\,. \tag{4.9}$$

We notice that equation (4.9) is different from the form in Ref. [81], where the authors identified our $\widetilde{\mathcal{G}}^{N\mathrm{s}}$ as $\widetilde{\mathcal{G}}_2^{N\mathrm{c}}$. This identification is incorrect because $\widetilde{\mathcal{G}}^{N\mathrm{s}}$ and $\widetilde{\mathcal{G}}_2^{N\mathrm{c}}$ have different initial conditions as we shall see. With this decomposition of density perturbations, the evolution equations can be solved as follows.

4.3 Solutions in flat $\Lambda = 0$ cosmology

4.3.1 Perturbation equations

For simplicity, we first assume no baryons and therefore set $R = 0$ for now. With the change of variable $y = 1 + A\eta/2$ where $A = 2(\sqrt{2} - 1)/\eta_{\mathrm{eq}}$ leading to $a/a_{\mathrm{eq}} = y^2 - 1$, in Fourier space equations (4.1) and (4.2) with $K = \Lambda = R = 0$ become

$$\widetilde{\mathcal{G}}^{\mathrm{r}\prime\prime} - \frac{4}{3}\widetilde{\mathcal{G}}^{\mathrm{c}\prime\prime} + \frac{4k^2}{3A^2}\widetilde{\mathcal{G}}^{\mathrm{r}} = 0 \ , \tag{4.10}$$

$$(1 - y^2)\widetilde{\mathcal{G}}^{\mathrm{c}\prime\prime} - 2y\widetilde{\mathcal{G}}^{\mathrm{c}\prime} + \left[6 - \frac{12\widetilde{\mathcal{G}}^{\mathrm{r}}/\widetilde{\mathcal{G}}^{\mathrm{c}}}{1 - y^2}\right]\widetilde{\mathcal{G}}^{\mathrm{c}} = 0 \ , \tag{4.11}$$

where a prime represents a derivative with respect to y, $\widetilde{\mathcal{G}}^{\mathrm{c}} \equiv \widetilde{\mathcal{G}}_1^{\mathrm{c}N}$, $\widetilde{\mathcal{G}}_2^{\mathrm{c}N}$ or $\widetilde{\mathcal{G}}^{\mathrm{cs}}$, and $\widetilde{\mathcal{G}}^{\mathrm{r}} \equiv \widetilde{\mathcal{G}}_1^{\mathrm{r}N}$, $\widetilde{\mathcal{G}}_2^{\mathrm{r}N}$ or $\widetilde{\mathcal{G}}^{\mathrm{rs}}$. According to equations (4.8) and (4.9), the initial conditions (4.4) and (4.5) can be rewritten as:

$$\widetilde{\mathcal{G}}_1^{\mathrm{cc}} = \dot{\widetilde{\mathcal{G}}}_2^{\mathrm{cc}} = \widetilde{\mathcal{G}}_1^{\mathrm{rr}} = \dot{\widetilde{\mathcal{G}}}_2^{\mathrm{rr}} = 1 \qquad \text{at} \qquad \eta = \eta_{\mathrm{i}}, \tag{4.12}$$

$$\dot{\widetilde{\mathcal{G}}}^{\mathrm{cs}} = \frac{3}{4}\dot{\widetilde{\mathcal{G}}}^{\mathrm{rs}} = 1 \qquad \text{at} \qquad \eta = \hat{\eta}, \tag{4.13}$$

with all the other Green functions and their time derivatives vanishing. Note that $\widetilde{\mathcal{G}}_i^{NN'}(k; \eta, \eta_{\mathrm{i}}) = 0$ for $\eta \leq \eta_{\mathrm{i}}$, and $\widetilde{\mathcal{G}}^{N\mathrm{s}}(k; \eta, \hat{\eta}) = 0$ for $\eta \leq \hat{\eta}$. We also notice that the Green functions $\widetilde{\mathcal{G}}_i^{NN'}$ only describe the time dependence of the homogeneous version of equations (4.1) and (4.2), while the Green functions $\widetilde{\mathcal{G}}^{N\mathrm{s}}$ are, by the conventional definition, the true Green functions used to solve the inhomogeneous equations (4.1) and (4.2). Moreover, since there are only four variables in equations (4.10) and (4.11) (i.e. $\widetilde{\mathcal{G}}^{\mathrm{c}}$, $\widetilde{\mathcal{G}}^{\mathrm{r}}$, $\dot{\widetilde{\mathcal{G}}}^{\mathrm{c}}$ and $\dot{\widetilde{\mathcal{G}}}^{\mathrm{r}}$), there must exist some dependence among the five sets of Green functions (i.e. $\widetilde{\mathcal{G}}_1^{N\mathrm{c}}$, $\widetilde{\mathcal{G}}_1^{N\mathrm{r}}$, $\widetilde{\mathcal{G}}_2^{N\mathrm{c}}$, $\widetilde{\mathcal{G}}_2^{N\mathrm{r}}$ and $\widetilde{\mathcal{G}}^{N\mathrm{s}}$). This dependence can be observed from the initial conditions (4.12) and (4.13), which yield

$$\widetilde{\mathcal{G}}^{N\mathrm{s}} = \widetilde{\mathcal{G}}_2^{N\mathrm{c}} + \frac{4}{3}\widetilde{\mathcal{G}}_2^{N\mathrm{r}} \, . \tag{4.14}$$

We note that the initial conditions (4.12) and (4.13) are different from those in Ref. [81], where the authors ignored the fact that $\dot{\widetilde{\mathcal{G}}}^{\mathrm{rs}} = 4/3$. This led to the absence of the second term in equation (4.14) and thus the identification of $\widetilde{\mathcal{G}}^{N\mathrm{s}}$ as $\widetilde{\mathcal{G}}_2^{N\mathrm{c}}$, which resulted in incorrect solutions of Green functions in their final results. We shall correct this derivation using our initial conditions (4.12) and (4.13), and further derive a complete set but with an effectively smaller number of Green functions, whose analytic solutions are revealed and numerically verified for the full gamut of k, η_{i} and $\hat{\eta}$.

4.3.2 Super-horizon and sub-horizon modes

For the opposite limits $k\eta \gg 1$ and $k\eta \ll 1$, $\widetilde{\mathcal{G}}^{\mathrm{r}}/\widetilde{\mathcal{G}}^{\mathrm{c}}$ approaches a constant (see below), so equation (4.11) becomes the associated Legendre equation, with solutions composed of the associated Legendre functions $P_2^{-\mu}(y)$ and $Q_2^{\mu}(y)$ where $\mu = \sqrt{12\widetilde{\mathcal{G}}^{\mathrm{r}}/\widetilde{\mathcal{G}}^{\mathrm{c}}}$. We shall use subscripts ∞ and 0 to denote solutions in the limits $k\eta \gg 1$ and $k\eta \ll 1$ respectively. For simplicity, we shall denote both $\hat{\eta}$ and η_{i} as $\hat{\eta}$ in the following solutions.

1. $k\eta \gg 1$: When the wavelengths are much smaller than the horizon size, the radiation oscillates many times per expansion time and its effect is therefore negligible. By setting $\widetilde{\mathcal{G}}^{\mathrm{r}}/\widetilde{\mathcal{G}}^{\mathrm{c}} = 0$, equation (4.11) can be solved as
$$\widetilde{\mathcal{G}}_\infty^{\mathrm{c}}(\eta, \hat{\eta}) = E(\hat{\eta})P_2^0(y) + F(\hat{\eta})Q_2^0(y), \tag{4.15}$$
where $E(\hat{\eta})$ and $F(\hat{\eta})$ are functions of $\hat{\eta}$. This gives the sub-horizon solutions.

2. $k\eta \ll 1$: When the wavelengths are much longer than the horizon size, we have $\widetilde{\mathcal{G}}^{\mathrm{r}}/\widetilde{\mathcal{G}}^{\mathrm{c}} = 4/3$ for the superhorizon modes, giving $\mu = 4$. Thus equations (4.10) and (4.11) yield

$$\widetilde{\mathcal{G}}_0^{\mathrm{r}}(\eta, \hat{\eta}) \ = \ \frac{4}{3}\widetilde{\mathcal{G}}_0^{\mathrm{c}}(\eta, \hat{\eta}) + \alpha_{\mathrm{i}}(\eta - \hat{\eta}) + \beta_{\mathrm{i}} \ , \tag{4.16}$$

$$\begin{aligned}\widetilde{\mathcal{G}}_0^{\rm c}(\eta,\hat\eta) &= G(\hat\eta)P_2^{-4}(y)+H(\hat\eta)Q_2^4(y)\\ &+ 12\int_{y_{\rm i}}^{y}\frac{Q_2^4(x)P_2^{-4}(y)-P_2^{-4}(x)Q_2^4(y)}{Q_2^4(x)P_2^{-4'}(x)-P_2^{-4}(x)Q_2^{4'}(x)}\frac{A\beta_{\rm i}+2\alpha_{\rm i}(x-y_{\rm i})}{A(x^2-1)^2}dx,\end{aligned}\tag{4.17}$$

where $\alpha_{\rm i}$ and $\beta_{\rm i}$ are constants, and $G(\hat\eta)$ and $H(\hat\eta)$ are functions of $\hat\eta$, all determined by the initial conditions. This provides the super-horizon solutions.

Combining with the initial conditions (4.12) and (4.13), we solve $\widetilde{\mathcal{G}}_0^{\rm c}$ and $\widetilde{\mathcal{G}}_\infty^{\rm c}$ as follows. For clarity, we shall denote $\hat y = 1+A\hat\eta/2$ in $\widetilde{\mathcal{G}}^{N{\rm s}}$ and $y_{\rm i}=1+A\eta_{\rm i}/2$ in $\widetilde{\mathcal{G}}_i^{NN'}$ both as w :

$$\widetilde{\mathcal{G}}_\infty^{\rm cs} = \frac{1}{4A}(w^2-1)\left\{(3w^2-1)(3y^2-1)\log\left[\frac{(w+1)(y-1)}{(w-1)(y+1)}\right]-6(y-w)(3wy+1)\right\},\tag{4.18}$$

$$\widetilde{\mathcal{G}}_0^{\rm cs} = \frac{2(y^6w-w^6y-5y^4w+5w^4y+15y^2w-15yw^2+5w-5y)}{5A(y^2-1)^2(w^2-1)},\tag{4.19}$$

$$\begin{aligned}\widetilde{\mathcal{G}}_{1\infty}^{\rm cc} &= \frac{1}{2}(3y^2-1)(3w^2-2)-\frac{9}{2}wy(w^2-1)\\ &+\frac{3}{4}w(w^2-1)(3y^2-1)\log\left[\frac{(y+1)(w-1)}{(y-1)(w+1)}\right],\end{aligned}\tag{4.20}$$

$$\widetilde{\mathcal{G}}_{10}^{\rm cc} = \frac{2yw^5-20yw^3+20y^2w^2+20w^2-30yw-15y^4-5+3y^6+25y^2}{5(y^2-1)^2(w^2-1)},\tag{4.21}$$

$$\widetilde{\mathcal{G}}_{2\infty}^{\rm cc} = \widetilde{\mathcal{G}}_\infty^{\rm cs},\tag{4.22}$$

$$\widetilde{\mathcal{G}}_{20}^{\rm cc} = \widetilde{\mathcal{G}}_0^{\rm cs}-\frac{4}{3}\widetilde{\mathcal{G}}_{20}^{\rm cr},\tag{4.23}$$

$$\widetilde{\mathcal{G}}_{1\infty}^{\rm cr} = 0,\tag{4.24}$$

$$\widetilde{\mathcal{G}}_{10}^{\rm cr} = \frac{3(y^6-5y^2w^4+4yw^5+10y^2w^2-5y^4-5w^4+10y^2-20yw+10w^2)}{5(y^2-1)^2(w^2-1)^2},\tag{4.25}$$

$$\widetilde{\mathcal{G}}_{2\infty}^{\rm cr} = 0,\tag{4.26}$$

$$\begin{aligned}\widetilde{\mathcal{G}}_{20}^{\rm cr} &= -\frac{3}{10A}\left[\frac{(y^2+4y+5)(y-1)^2}{(y+1)^2}\log\left(\frac{y-1}{w-1}\right)+\frac{(y^2-4y+5)(y+1)^2}{(y-1)^2}\log\left(\frac{w+1}{y+1}\right)\right.\\ &\left.+\,2\,(w-y)\frac{(4yw^3-6y^2w^2-10w^2+y^5w-4y^2w+7yw+6y^2-4y^3+5+y^4)}{(w^2-1)(y^2-1)^2}\right].\end{aligned}\tag{4.27}$$

We note that equations (4.22) and (4.23) result directly from the initial conditions (4.12) and (4.13), which lead to equation (4.14).

Despite the complicated forms presented here, all these Green functions have simple asymptotic behaviours in the radiation- or matter-dominated regimes. Since we are more interested in the matter perturbations today and we know from equation (4.11) that $\widetilde{\mathcal{G}}^{\rm c}\propto\eta^2\propto a$ when $\eta/\eta_{\rm eq}\to\infty$, we can design a 'source transfer function' as

$$\widetilde{T}^{\rm c}(k;\hat\eta)\equiv\lim_{\eta/\eta_{\rm eq}\to\infty}\frac{a_{\rm eq}}{a}\,\widetilde{\mathcal{G}}^{\rm c}(k;\eta,\hat\eta).\tag{4.28}$$

Note that this is different from the definition of the standard CDM transfer function (2.39). Equations (4.18)–(4.27) then lead to the source transfer functions:

$$\widetilde{T}_\infty^{\rm cs} = \frac{3}{4A}(w^2-1)\left[(3w^2-1)\log\left(\frac{w+1}{w-1}\right)-6w\right],\tag{4.29}$$

$$\widetilde{T}_0^{\rm cs} = \frac{2w}{5A(w^2-1)},\tag{4.30}$$

$$\widetilde{T}_{1\infty}^{\rm cc} = \frac{3}{2}(3w^2-2)+\frac{9}{4}w(w^2-1)\log\left(\frac{w-1}{w+1}\right),\tag{4.31}$$

$$\widetilde{T}_{10}^{\rm cc} = \frac{3}{5(w^2-1)},\tag{4.32}$$

$$\widetilde{T}_{2\infty}^{\rm cc} = \widetilde{T}_\infty^{\rm cs},\tag{4.33}$$

$$\widetilde{T}_{20}^{\rm cc} = \widetilde{T}_0^{\rm cs}-\frac{4}{3}\widetilde{T}_{20}^{\rm cr} = -\frac{2}{5A}\left[\frac{w}{w^2-1}-\log\left(\frac{w+1}{w-1}\right)\right],\tag{4.34}$$

$$\widetilde{T}_{1\infty}^{\rm cr} = 0,\tag{4.35}$$

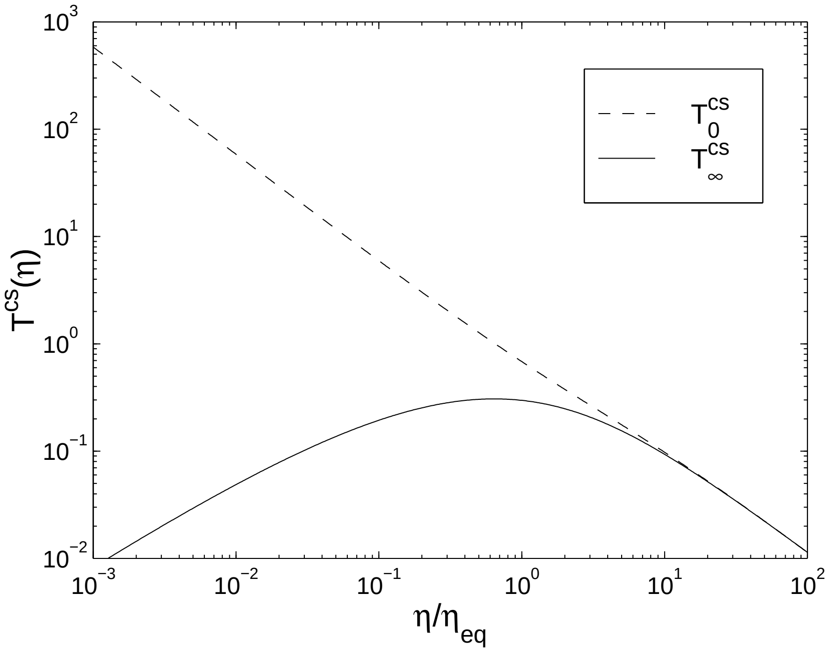

Figure 4.1: The source transfer functions $\widetilde{T}^{\rm cs}_\infty$ ($k\eta \gg 1$, solid line) and $\widetilde{T}^{\rm cs}_0$ ($k\eta \ll 1$, dashed line).

$$\widetilde{T}^{\rm cr}_{10} = \frac{3}{5(w^2-1)^2}\,, \tag{4.36}$$

$$\widetilde{T}^{\rm cr}_{2\infty} = 0\,, \tag{4.37}$$

$$\widetilde{T}^{\rm cr}_{20} = \frac{3}{10A}\left[\frac{2w}{w^2-1} - \log\left(\frac{w+1}{w-1}\right)\right]. \tag{4.38}$$

We plot these source transfer functions in Figures 4.1, 4.2, and 4.3. In the context of topological defects, the defect source was formed at $\eta_{\rm i} \ll \eta_{\rm eq}$. Therefore it would be interesting to investigate the asymptotic behaviours of the source transfer functions at very early times. For $\eta_{\rm i} \ll \eta_{\rm eq}$, equations (4.29)–(4.38) become:

$$\widetilde{T}^{\rm cc}_{1\infty({\rm i})} = \frac{3}{2}\,, \quad \widetilde{T}^{\rm cs}_{\infty({\rm i})} = \widetilde{T}^{\rm cc}_{2\infty({\rm i})} = \frac{3\eta_{\rm i}}{2\eta_{\rm eq}}\log\frac{4\eta_{\rm eq}}{A\eta_{\rm i}}\,, \quad \widetilde{T}^{\rm cr}_{1\infty({\rm i})} = \widetilde{T}^{\rm cr}_{2\infty({\rm i})} = 0\,, \tag{4.39}$$

$$\frac{3A}{2}\widetilde{T}^{\rm cs}_{0({\rm i})} = \widetilde{T}^{\rm cc}_{10({\rm i})} = -\frac{3A}{2}\widetilde{T}^{\rm cc}_{20({\rm i})} = A\frac{\eta_{\rm i}}{\eta_{\rm eq}}\widetilde{T}^{\rm cr}_{10({\rm i})} = A\widetilde{T}^{\rm cr}_{20({\rm i})} = \frac{3\eta_{\rm eq}}{5A\eta_{\rm i}}\,, \tag{4.40}$$

where the subscript (i) denotes the condition $\eta_{\rm i} \ll \eta_{\rm eq}$. These asymptotic behaviours can be clearly seen in Figures 4.2 and 4.3. We note that on sub-horizon scales, $\widetilde{T}^{\rm cs}_\infty$ has a maximum at $\eta \sim \eta_{\rm eq}$ as seen in Figure 4.1. Adding the fact that cosmic defects seed matter perturbations only on sub-horizon modes due to compensation, it follows that defects induce power in matter perturbations mainly during the radiation-matter transition era. This is a generically different mechanism from inflationary models, in which matter perturbations are seeded during inflation in the deep radiation era when all the modes are well outside the horizon. Nevertheless, the defect and inflationary models both provide scale-invariant perturbations at horizon crossing, and these perturbations evolve similarly after horizon crossing.

4.3.3 Degeneracy of the Green functions

In principle we need ten Green functions (five for $\delta_{\rm c} = \delta_{\rm c}^{\rm I} + \delta_{\rm c}^{\rm S}$ and five for $\delta_{\rm r} = \delta_{\rm r}^{\rm I} + \delta_{\rm r}^{\rm S}$) in order to solve equations (4.1) and (4.2) by using the formalism (4.8) and (4.9). However, apart from the dependence (4.14) by which we can reduce the effective number of Green functions, there is another constraint we can invoke—the zero entropy fluctuation on superhorizon scales in the initial conditions, i.e. $s = \dot{s} = 0$ at $\eta_{\rm i}$ for modes $k \ll 1/\eta_{\rm i}$. This allows us to rewrite equation (4.8) as

$$\widetilde{\delta}^{\rm I}_N(\mathbf{k},\eta) = \widetilde{\mathcal{G}}^N_3(k;\eta,\eta_{\rm i})\widetilde{\delta}_{\rm c}(\mathbf{k},\eta_{\rm i}) + \widetilde{\mathcal{G}}^N_4(k;\eta,\eta_{\rm i})\dot{\widetilde{\delta}}_{\rm c}(\mathbf{k},\eta_{\rm i})\,, \tag{4.41}$$

where

$$\widetilde{\mathcal{G}}^N_i = \widetilde{\mathcal{G}}^{N{\rm c}}_{i-2} + \frac{4}{3}\widetilde{\mathcal{G}}^{N{\rm r}}_{i-2}\,, \quad i = 3,4\,. \tag{4.42}$$

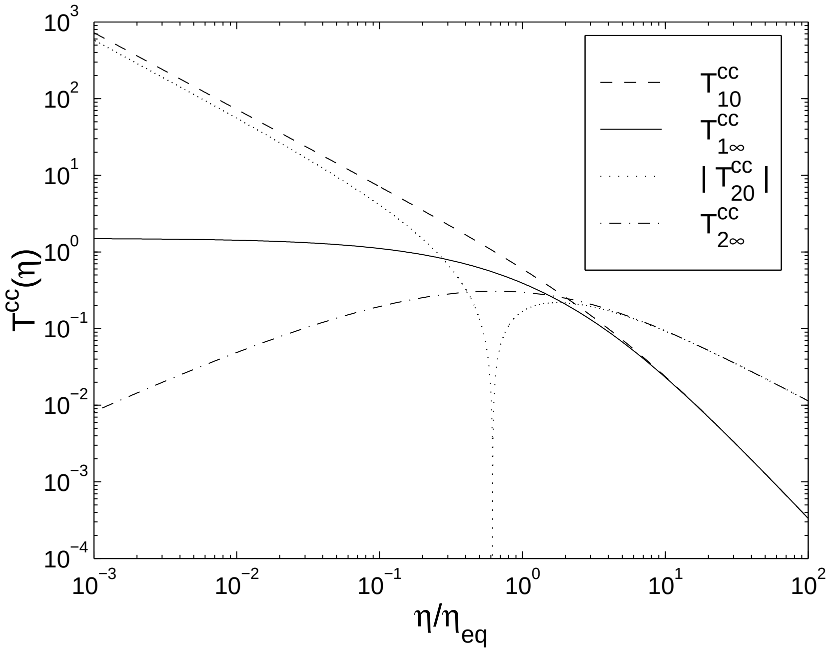

Figure 4.2: The source transfer functions $\widetilde{T}^{\mathrm{cc}}_{1\infty}$, $\widetilde{T}^{\mathrm{cc}}_{10}$, $\widetilde{T}^{\mathrm{cc}}_{2\infty}$ and $\widetilde{T}^{\mathrm{cc}}_{20}$. We have taken the absolute value of $\widetilde{T}^{\mathrm{cc}}_{20}$, because it becomes negative when $\eta \lesssim 0.6\eta_{\mathrm{eq}}$.

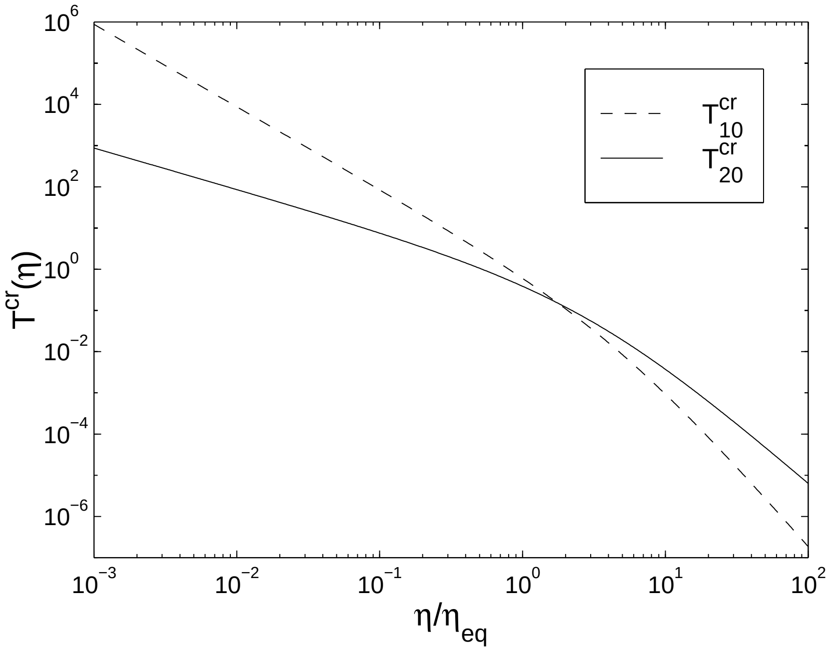

Figure 4.3: The source transfer functions $\widetilde{T}^{\mathrm{cr}}_{10}$ and $\widetilde{T}^{\mathrm{cr}}_{20}$. We note that $\widetilde{T}^{\mathrm{cr}}_{1\infty} = \widetilde{T}^{\mathrm{cr}}_{2\infty} = 0$.

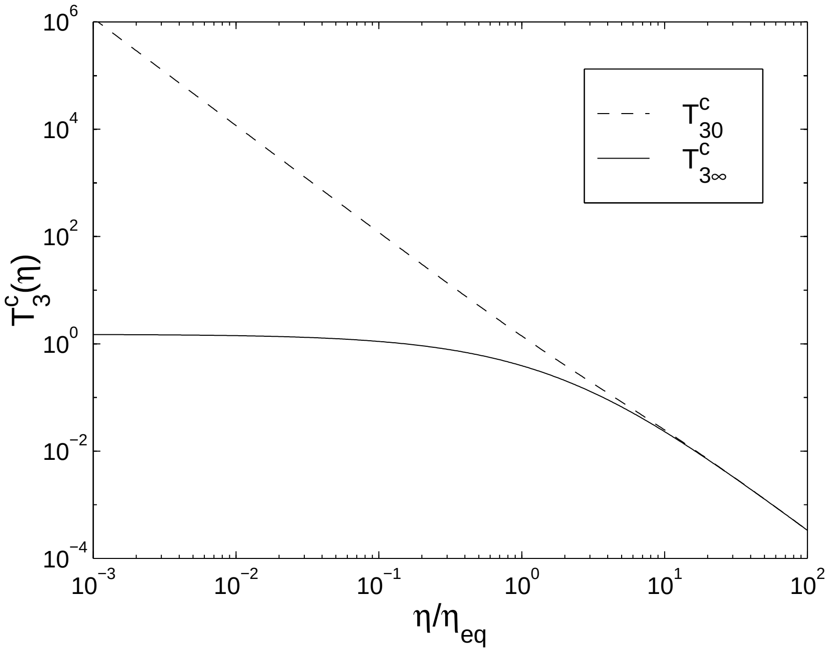

Figure 4.4: The source transfer functions $\widetilde{T}_{3\infty}^{c}$ ($=\widetilde{T}_{1\infty}^{cc}$) and $\widetilde{T}_{30}^{c}$.

Using equation (4.14), we obtain $\widetilde{\mathcal{G}}_4^{c} = \widetilde{\mathcal{G}}^{cs}$, so that $\widetilde{\mathcal{G}}_{4\infty}^{c} = \widetilde{\mathcal{G}}_{\infty}^{cs}$ and $\widetilde{\mathcal{G}}_{40}^{c} = \widetilde{\mathcal{G}}_{0}^{cs}$; from equations (4.20), (4.21), (4.24) and (4.25), we get $\widetilde{\mathcal{G}}_{3\infty}^{c} = \widetilde{\mathcal{G}}_{1\infty}^{cc}$ and

$$\begin{aligned}\widetilde{\mathcal{G}}_{30}^{c} &= \left[5(w^2-1)^2(y^2-1)^2\right]^{-1}\left(y^6+3y^6w^2-5y^4-15w^2y^4+15y^2\right.\\ &\quad \left.+45w^2y^2+2w^7y-10w^3y-6yw^5-50yw+5+15w^2\right). \end{aligned} \tag{4.43}$$

This further yields $\widetilde{T}_{3\infty}^{c} = \widetilde{T}_{1\infty}^{cc}$, $\widetilde{T}_{3\infty(\mathrm{i})}^{c} = \widetilde{T}_{1\infty(\mathrm{i})}^{cc}$, and

$$\widetilde{T}_{30}^{c} = \frac{3w^2+1}{5(w^2-1)^2}, \quad \widetilde{T}_{30(\mathrm{i})}^{c} = \frac{4\eta_{\mathrm{eq}}^2}{5A^2\eta_{\mathrm{i}}^2}. \tag{4.44}$$

Figure 4.4 shows the solutions of $\widetilde{T}_{30}^{c}$ and $\widetilde{T}_{3\infty}^{c}$ ($=\widetilde{T}_{1\infty}^{cc}$). Therefore, the original ten Green functions for solving $\widetilde{\delta}_{\mathrm{c}}$ and $\widetilde{\delta}_{\mathrm{r}}$ have now been reduced to four functions, among which we need only two to solve $\widetilde{\delta}_{\mathrm{c}}$: $\widetilde{\mathcal{G}}^{cs}$ and $\widetilde{\mathcal{G}}_3^{c}$. To calculate $\widetilde{\delta}_{\mathrm{c}}^{\mathrm{S}}$ we need $\widetilde{\mathcal{G}}^{cs}$ from equation (4.9); to calculate $\widetilde{\delta}_{\mathrm{c}}^{\mathrm{I}}$ we need $\widetilde{\mathcal{G}}_4^{c} = \widetilde{\mathcal{G}}^{cs}$ and $\widetilde{\mathcal{G}}_3^{c}$ from equation (4.41). In solving $\widetilde{\delta}_{\mathrm{c}}^{\mathrm{I}}$, we notice that $\widetilde{\mathcal{G}}_3^{c}$ transfers the initial perturbations in both matter and radiation $[\widetilde{\delta}_{\mathrm{c}}(\eta_{\mathrm{i}}) + \widetilde{\delta}_{\mathrm{r}}(\eta_{\mathrm{i}})]$ to today, while $\widetilde{\mathcal{G}}^{cs}$ transfers the initial perturbations in their time derivatives $[\dot{\widetilde{\delta}}_{\mathrm{c}}(\eta_{\mathrm{i}}) + \dot{\widetilde{\delta}}_{\mathrm{r}}(\eta_{\mathrm{i}})]$ to the present.

4.3.4 Solutions on intermediate scales

With $\widetilde{\mathcal{G}}_3^{c}$ and $\widetilde{\mathcal{G}}_4^{c}$ ($=\widetilde{\mathcal{G}}^{cs}$) as the two basis Green functions, we can now work out the solutions on intermediate scales, using results derived in previous sections. In the matter era, $\widetilde{\mathcal{G}}_i^{r} \ll \widetilde{\mathcal{G}}_i^{c}$ on all scales so from equation (4.11) we know that $\widetilde{\mathcal{G}}_{i0}^{c}(k;\eta_0,\eta_{\mathrm{i}}) \approx \widetilde{\mathcal{G}}_{i\infty}^{c}(k;\eta_0,\eta_{\mathrm{i}})$ when $\eta_{\mathrm{i}} \gg \eta_{\mathrm{eq}}$. This can be clearly seen from Figures 4.1 and 4.4. In the radiation era, the perturbations $[\widetilde{\delta}_{\mathrm{c}}(\eta_{\mathrm{i}}) + \widetilde{\delta}_{\mathrm{r}}(\eta_{\mathrm{i}})]$ or $[\dot{\widetilde{\delta}}_{\mathrm{c}}(\eta_{\mathrm{i}}) + \dot{\widetilde{\delta}}_{\mathrm{r}}(\eta_{\mathrm{i}})]$ seeded well before horizon crossing will evolve in the same way as in the standard CDM model due to the same zero entropy fluctuation initial condition. Therefore the solution interpolating between $\widetilde{\mathcal{G}}_{i0}^{c}(k;\eta_0,\eta_{\mathrm{i}})$ and $\widetilde{\mathcal{G}}_{i\infty}^{c}(k;\eta_0,\eta_{\mathrm{i}})$ for $\eta_{\mathrm{i}} \ll \eta_{\mathrm{eq}}$ will be the standard CDM transfer function. Thus we can write down a fit of the solution for the full gamut of k and η_{i} as

$$\widetilde{\mathcal{G}}_i^{c}(k;\eta_0,\eta_{\mathrm{i}}) = \widetilde{\mathcal{G}}_{i\infty}^{c}(\eta_0,\eta_{\mathrm{i}}) + \left[\widetilde{\mathcal{G}}_{i0}^{c}(\eta_0,\eta_{\mathrm{i}}) - \widetilde{\mathcal{G}}_{i\infty}^{c}(\eta_0,\eta_{\mathrm{i}})\right] T(k) I(k;\eta_{\mathrm{i}}), \quad i = 3, 4, \tag{4.45}$$

where

$$T(k) = \left[1 + \frac{(0.0534 + \frac{2.75}{1+3.83k})k^2}{\ln(2e+0.11k)}\right]^{-1} \tag{4.46}$$

is the standard CDM transfer function without baryons (modified from Ref. [91]), k is in units of $\eta_{\rm eq}^{-1}$ (see equation (3.11)), and

$$I(k;\eta_{\rm i}) = \frac{1+30\eta_{\rm i}}{1+30\eta_{\rm i}(1+\frac{k\eta_{\rm i}}{2\pi})} \tag{4.47}$$

is a small correction near horizon crossing to make the analytic solutions (4.45) fit the numerical results. For a given mode which is initially outside the horizon, the background contents of the universe compensate the defect source until horizon crossing. Therefore the detailed behaviour of these Green functions near the horizon scale will affect the so-called compensation scale, beyond which no perturbations can grow. This means that the correction function $I(k;\eta_{\rm i})$ in equation (4.45) actually plays an important role in getting the compensation scale right, and we shall discuss this further in section 4.5.3. We have verified numerically for both $\widetilde{\mathcal{G}}_3^{\rm c}$ and $\widetilde{\mathcal{G}}_4^{\rm c}$ that the fit (4.45) is accurate within a 4% error for any k and $\eta_{\rm i}$ (note that the initial conditions for the numerical verifications can be obtained from equations (4.12), (4.13) and (4.42)). Figure 4.5 shows the numerical solutions of $\widetilde{\mathcal{G}}_3^{\rm c}$ and $\widetilde{\mathcal{G}}_4^{\rm c}$ ($=\widetilde{\mathcal{G}}^{\rm cs}$) with a chosen domain in $(k,\eta_{\rm i})$. It confirms the asymptotic behaviours indicated by equations (4.29), (4.30), (4.31) and (4.44), and plotted in Figures 4.1 and 4.4. In particular, one can verify from Figure 4.5 the asymptotic behaviours (4.39), (4.40) and (4.44) when $\eta_{\rm i} \ll \eta_{\rm eq}$:

$$\widetilde{T}^{\rm c}_{30({\rm i})} = \frac{4\eta_{\rm eq}^2}{5A^2\eta_{\rm i}^2} \propto \eta_{\rm i}^{-2}, \qquad \widetilde{T}^{\rm c}_{3\infty({\rm i})} = \tfrac{3}{2} \propto \eta_{\rm i}^{0}, \tag{4.48}$$

$$\widetilde{T}^{\rm c}_{40({\rm i})} = \frac{2\eta_{\rm eq}}{5A^2\eta_{\rm i}} \propto \eta_{\rm i}^{-1}, \qquad \widetilde{T}^{\rm c}_{4\infty({\rm i})} = \tfrac{3\eta_{\rm i}}{2\eta_{\rm eq}} \log\tfrac{4\eta_{\rm eq}}{A\eta_{\rm i}} \propto \eta_{\rm i}\,, \tag{4.49}$$

where the last proportionality is only an approximation. Schematically, we can divide the $(k,\eta_{\rm i})$-plane into three regions for the solutions of $\widetilde{\mathcal{G}}_i^{\rm c}$ ($i=3,4$). As shown in Figure 4.6, these three domains are: Region I ($k<k_{\rm eq}=1/\eta_{\rm eq}$), Region II ($k>k_{\rm eq}$ and $k>1/\eta_{\rm i}$), and Region III ($k_{\rm eq}<k<1/\eta_{\rm i}$). In Region I, the solution of $\widetilde{\mathcal{G}}_i^{\rm c}$ is $\widetilde{\mathcal{G}}_{i0}^{\rm c}$ because the horizon crossing happens after $\eta_{\rm eq}$, after which $\widetilde{\mathcal{G}}_{i0}^{\rm c} \approx \widetilde{\mathcal{G}}_{i\infty}^{\rm c}$. In Region II, the solution is $\widetilde{\mathcal{G}}_{i\infty}^{\rm c}$ because all modes in this region are inside the horizon all the time. We notice that $\widetilde{\mathcal{G}}_{i0}^{\rm c}$ merges with $\widetilde{\mathcal{G}}_{i\infty}^{\rm c}$ at the boundary of Regions I and II where $\eta_{\rm i} > \eta_{\rm eq}$. In Region III, the solution is in the same form as the standard CDM transfer function along the k direction. This is because modes with larger k cross the horizon earlier, so that their perturbations are suppressed after the horizon crossing for longer until $\eta_{\rm eq}$. In addition, the solution in Region III is in the same form as $\widetilde{\mathcal{G}}_{i0}^{\rm c}$ along the $\eta_{\rm i}$ direction. This is because modes in this region are initially on superhorizon scales, and a given mode with different initial time $\eta_{\rm i}$ will experience the same amount of suppression resulting from the period between the horizon crossing and $\eta_{\rm eq}$. Therefore, Regions I, II and III illustrate the intrinsic property of the solution (4.45).

4.3.5 The effect of baryons

There is one important issue which we have not included—the effect of baryons. Prior to the photon-baryon decoupling at $\eta_{\rm d}$, the CDM and baryons are dynamically independent. In this era, the photon-baryon fluid propagates as acoustic waves with a sound speed given by equation (3.31), preventing baryons from collapsing. Therefore there exists a sound horizon at the decoupling epoch $d_{\rm s(d)}$ (hereafter simply the sound horizon) which is the distance such waves can travel prior to $\eta_{\rm d}$, and which is the largest scale at which the baryons can affect the evolution of density perturbations. It has been shown that inside the sound horizon $d_{\rm s(d)}$, not only are the CDM perturbations seeded before $\eta_{\rm d}$ suppressed due to the presence of baryons (e.g. [35, 91, 93]), but also the baryons themselves have an exponential decay due to the Silk damping [94], with acoustic oscillations due to the velocity overshoot [96, 97]. After the decoupling, baryons evolve in the same way as the CDM does, so the matter perturbations today can be obtained by linearly combining the CDM and baryonic fluctuations at $\eta_{\rm d}$ (see section 3.3.6 and equation (3.76)), and then evolving them to today.

We thus see that baryonic effects tend to suppress the matter perturbations seeded before the decoupling epoch ($\eta<\eta_{\rm d}$, see the horizontal dotted line in Figure 4.6) on scales inside the sound horizon (i.e. for $k > k_{{\rm s(d)}\equiv 2\pi/{\rm d}_{\rm s(d)}}$, see the vertical dotted line in Figure 4.6). The perturbations seeded after $\eta_{\rm d}$ or on scales $k<k_{\rm s(d)}$ will not be affected by the baryons. With this argument, we can impose a suppression factor on our current solution (4.45) to account for the effect of baryons, i.e. the solution with the inclusion of baryons can be written as

$$\widetilde{\mathcal{G}}_i^{\rm c(B)}(k;\eta,\eta_{\rm i};h,\Omega_{\rm m0},\Omega_{\rm B0}) = \widetilde{\mathcal{G}}_i^{\rm c}(k;\eta,\eta_{\rm i})B(k,\eta_{\rm i};h,\Omega_{\rm m0},\Omega_{\rm B0}), \tag{4.50}$$

where $B(k,\eta_{\rm i};h,\Omega_{\rm c0},\Omega_{\rm B0})$ accounts for the baryonic suppression:

$$B(k,\eta_{\rm i};h,\Omega_{\rm m0},\Omega_{\rm B0}) = \begin{cases} \frac{T(k;h,\Omega_{\rm m0},\Omega_{\rm B0})}{T(k;h,0,1)}, & \text{for } \eta_{\rm i} \ll \eta_{\rm d}\,, \\ 1, & \text{for } \eta_{\rm i} \gtrsim \eta_{\rm d}\,, \end{cases} \tag{4.51}$$

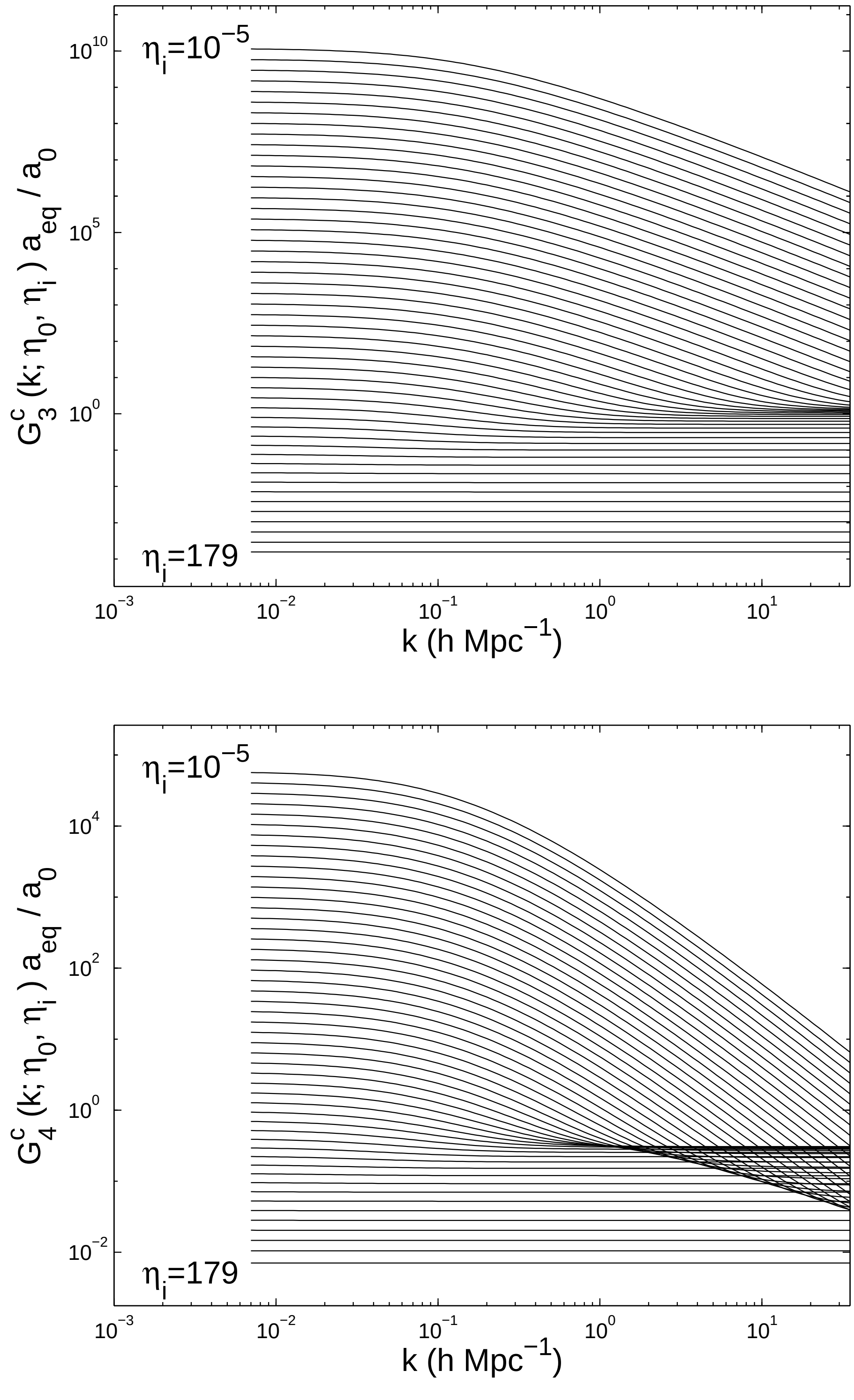

Figure 4.5: The numerical solutions of $\widetilde{\mathcal{G}}_3^{\rm c}(k;\eta_0,\eta_{\rm i})$ and $\widetilde{\mathcal{G}}_4^{\rm c}(k;\eta_0,\eta_{\rm i})$ $(=\widetilde{\mathcal{G}}^{\rm cs}(k;\eta_0,\eta_{\rm i}))$. They both have been normalised to the scale factor today, $a_0/a_{\rm eq}$. Each line has a different initial time $\eta_{\rm i}$, whose smallest and largest values are labelled in both plots. Successive lines have even logarithmic time intervals, and $\eta_{\rm i}$ is in units of $\eta_{\rm eq}$.

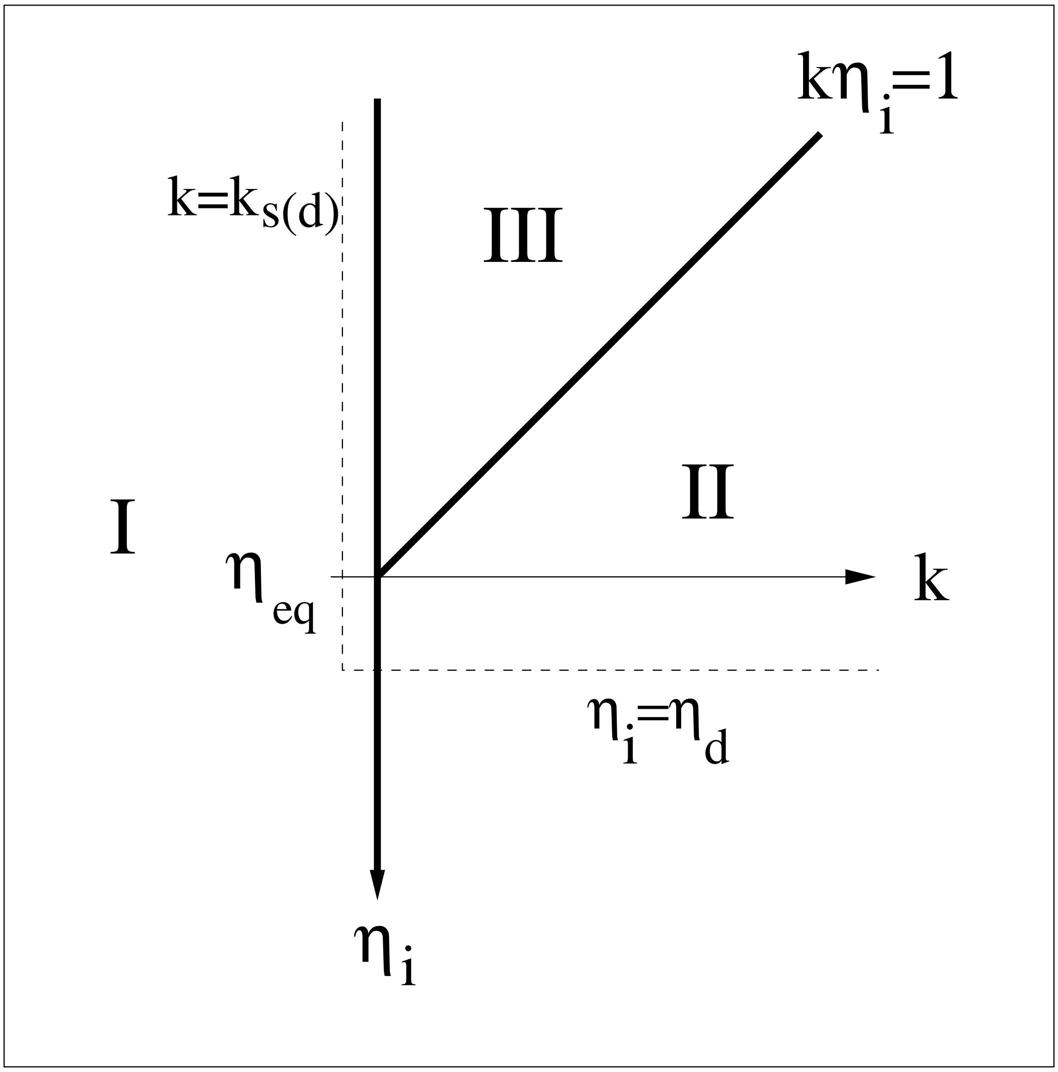

Figure 4.6: Three domains on the $(k, \eta_{\rm i})$-plane for the solutions of Green functions $\widetilde{\mathcal{G}}_3^{\rm c}$ and $\widetilde{\mathcal{G}}_4^{\rm c}$: Region I ($k < k_{\rm eq} = 1/\eta_{\rm eq}$), Region II ($k > k_{\rm eq}$ and $k > 1/\eta_{\rm i}$), and Region III ($k_{\rm eq} < k < 1/\eta_{\rm i}$). These three regions are divided by the thick solid lines as shown.

where $T(k;h,\Omega_{\rm m0},\Omega_{\rm B0})$ is the usual standard CDM transfer function but with baryonic dependence this time. One accurate fit of $T(k;h,\Omega_{\rm m0},\Omega_{\rm B0})$ is provided in Ref. [91]. We note that the ratio $T(k;h,\Omega_{\rm m0},\Omega_{\rm B0})/T(k;h,0,)$ is unity outside the sound horizon ($k < k_{\rm s(d)} \equiv 2\pi/d_{\rm s(d)}$), and is less than unity inside the sound horizon. Referring to Figure 4.6, equation (4.51) means that the value of $B(k,\eta_{\rm i};h,\Omega_{\rm m0},\Omega_{\rm B0})$ is unity in the region to the left and below the dotted lines, and is less than unity otherwise. We also note that the sound horizon can be smaller than the equality horizon ($k_{\rm s(d)} \equiv 2\pi/d_{\rm s(d)} > k_{\rm eq}$) in low-$\Omega_{\rm m0}$ models [91]. In addition, there is a transition era ($\eta_{\rm i} \lesssim \eta_{\rm d}$) which is not included in equation (4.51). This is because in this era the baryonic effects do not fully operate as in the regime $\eta_{\rm i} \ll \eta_{\rm d}$ so that a good fit is not trivial to obtain. We have numerically verified equation (4.51), although an accurate fit to the missing era $\eta_{\rm i} \lesssim \eta_{\rm d}$ has yet to be found [6].

4.4 Solutions in $K \neq 0$ or $\Lambda \neq 0$ models

As discussed before, in $K \neq 0$ or $\Lambda \neq 0$ models, the growth of CDM perturbations departs from that of a flat $\Lambda = 0$ model only at very late times in the matter era. This allows us to apply a universal suppression factor on $\widetilde{\mathcal{G}}^{\rm c(B)}(k;\eta,\hat{\eta})$ to account for the effects of curvature or Λ:

$$\widetilde{\mathcal{G}}_i^{\rm c(B)}(k;\eta_0,\hat{\eta};h,\Omega_{\rm m0},\Omega_{\rm B0},\Omega_{\Lambda0}) = \Omega_{\rm m0}h^2 \times g(\Omega_{\rm m0},\Omega_{\Lambda0}) \times \widetilde{\mathcal{G}}_i^{\rm c(B)}(k;\eta_0,\hat{\eta};1,1,\Omega_{\rm B0},0), \tag{4.52}$$

where k is in units of $\Omega_{\rm m0}h^2\,{\rm Mpc}^{-1}$, and $g(\Omega_{\rm m0},\Omega_{\Lambda0})$ is given by [98]

$$g(\Omega_{\rm m0},\Omega_{\Lambda0}) = \frac{5\Omega_{\rm m0}}{2\left[\Omega_{\rm m0}^{4/7} - \Omega_{\Lambda0} + (1+\Omega_{\rm m0}/2)(1+\Omega_{\Lambda0}/70)\right]}\,. \tag{4.53}$$

In equation (4.52), the leading factor $\Omega_{\rm m0}h^2$ results from the fact that the ratio of scale factors $a_0/a_{\rm eq}$ is proportional to $\Omega_{\rm m0}h^2$ and that the Green function $\widetilde{\mathcal{G}}_i^{\rm c(B)} = \widetilde{T}_i^{\rm c(B)} a_0/a_{\rm eq}$ is proportional to this ratio. The factor $g(\Omega_{\rm m0},\Omega_{\Lambda0})$ accounts for the fact that in a $K \neq 0$ or Λ-universe the linear growth of density perturbations is suppressed relative to an $\Omega_{\rm m0} = 1$ and $\Omega_{\Lambda0} = 0$ universe [98] (also verified in ref. [99] for primordial perturbations). The reason for k to have the unit $\Omega_{\rm m0}h^2\,{\rm Mpc}^{-1}$ in equation (4.52) is that the horizon size at radiation-matter equality $\eta_{\rm eq}$ is proportional to $(\Omega_{\rm m0}h^2)^{-1}$ (see section 3.2).

The extrapolation scheme (4.52) will be invalid at late times when the background dynamics departs from that of a flat $\Lambda = 0$ model. Nevertheless, because at this time the defect source does not seed much power in CDM perturbations on scales of our interest (as argued in section 4.2), equation (4.52) should still provide a very good approximation to the real solution. Indeed, we have numerically verified that up to 5% accuracy, the extrapolation scheme (4.52) is good for $\eta_{\rm i} \leq 60\eta_{\rm eq}$ and $\Omega_\Lambda \leq 0.85$ in Λ-models, for $\eta_{\rm i} \leq 20\eta_{\rm eq}$ and $\Omega_{\rm m0} \geq 0.2$ in open $\Lambda = 0$ models, and for $\eta_{\rm i} \leq 200\eta_{\rm eq}$ and $\Omega_{\rm m0} \leq 2$ in closed $\Lambda = 0$ models.

4.5 Important properties

With the Green function solutions we have found, we can now analytically investigate some important aspects of structure formation by causal seeds.

4.5.1 The standard CDM model

First it would be interesting to investigate the relationship between our Green functions and the standard CDM transfer function, and thereby to justify the use of the standard CDM transfer function in the analytic solution (4.45). In the standard CDM model, there are no subsequent perturbations, so we have $\widetilde{\delta}_N = \widetilde{\delta}_N^{\rm I} + \widetilde{\delta}_N^{\rm S} = \widetilde{\delta}_N^{\rm I}$. As discussed in equations (3.47) and (3.48), we also know that the CDM perturbations have a growing mode $\widetilde{\delta}_{\rm c}(k;\eta) \propto \eta^2$ both for $\eta \gg \eta_{\rm eq}$, and for $\eta \ll \eta_{\rm eq}$ and $k\eta \ll 1$. For superhorizon modes in the radiation era and all modes in the matter era, this allows us to write

$$\widetilde{\delta}_{\rm c}(k;\eta) = A_j(k)\eta^2, \quad j = {\rm R},{\rm M}\,, \tag{4.54}$$

where A_j is the coefficient of the growing mode in the radiation era (j =R: $\eta \ll \eta_{\rm eq}$ and $k\eta \ll 1$) or in the matter era (j =M: $\eta \gg \eta_{\rm eq}$). Thus using our Green function solutions (4.41) and (4.45) with the initial conditions $s = \dot{s} = 0$ and $\dot{\widetilde{\delta}}_N(k;\eta_{\rm i}) = 2\widetilde{\delta}_N(k;\eta_{\rm i})/\eta_{\rm i}$, we can derive the standard CDM transfer function as

$$\begin{aligned}\frac{A_{\rm M}}{A_{\rm R}} &= \frac{\widetilde{\delta}_{\rm c}^{\rm I}(k;\eta)\eta_{\rm i}^2}{\widetilde{\delta}_{\rm c}(k;\eta_{\rm i})\eta^2} = \frac{A^2\eta_{\rm i}^2 a_{\rm eq}}{4\eta_{\rm eq}^2 a}\left[\widetilde{\mathcal{G}}_3^{\rm c} + \frac{2}{\eta_{\rm i}}\widetilde{\mathcal{G}}_4^{\rm c}\right] \\ &= \frac{1}{4\eta_{\rm eq}^2}\left[A^2\eta_{\rm i}^2\widetilde{T}_{30({\rm i})}^{\rm c} + 2A^2\eta_{\rm i}\widetilde{T}_{40({\rm i})}^{\rm c}\right]T(k) = \frac{2}{5}T(k)\,,\end{aligned} \tag{4.55}$$

where we have used $\eta_{\rm i} \ll \eta_{\rm eq} \ll \eta$ and equations (4.48) and (4.49). Here $T(k)$ is the transfer function of the standard CDM model, whose fit is given in equation (4.46). First, we note that the two terms involving $\widetilde{T}^{\rm c}_{30({\rm i})}$ and $\widetilde{T}^{\rm c}_{40({\rm i})}$ contribute equally to the final result, meaning that $[\widetilde{\delta}_{\rm c}(\eta_{\rm i}) + \widetilde{\delta}_{\rm r}(\eta_{\rm i})]$ and $[\dot{\widetilde{\delta}}_{\rm c}(\eta_{\rm i}) + \dot{\widetilde{\delta}}_{\rm r}(\eta_{\rm i})]$ yield an equal quantity of perturbations today. Second, the coefficient 2/5 in the final result of equation (4.55) is well known (e.g. [82, 35]), and here we obtained it using our Green function solutions. This coefficient can be also obtained by knowing from equation (3.56) that τ_{00} is a constant on superhorizon scales, and then comparing its expressions using its definition (3.49) and equation (4.54) for j =R, M. One will find $\tau_{00} = A_{\rm R}/\pi G = 5A_{\rm M}/2\pi G$.

4.5.2 Independence of the initial conditions

One important problem for structure formation with seeds is to investigate how the source energy was compensated into the radiation and matter background when the seeds were formed at $\eta_{\rm i}$. From the result (3.59) we know that the power spectrum of the pseudo energy $\widetilde{\tau}_{00}$ should decay as k^4 on superhorizon modes, and we can thus take $\widetilde{\tau}_{00} = 0$ as part of the initial conditions provided that the scales of interest are well outside the horizon initially. For similar reasons we can take $\widetilde{s} = \dot{\widetilde{s}} = 0$, where $s = 3\delta_{\rm r}/4 - \delta_{\rm c}$. However, from equation (3.49) with no baryons, we have

$$\tau_{00} = \Theta_{00} + \frac{3}{8\pi G}\left(\frac{\dot{a}}{a}\right)^2 \sum_{N={\rm c,r}} \Omega_N \delta_N + \frac{1}{4\pi G}\frac{\dot{a}}{a}\dot{\delta}_{\rm c}. \tag{4.56}$$

This means that for a given $\Theta_{00}(\mathbf{x}, \eta_{\rm i})$, one can have different ways of compensating it between δ_N and $\dot{\delta}_N$. It is thus vital to check the dependence of the resulting initial perturbations $\delta^{\rm I}_{\rm c}(\eta)$ on the way we compensate $\Theta_{00}(\mathbf{x}, \eta_{\rm i})$ into the background. Consider the following two extreme cases, both satisfying $\widetilde{\tau}_{00} = \widetilde{s} = \dot{\widetilde{s}} = 0$ on super-horizon scales at $\eta_{\rm i}$:

1. $\widetilde{\delta}_{\rm c} = 3\widetilde{\delta}_{\rm r}/4 = 0$, $\dot{\widetilde{\delta}}_{\rm c} = 3\dot{\widetilde{\delta}}_{\rm r}/4 = [-4\pi G(a/\dot{a})\widetilde{\Theta}_{00}]_{\rm i}$: Using equation (4.41), the normalized resulting initial perturbations can be calculated as

$$\Delta_1(\eta, \eta_{\rm i}) = \frac{\widetilde{\delta}_{\rm c}^{\rm I}(k, \eta)}{[-4\pi G(a/\dot{a})\widetilde{\Theta}_{00}]_{\rm i}} = \widetilde{\mathcal{G}}^{\rm c}_4 = \widetilde{\mathcal{G}}^{\rm cs}\,. \tag{4.57}$$

2. $\widetilde{\delta}_{\rm c} = 3\widetilde{\delta}_{\rm r}/4 = [-8\pi G(a/\dot{a})^2\widetilde{\Theta}_{00}/(4 - \Omega_{\rm c})]_{\rm i}$, $\dot{\widetilde{\delta}}_{\rm c} = 3\dot{\widetilde{\delta}}_{\rm r}/4 = 0$: Similarly we have

$$\Delta_2(\eta, \eta_{\rm i}) = \frac{\widetilde{\delta}_{\rm c}^{\rm I}(k, \eta)}{[-4\pi G(a/\dot{a})\widetilde{\Theta}_{00}]_{\rm i}} = \widetilde{\mathcal{G}}^{\rm c}_3 \frac{2w(w^2-1)}{A(3w^2+1)}\,. \tag{4.58}$$

To see the difference in $\widetilde{\delta}_{\rm c}^{\rm I}(k, \eta_0)$ today resulting from these two cases, one can calculate

$$D_{12}(\eta_0, \eta_{\rm i}) = \frac{\Delta_2}{\Delta_1} - 1 = \frac{2w(w^2-1)}{A(3w^2+1)}\frac{\widetilde{T}^{\rm c}_{30}}{\widetilde{T}^{\rm cs}_0} = 0\,, \tag{4.59}$$

where we have used equations (4.30), (4.44) and (4.45). This implies that no matter how the source $\Theta_{00}(\mathbf{x}, \eta_{\rm i})$ is compensated into the background when it was formed (i.e. with any portions between δ_N and $\dot{\delta}_N$ initially), it results in the same $\widetilde{\delta}_{\rm c}^{\rm I}(k, \eta_0)$ today for initially superhorizon scales. We note that this independence of the initial conditions was first numerically observed in Ref. [82], and here we have provided an analytic proof.

4.5.3 Compensation and total matter perturbations

With a complete set of Green functions for both initial and subsequent perturbations, we can now investigate the resulting total CDM perturbations and therefore the compensation mechanism. Having seen the independence of the resulting $\widetilde{\delta}_{\rm c}^{\rm I}(\mathbf{k}, \eta_0)$ on the way the source energy is initially compensated into various background components, we can invoke equation (4.57) for $\widetilde{\delta}^{\rm I}_{\rm c}$, and equation (4.9) for $\widetilde{\delta}^{\rm S}_{\rm c}$ to obtain $\widetilde{\delta}_{\rm c}(\mathbf{k}, \eta_0) = \widetilde{\delta}^{\rm I}_{\rm c}(\mathbf{k}, \eta_0) + \widetilde{\delta}^{\rm S}_{\rm c}(\mathbf{k}, \eta_0)$. For a given mode at which $k\eta_{\rm i} \ll 1$ initially, we have:

$$\begin{aligned}
\widetilde{\delta}_{\rm c}(\mathbf{k}, \eta_0) &= \widetilde{\delta}^{\rm I}_{\rm c}(\mathbf{k}, \eta_0) + \widetilde{\delta}^{\rm S}_{\rm c}(\mathbf{k}, \eta_0) \\
&= 4\pi G\left[-\frac{a(\eta_{\rm i})}{\dot{a}(\eta_{\rm i})}\widetilde{\mathcal{G}}^{\rm cs}(k; \eta_0, \eta_{\rm i})\widetilde{\Theta}_{00}(\mathbf{k}, \eta_{\rm i}) + \int_{\eta_{\rm i}}^{\eta_0} \widetilde{\mathcal{G}}^{\rm cs}(k; \eta_0, \hat{\eta})\widetilde{\Theta}_{+}(\mathbf{k}, \hat{\eta})d\hat{\eta}\right]
\end{aligned} \tag{4.60}$$

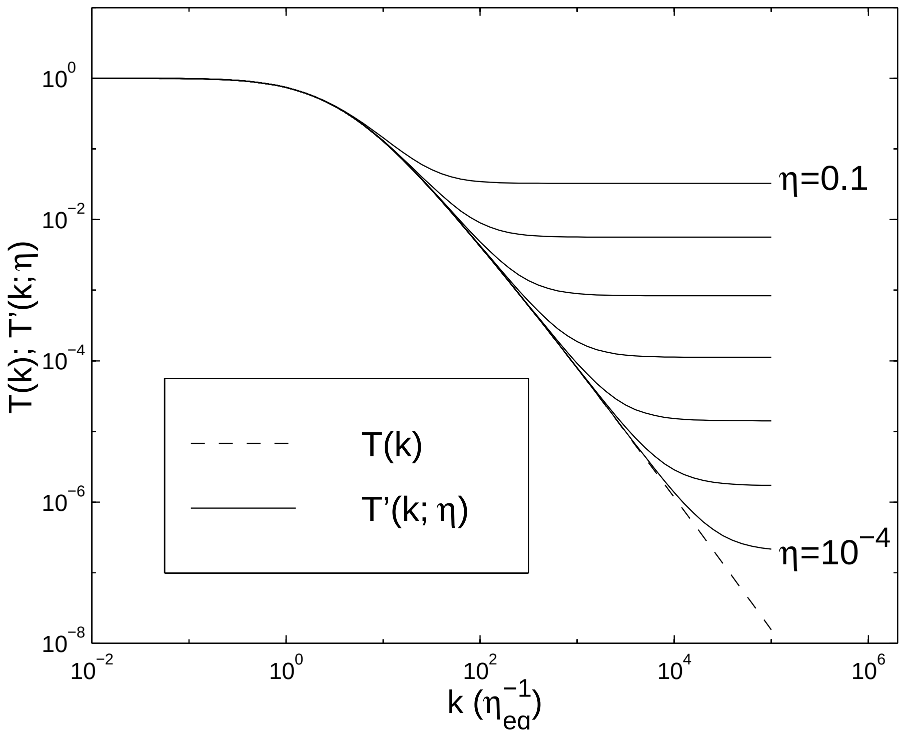

Figure 4.7: The function $T'(k;\eta)$ (solid lines) and the standard CDM transfer function $T(k)$ (the dashed line). Each solid line has different η, whose highest and lowest values are labelled in units of $\eta_{\rm eq}$. Successive lines have even logarithmic time intervals.

$$= \frac{8\pi G a_0}{5A^2 a_{\rm eq}} \left[-T(k)\widetilde{\Theta}_{00}(\mathbf{k},\eta_{\rm i}) + \int_{\eta_{\rm i}}^{\eta_0} T'(k;\hat{\eta}) \frac{\dot{a}(\hat{\eta})}{a(\hat{\eta})} \widetilde{\Theta}_+(\mathbf{k},\hat{\eta}) d\hat{\eta} \right] \tag{4.61}$$

$$= \frac{8\pi G a_0}{5A^2 a_{\rm eq}} \Big\{ -T(k)\widetilde{\Theta}_{00}(\mathbf{k},\eta_0) + \int_{\eta_{\rm i}}^{\eta_0} \left[T'(k;\hat{\eta}) \frac{\dot{a}(\hat{\eta})}{a(\hat{\eta})} \widetilde{\Theta}_+(\mathbf{k},\hat{\eta}) + T(k)\dot{\widetilde{\Theta}}_{00}(\mathbf{k},\hat{\eta}) \right] d\hat{\eta} \Big\} , \tag{4.62}$$

where

$$T'(k;\hat{\eta}) = \frac{\widetilde{\mathcal{G}}_4^{\rm c}(k;\eta_0,\hat{\eta})}{\widetilde{\mathcal{G}}_{40}^{\rm c}(k;\eta_0,\hat{\eta})} , \tag{4.63}$$

and $\widetilde{\mathcal{G}}_4^{\rm c}(k;\eta_0,\hat{\eta})$ is given by (4.45). The function $T'(k;\hat{\eta})$ is plotted in Figure 4.7. Here we notice that the quantities inside the outer most brackets in equations (4.61) and (4.62) are equivalent to nothing but the coefficient of the growing mode in CDM perturbations. Using equation (4.61), one can obtain the resulting perturbations $\widetilde{\delta}_{\rm c}(\mathbf{k},\eta_0)$ by knowing the initial $\widetilde{\Theta}_{00}(\mathbf{k},\eta_{\rm i})$ and integrating the evolution history of $\widetilde{\Theta}_+(\mathbf{k},\hat{\eta})$. Hence this expression is convenient for numerical purposes. In addition, we see that the first term in equation (4.61) comes simply from the initial source energy, serving with an opposite sign to account for energy conservation. This is the so-called compensation. On the other hand, the second term results from the subsequent evolution of $\widetilde{\Theta}_+(\mathbf{k},\hat{\eta})$, which actively creates the CDM density perturbations on subhorizon scales (see later). This term also provides a way for defects to create non-Gaussianity.

Alternatively, equation (4.62) provides both a physically and mathematically transparent way of interpreting how the perturbations are seeded by the source. First consider the integral term for a given mode k. When the mode is well outside the horizon, i.e. $\hat{\eta} \ll 1/k$, $T'(k;\eta)$ equals $T(k)$ by definition. Hence the two terms inside the brackets reduce to $\widetilde{\Theta}_{0i,i}(\mathbf{k},\hat{\eta})T(k)$ due to source stress-energy conservation (3.41). Since the power spectrum of $\widetilde{\Theta}_{0i,i}(\mathbf{k},\hat{\eta})$ falls off as k^4 outside the horizon (see (3.58)), we expect the quantity inside the brackets to be negligible until the given mode approaches horizon crossing. Near horizon crossing, $\widetilde{\Theta}_{0i,i}(\mathbf{k},\hat{\eta})$ is no longer small, and $T'(k;\eta)$ starts departing from $T(k)$ (i.e. $T'(k;\eta) \approx$ constant $> T(k) \propto k^{-2}$, see Figure 4.7), so the two terms begin to contribute to the integral. This also explains why the correction function $I(k;\eta_{\rm i})$ in equation

(4.45) is important in affecting the compensation scale. After horizon crossing, the significance of the two terms then depends on the subhorizon behaviours of their power spectra.

As for the first term in equation (4.62), we see that for a superhorizon mode today, the integral in (4.62) is negligible as argued above so that only the first term contributes. It serves to give the opposite sign to the source energy so as to account for energy conservation on superhorizon scales today, and thus for the compensation at the present epoch η_0. On the other hand, if a given mode is well inside the horizon today, then the first term will be negligible provided that the source energy $\widetilde{\Theta}_{00}(\mathbf{k},\eta_0)$ has a power-law fall-off inside the horizon, as it does for cosmic strings. Therefore in calculating CDM perturbations on scales of our interest, which are well inside the horizon today, the first term in equation (4.62) is negligible, so that it will not affect our compensation argument observed from the integral.

This argument can be further strengthened by deriving the pseudo-energy today. From the definition of τ_{00} (4.56) and the final result of equation (4.62), one obtains

$$\begin{aligned}\widetilde{\tau}_{00}(\mathbf{k},\eta_0) \;&=\; (1-T(k))\,\widetilde{\Theta}_{00}(\mathbf{k},\eta_0)\,+ \\ &\int_{\eta_{\rm i}}^{\eta_0}\left[T'(k;\hat{\eta})\frac{\dot{a}(\hat{\eta})}{a(\hat{\eta})}\widetilde{\Theta}_{+}(\mathbf{k},\hat{\eta})+T(k)\dot{\widetilde{\Theta}}_{00}(\mathbf{k},\hat{\eta})\right]d\hat{\eta}\,. \end{aligned} \tag{4.64}$$

From this result, one can clearly see that for superhorizon modes, $T(k)$ is unity by definition so that only the integral survives. We have also seen from an earlier argument that the quantity inside the square brackets is nothing but the $\widetilde{\Theta}_{0i,i}(\mathbf{k},\hat{\eta})$, which has a k^4 fall-off power spectrum on superhorizon modes. It follows immediately from equation (4.64) that the pseudo-energy today, $\tau_{00}(\eta_0)$, has a k^4 decay power spectrum outside the horizon. This result confirms the superhorizon behaviour of τ_{00} presented in equation (3.59). On the other hand, although $(1-T(k))$ is approximately unity for subhorizon modes, the usual subhorizon power-law decay in $\widetilde{\Theta}_{00}(\mathbf{k},\eta_0)$ (as in the case of cosmic strings) will still make the first term in equation (4.64) negligible inside the horizon.

Thus we can see explicitly in a neat mathematical form how compensation acts on a given lengthscale. From this analysis we can also see that the compensation scale is determined not only by the functions $T'(k;\eta)$ and $T(k)$, but also by the properties of the source near the horizon scale. Once the detailed behaviour of the source near the horizon scale is known, we can accurately locate the compensation scale using equation (4.62) or (4.64). We note that this result is different from the claim in Ref. [100], where multi-fluid compensation back-reaction effects were studied to show that the compensation scale arisen naturally and uniquely from an algebraic identity in the perturbation analysis. Ref. [101] also investigated the compensation scale, and found constraints on the generation of super-causal-horizon energy perturbations from a smooth initial state, under a simple physical scheme. The compensation wavenumber was found to be constrained with $k_{\rm c} \gtrsim 2\eta^{-1}$ due to causality, depending on the behaviour of the causal events. This result is not inconsistent with our finding above, where we further provide a quantitative way to locate the compensation scale for a specific model.

4.6 Conclusion

In this chapter, we derived analytic solutions in the synchronous gauge for the matter density perturbations for a flat $\Lambda = 0$ cosmology. The errors in Ref. [81] were corrected to yield a set of correct and further simplified Green function solutions. These solutions were then extrapolated to $K \neq 0$ or $\Lambda \neq 0$ models, and shown to be accurate for reasonable choices of cosmological parameters. Using these Green function solutions, we investigated several important aspects of structure formation with causal seeds. In particular, we showed that the resulting matter density today is independent of the way the source was initially compensated in the background contents of the universe. We also investigated the compensation mechanism in an explicit manner by considering the Green function formalism. The compensation scale was shown to be not only dependent on the dynamics of the universe, but also on the properties of the source near the horizon scale.

Part III

LARGE-SCALE STRUCTURE WITH COSMIC STRINGS

Chapter 5

Properties of Cosmic String Networks

5.1 Introduction

In Part II, we developed a theoretical framework for the study of cosmological perturbations in the context of causal models. In this part, we shall concentrate on cosmic strings, especially their cosmological implications for large-scale structure. To this end, we first need to study the characteristics of string networks, and this is the theme of this chapter. As mentioned in section 2.3, a string network can be considered as being composed of two parts, long strings and small loops, and we shall investigate each in turn.

The structure of the chapter is as follows. In section 5.2, we begin with a fundamental analysis for objects of different topologies, using the different behaviours of their power spectra to distinguish them. Some important features of long-string networks are then investigated in section 5.3, including their topologies, their power spectra, the correlation length, the correlation time, the thickness of long string paths, the evolution of their energy density, and the unequal time correlator (UETC) of the main seed for large-scale structure formation $\langle\widetilde{\Theta}_+(\mathbf{k},\eta)\widetilde{\Theta}_+(\mathbf{k},\eta')\rangle$. We then move on to investigate loops in section 5.4. We discuss the transfer rate of the energy density from long strings into loops, and describe the method which we use to model the loop decay into gravitational waves. This treatment is missing from the AS simulations, and will help us understand the role of small loops in structure formation of the universe. Finally the strong correlation between loops and long strings are revealed. A conclusion is given in section 5.5.

5.2 Topological analysis

Before examining cosmic strings as the external source, let us consider some fundamental features of sources with various possible topologies. Since the power spectrum is the most common statistic used to describe observations, we shall also use it to investigate the properties of the source. The types of external sources we shall come across can be roughly classified into four categories, according to their topologies: zero-, one-, two-, or three-dimensional objects. For the first three types, Ref. [102] gave the forms of their power spectra by Fourier transforming a set of randomly distributed points, straight filaments, and spherical shells:

$$\mathcal{P}_0(k) = \frac{V^2}{8\pi^3} , \tag{5.1}$$

$$\mathcal{P}_1(k) = \frac{V^2}{8\pi^3}\frac{\tan^{-1}(kc_*)}{kc_*} , \tag{5.2}$$

$$\mathcal{P}_2(k) = \frac{V^2}{8\pi^3}\frac{1}{(1+4k^2R_*^2)} , \tag{5.3}$$

where V is the volume of the box to be Fourier transformed, and the filament length c and the shell radius R of each objects are drawn from the probability distribution functions (PDF)

$$p_1(c) = \frac{c^2}{2c_*^3}\exp(-\frac{c}{c_*}) , \qquad \text{for filament length } c, \tag{5.4}$$

$$p_2(R) = \frac{R^2}{2R_*^3}\exp(-\frac{R}{R_*}) , \qquad \text{for shell radius } R. \tag{5.5}$$

These distributions lead to the rms amplitudes $\langle c\rangle = 2\sqrt{3}\,c_*$ and $\langle R\rangle = 2\sqrt{3}\,R_*$. Here, we further consider the fourth type: the three-dimensional objects. By Fourier transforming a set of randomly distributed spheres with a radius drawn from a PDF

$$p_3(r) = \frac{r^6}{720 r_*^7} \exp(-\frac{r}{r_*}) \ , \tag{5.6}$$

which gives an rms radius $\langle r \rangle = 2\sqrt{14}\, r_*$, we obtain the power spectrum for a set of three-dimensional objects as:

$$\mathcal{P}_3(k) = \frac{V^2}{8\pi^3} \frac{(1 + 0.8k^2 r_*^2)}{(1 + 4k^2 r_*^2)^3} \ . \tag{5.7}$$

Thus (5.1), (5.2), (5.3) and (5.7) represent a complete set for the topological analysis of the power spectrum of any possible non-fractal sources. We can see that for sufficiently small k, they all give a white-noise power spectrum $\mathcal{P}(k) = V^2/8\pi^3 \propto k^0$. This is because on lengthscales sufficiently larger than their characteristic scales $2\sqrt{3}c_*$, $2\sqrt{3}R_*$ and $2\sqrt{14}r_*$, all objects are essentially point-like. On the other hand, at larger k, they give the characteristic behaviours of zero-, one-, two- and three-dimensional objects: $\mathcal{P}(k) \propto k^0, k^{-1}, k^{-2}$ and k^{-4} respectively.

Knowing this, we can then debug and quantitatively test our simulation code by using an artificial network which has certain assigned topology. In Figure 5.1, we show the power spectra of randomly distributed zero-, one-, two-, and three-dimensional objects. They agree very well with the analytic behaviours of (5.1), (5.2), (5.3) and (5.7). Thus we see that on small scales, the higher the dimension that objects are, the steeper the slope of their power spectrum will be. With a fixed overall mean density inside the volume V, this means that objects of higher dimension will have less power on small scales. In general, this is because the mass of higher dimensional objects is less concentrated, so that they have less power towards small scales. Therefore, by looking at the power spectrum of a field and, in particular its turn-over scales, we can recognise the topology of objects contained in the field, as well as at what scale they can be identified. This is the technique which we shall use to help interpret the power spectra induced by cosmic string networks.

5.3 Long-string networks

5.3.1 The string source

The external source under consideration here is a network of cosmic strings. For local gauge strings, the microphysical width of the string will be many orders of magnitude smaller than its typical curvature radius, so that we can take the zero thickness limit. Therefore, the cosmic strings which dynamically source the subsequent perturbations $\delta^{\rm S}(\mathbf{x},\eta)$ in (4.3) have spacetime trajectories which we can represent as $x_{\rm s}^\mu = (\eta, \mathbf{x}_{\rm s}(\sigma,\eta))$, where σ is a spacelike parameter labelling points along the string. As shown in section 2.3.3, the stress energy tensor of cosmic strings can be derived from the Nambu action as [38]

$$\Theta_{\mu\nu}(\mathbf{x},\eta) = \mu \int d\sigma (\epsilon \dot{x}_{\rm s}^\mu \dot{x}_{\rm s}^\nu - \epsilon^{-1} {x'}_{\rm s}^\mu {x'}_{\rm s}^\nu) \delta^{(3)}(\mathbf{x} - \mathbf{x}_{\rm s}(\sigma,\eta)), \tag{5.8}$$

where μ is the string linear energy density, a dot and a prime represent derivatives with respect to η and σ respectively, $\epsilon = [{\mathbf{x}'_{\rm s}}^2/(1 - \dot{\mathbf{x}}_{\rm s}^2)]^{1/2}$, and we have taken $\dot{\mathbf{x}}_{\rm s} \cdot \mathbf{x}_{\rm s}{}' = 0$. It is then straightforward to compute Θ_+ in (4.1) as

$$\Theta_+(\mathbf{x},\eta) = \Theta_{00} + \Theta_{ii} = 2\mu \int d\sigma\, \epsilon \dot{\mathbf{x}}_{\rm s}^2\, \delta^{(3)}(\mathbf{x} - \mathbf{x}_{\rm s}(\sigma,\eta)). \tag{5.9}$$

This source term can be calculated directly from the string network which was evolved using the AS string simulation [103]. These particular simulations have been shown to be in good quantitative agreement with the BB simulation, also of high resolution [78, 79]. Large-scale parallelized simulations were performed on the COSMOS supercomputer, a Silicon Graphics Origin2000 with 20 Gbytes main memory. Dynamic ranges exceeding one thousand expansion times were feasible because of the simulation size and by using a modified 'point-joining' algorithm which maintained fixed comoving resolution [104]. In string simulations, sampling points are used to trace cosmic strings. We thus invoke the "cloud in cell" (CIC) method to assign the sampling points of strings onto a comoving lattice, which is then fast Fourier transformed (FFT). To compensate for the smoothing effect induced by the CIC method, we multiply this Fourier transformed quantity on each grid by the following sharpening function [105]

$$W(\mathbf{k}) = \frac{(k/2)^2}{\sin^2(k_1/2) + \sin^2(k_2/2) + \sin^2(k_3/2)} \ , \tag{5.10}$$

where $\mathbf{k} = (k_1, k_2, k_3)$ and $k = |\mathbf{k}|$. This scheme has been verified against sets of zero-, one-, two- and three-dimensional objects as we have seen in Figure 5.1.

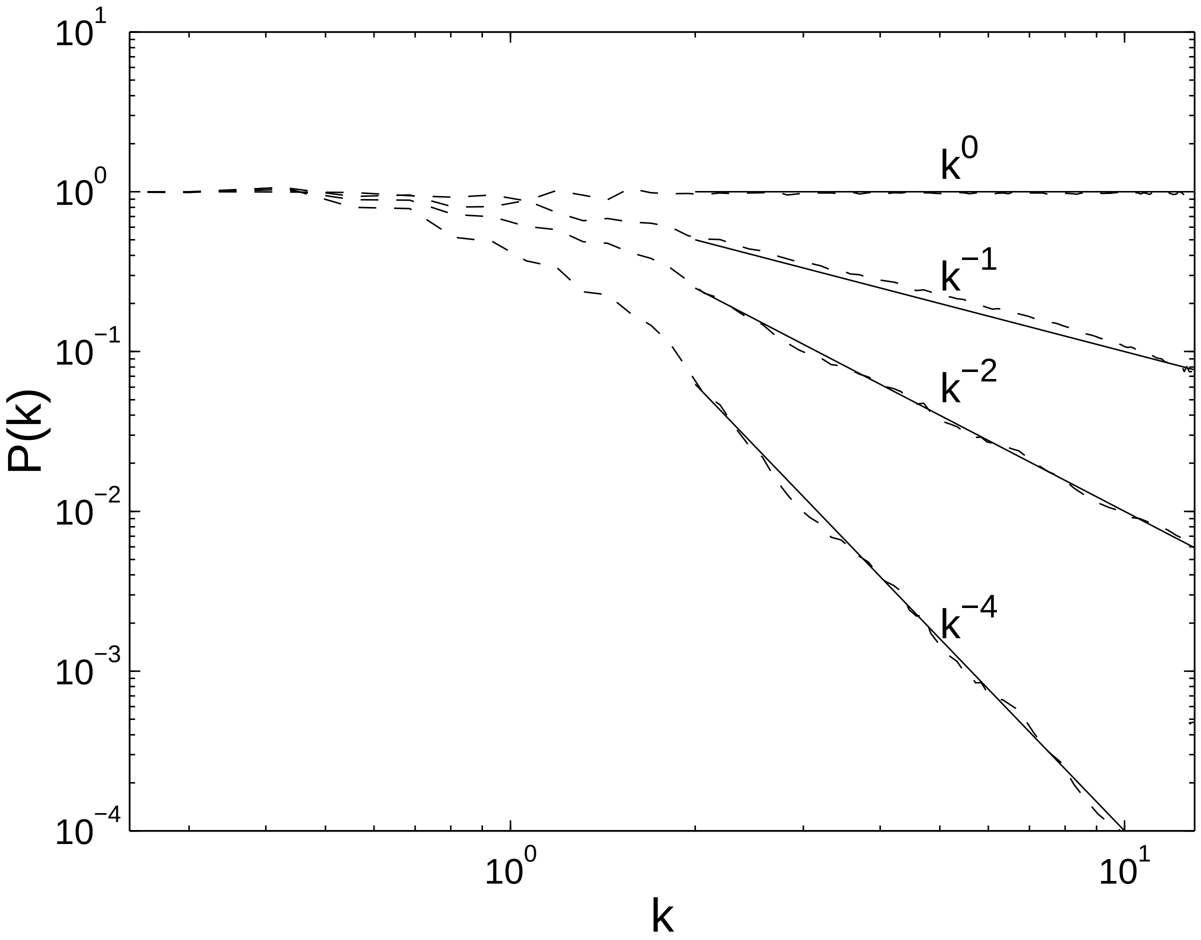

Figure 5.1: The power spectra of randomly distributed zero-, one-, two- and three-dimensional objects (dashed lines, downwards). Their characteristic sizes $\langle c \rangle$, $\langle R \rangle$ and $\langle r \rangle$ (see text) have been normalized to 2π, which corresponds to $k \sim 1$. We have also used (5.10) to deconvolve the lattice effect. These spectra agree very well on small scales with those solid lines, which have exact slopes of 0, -1, -2 and -4 respectively. On large scales, they all give white noise $\mathcal{P}(k) \propto k^0$.

We shall divide the string network into two parts: long strings and small loops. The latter is defined by taking those closed loops whose lengths are smaller than one tenth of the horizon size. We have numerically verified that a different threshold around this choice for separating long strings and loops will give rise to negligible differences in the resulting energy evolution in the long strings and loops. This is because the loop sizes are typically at least two orders below the horizon size in the scaling regime, and there are very few loops around the threshold we choose here. In the rest of this section, we shall investigate the properties of a long-string network, and return to discuss the role of loops in the next section.

5.3.2 Initial condition vs. scaling regime

We evolve the string network with Vachaspati-Vilenkin initial conditions [73] for a dynamic range of three in conformal time, and then compare the resulting power spectrum for the long strings with the initial power spectrum of the network. Figure 5.2 shows the initial and the final long-string configurations, as well as their power spectra. As one can clearly see, the power spectrum of the initial network has many periodic spikes, while that of the scaling network is very smooth. This phenomenon may be explained as follows. The initial network is a random walk with a correlation length ξ along the x, y or z directions, so its topology is very similar to that of a sampling function (the periodic Dirac Delta)

$$\Pi(x) = \sum_{n=-\infty}^{\infty} \delta(x - n\xi) \; , \tag{5.11}$$

where n is an integer. We know that its Fourier transform is also a sampling function. Hence, the power spectrum of this string network has periodic peaks, whose interval length should relate to the correlation length as $\Delta k = 2\pi/\xi$, as indeed we see in Figure 5.2. Due to the three-dimensional structure, there are also other minor peaks corresponding to diagonals such as $2\sqrt{2}\pi/\xi$, $2\sqrt{3}\pi/\xi$, etc. With a correlation length $\xi = \eta/8$, we also note that at large k ($\gtrsim 16\pi/\eta$) the power spectrum has an overall slope of -1 due to the line-like topology of cosmic strings; at small k ($\lesssim 16\pi/\eta$), the power spectrum turns flat because of the point-like nature of strings on lengthscales well above the correlation length (see section 5.2).

On the other hand, the scaling string network in Figure 5.2 shows different properties. Its power spectrum is much smoother than the previous case. This is because in the scaling network, the periodicity of the strings has no preferred directions (i.e. the x, y and z directions in the previous case), and therefore the periodic peaks degenerate to only one broad peak that corresponds to the mean spacing of strings. This smoothness of the power spectrum can be employed to ensure that when the structure formation simulation starts, the string network is already in the scaling regime so that the initial configuration of the network will have no effect on the subsequent results. In addition, we note that instead of turning from k^0 to k^{-1} very smoothly at the correlation scale as it does for randomly distributed filaments (see Figure 5.1), the power spectrum here has a broad peak on the turn over scale. This illustrates the fact that in the scaling regime, a real string network is not like random walk around the scale ξ, but has a configuration in which all strings remain at a roughly constant spacing (which is about the correlation length ξ) from each other, so that the power on the scale corresponding to this spacing is amplified to form a broad peak. We have verified that this broad peak appears at $k_\xi \approx 20/\eta$ persistently from deep in the radiation era through to deep into the matter era, with an amplitude about three times larger than the large-scale white noise. Thus the comoving mean spacing among strings is $\xi \approx 2\pi/(20/\eta) \approx \eta/3$. Consequently, previous work based on modelling cosmic strings with random filaments will have underestimated the resulting power spectrum by up to a factor of three (e.g. Ref. [106, 107]), because this additional power above the large-scale white noise will be missing. This underestimate by using toy models with random filaments becomes more apparent on smaller scales because, on large scales, the shape of the resulting density power spectrum is more dominated by the compensation scale, rather than the correlation length.

We also notice in Figure 5.2 that the slope of the spectrum in the scaling regime is not exactly -1 on small scales, but -1.13. This is because strings have a fractal contribution; they are very wiggly on small scales and therefore have an overall dimension slightly greater than one, so as to steepen the small-scale slope.

5.3.3 Correlation time and the UETC

Another important property of the string network is the correlation time. Because the strings' configurations are uncorrelated when separated by a sufficiently long time, we can define a correlation time η_c as the time period above which the strings' configurations become completely uncorrelated, i.e. in terms of the unequal time correlator (UETC)

$$\langle \widetilde{\Theta}_+(\mathbf{k},\eta)\widetilde{\Theta}_+(\mathbf{k},\eta')\rangle \approx 0 \quad \text{for } |\eta-\eta'| > \eta_c \, . \tag{5.12}$$

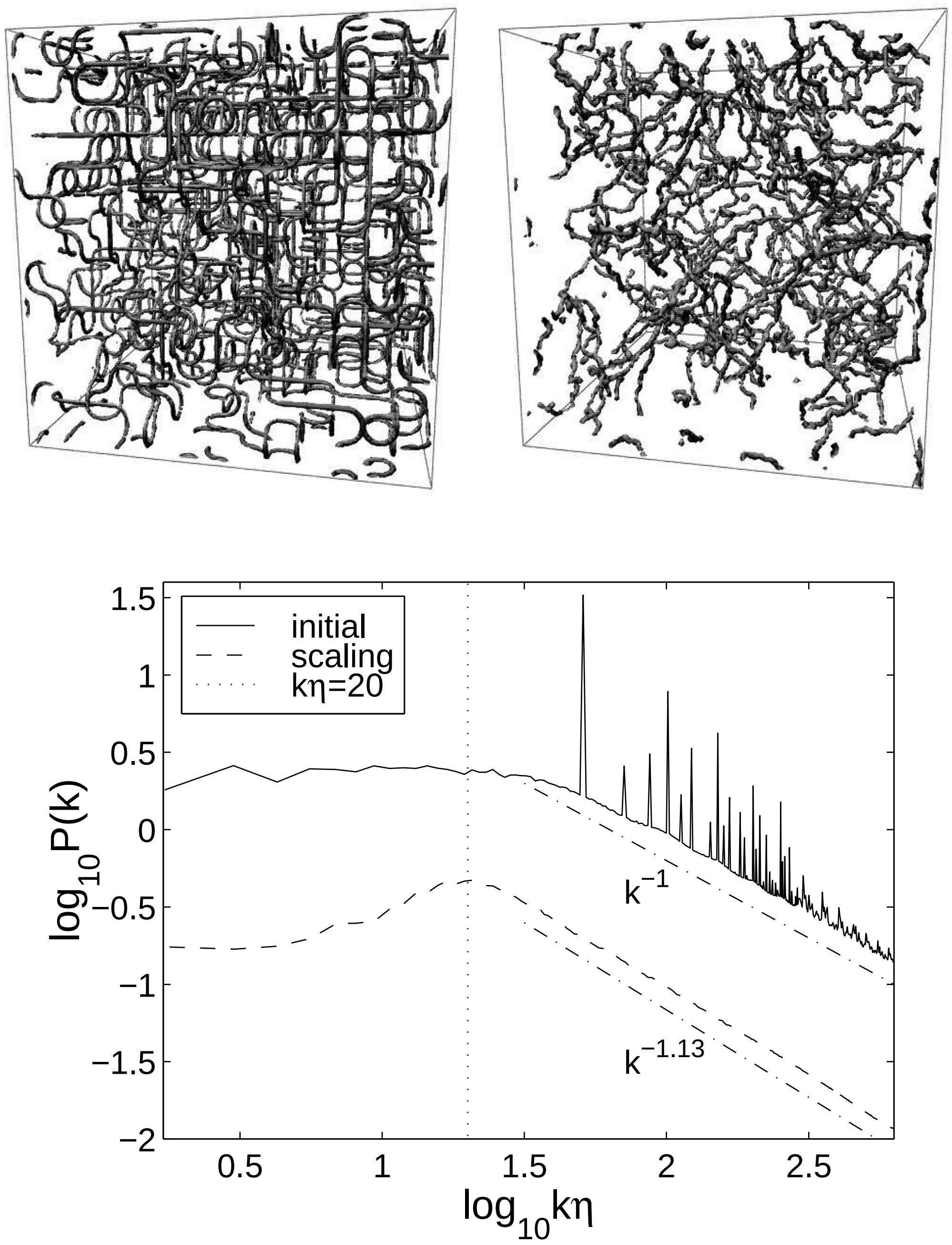

Figure 5.2: The top-left is an initial network of cosmic strings, while the top-right is a long-string network after a dynamic range of three in conformal time, when strings are in the scaling regime. At the bottom are their power spectra. The first peak in the initial spectrum appears at $k_\xi = 16\pi/\eta$, which accurately corresponds to the initial comoving correlation length $\xi_{\rm com} = \eta/8$ set by hand. The turnover scale in the scaling regime is clearly $k_\xi = 20/\eta$, and the slope of the spectrum turns from -1 in the initial network to -1.13 in the scaling regime. The normalization here is arbitrary.

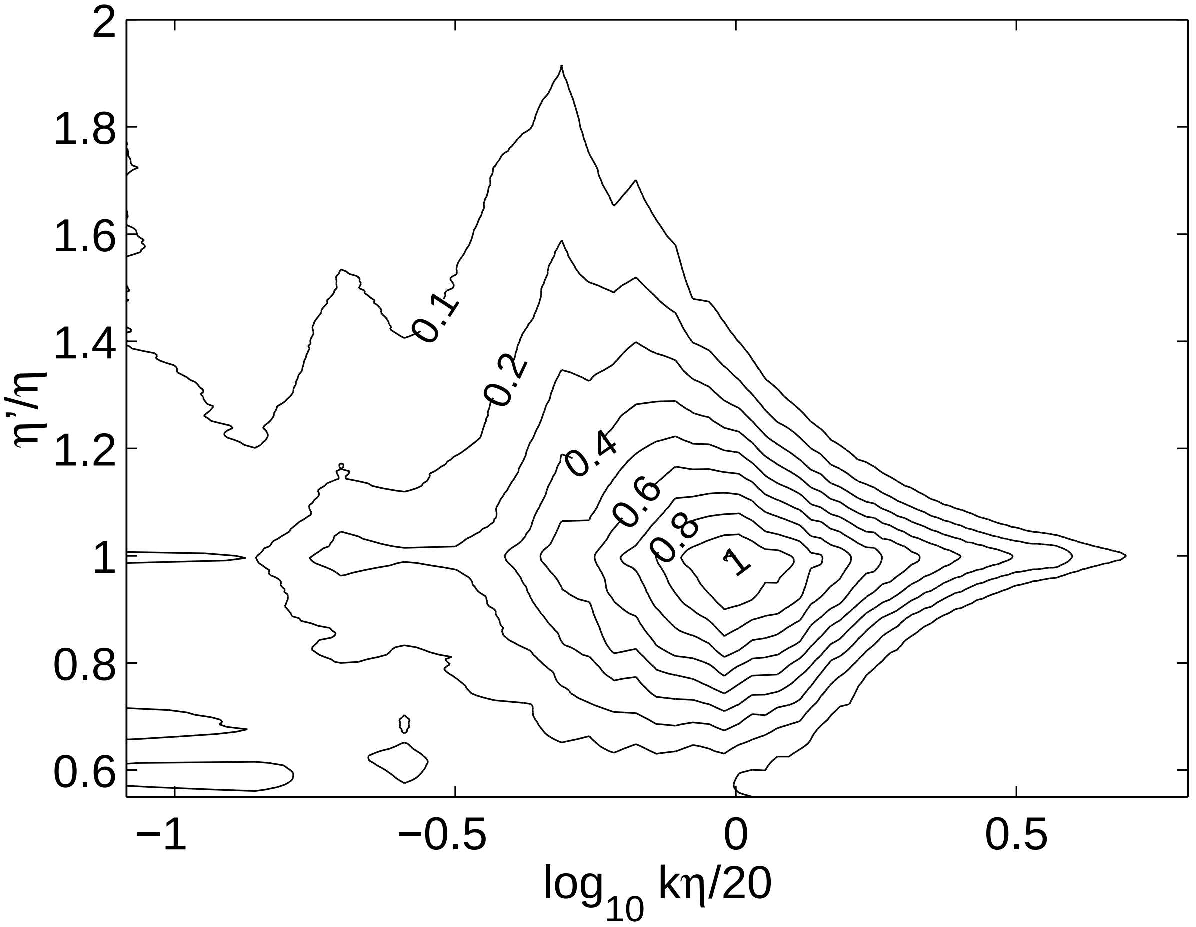

Figure 5.3: The UETC $\langle\widetilde{\Theta}_+(\mathbf{k},\eta)\widetilde{\Theta}_+(\mathbf{k},\eta')\rangle$ with a central time at η, and η' goes from 0.6η to 2η. The amplitude of the UETC has been normalized to with a maximum of 1, and the contours are taken with an interval of 0.1.

Figure 5.3 shows this behaviour for one string realization, but the same has been observed qualitatively from deep in the radiation era through to deep in the matter era. We see that on smaller scales, inside the horizon, the correlation time is shorter; on the horizon scale the correlation time η_c is about the same order as the conformal time η, although the UETC is dominated within a dynamic range $\eta' = (1 \pm 0.2)\eta$ by taking the half-maximum threshold, i.e. the "0.5" contour in Figure 5.3. Thus in a quantitative analysis, an accurate calculation of UETC's requires a dynamic range of at least 4 in conformal time in the scaling regime. This is equivalent to a dynamic range of 16 in radiation-era physical time, and 64 in matter-era physical time. So far in the literature, only our simulations have achieved this criterion for local cosmic strings, while keeping a resolution equivalent to 1000^3 points in a comoving box. This dynamic range/resolution also far exceeds that achieved to date for global defects. Each of our simulations typically has a dynamic range of more than 10 in conformal time. We also observe in Figure 5.3 that there is a hot spot located at around $\eta' = \eta$ and $k \approx 20/\eta$, which is exactly the correlation wavenumber k_ξ we have previously seen in the string source power spectrum. If we take a slice through $\eta' = \eta$ in Figure 5.3, then the profile of this slice gives the scaling spectrum in Figure 5.2, with the broad peak corresponding to the UETC hot spot here.

5.3.4 Thickness of long string paths

Figure 5.4 shows the integration of the UETC over time. This quantity reflects the topology of the path swept out by strings. Since a two-dimensional object will have a small-scale slope of -2, we expect such a quantity to behave in the same way. However, as we can see from the figure, the small-scale slope is not exactly -2, but -2.25. This is because the wakes swept out by strings are also wiggly due to the wiggliness of strings, so that their small-scale slope is further steepened from -2 to -2.25. We investigate the level of such a wiggliness in wakes, and it corresponds equivalently to planar wakes with a thickness of one tenth of their mean spacing.

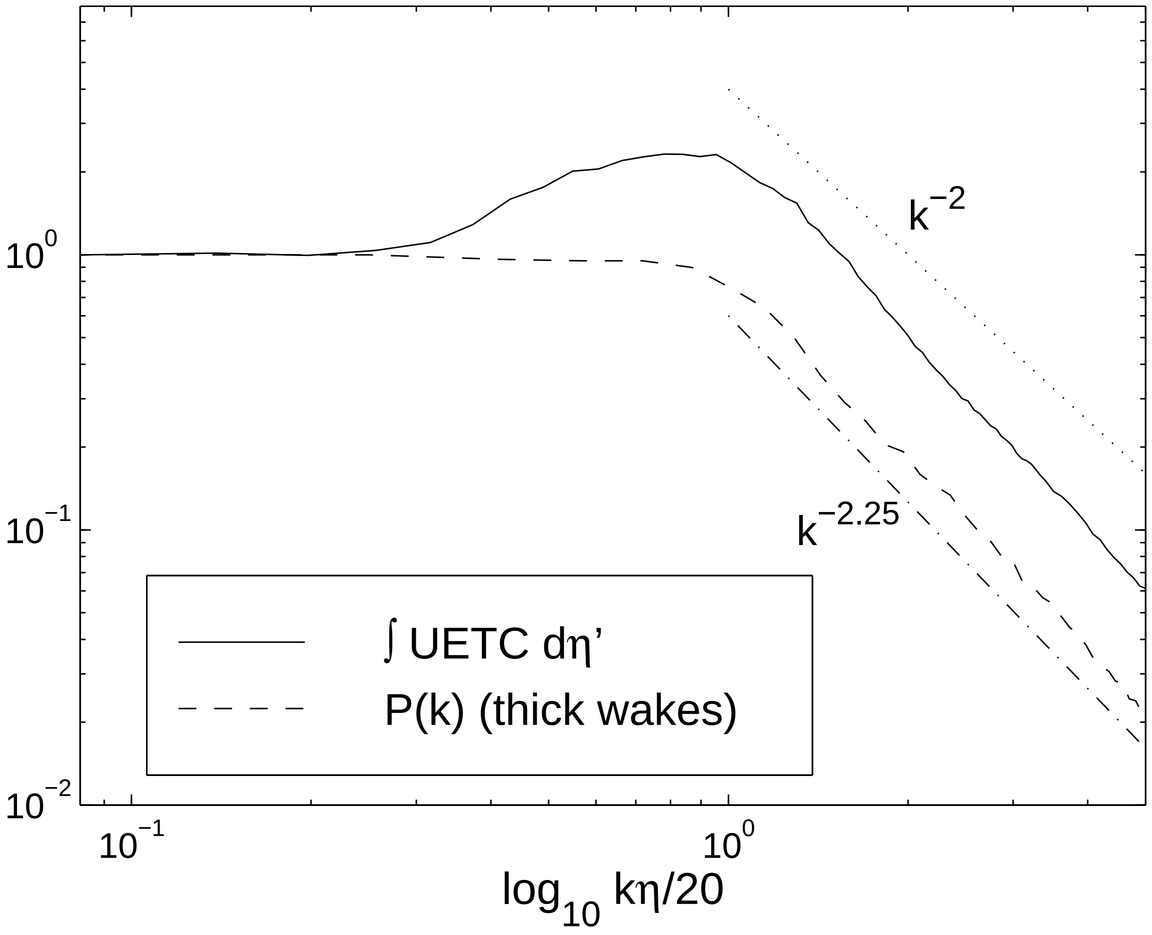

Figure 5.4: The integrated UETC, $\int \langle \widetilde{\Theta}_+(\mathbf{k},\eta)\widetilde{\Theta}_+(\mathbf{k},\eta')\rangle d\eta'$ (solid line), and the power spectrum of wakes with a thickness of 1/10 of their correlation length (dashed line). The normalization is arbitrary. The dotted and dot-dashed lines have exact slopes of -2 and -2.25 respectively.

This implies the wiggles of cosmic string wakes are about one tenth of their correlation length, i.e. $\xi/10 \approx \eta/30$.

5.3.5 Evolution of long-string energy density

Another important aspect of cosmic string evolution is the behaviour of the radiation-matter transition. The key observation for simulations spanning this period is the very slow relaxation to the matter era scaling density. Figure 5.5 shows the evolution of long-string energy density ρ_∞. A good analytic fit to the evolution of $\zeta(\eta) = \rho_\infty t^2/\mu$ is found to be

$$\zeta(\eta) = 14.7 - \frac{11.4}{1 + (4\eta_{\rm eq}/\eta)^{1.1}}\,. \tag{5.13}$$

This agrees well with the results by BB [78, 79]:

$$\zeta = \begin{cases} 13 \pm 2.5, & \text{radiation era,} \\ 3.5 \pm 1, & \text{matter era.} \end{cases} \tag{5.14}$$

When compared with AS simulations of higher resolution, we find that our values here are about 10% higher, while the overall shape remains unchanged. This is due to a lower kink density in our case, resulting in a shorter total length of strings. We also note from Figure 5.5 that the relaxation process extends well beyond recombination (for $h = 0.7$, $\log_{10}(\eta_{\rm rec}/\eta_{\rm eq}) \approx 0.76$ and $\log_{10}(\eta_0/\eta_{\rm eq}) \approx 2.4$). This has important implications for large-scale structure and CMBR anisotropies. As we shall show in the next chapter, due to the slow relaxation of the long-string energy density during the radiation-matter transition, the resulting string-seeded matter perturbations are dominated by the asymptotic value of the long-string energy density in the radiation era, rather than the matter era value as previously assumed.

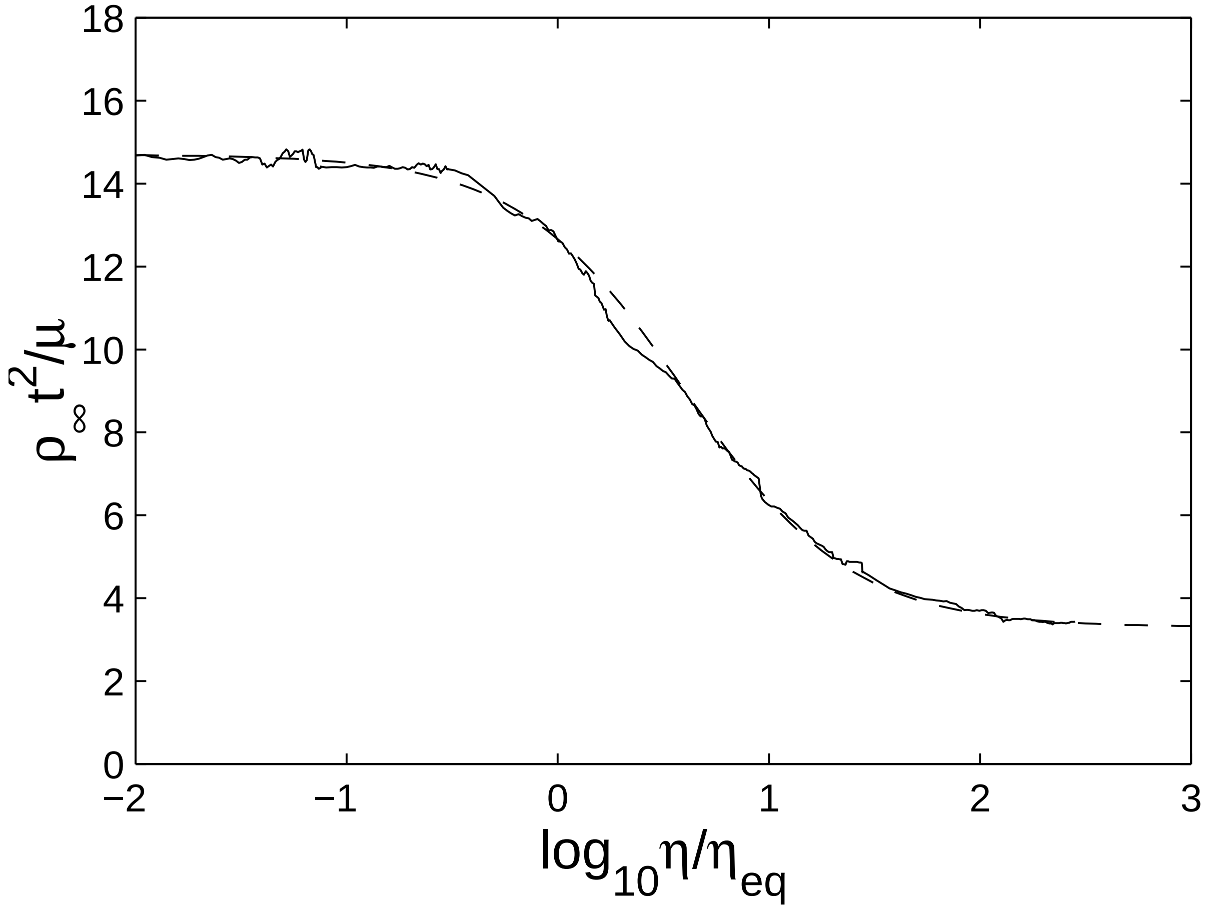

Figure 5.5: The evolution of long-string energy density observed from high-resolution simulations (solid), and the analytic fit (5.13) (dashed). It is clear that $\rho_\infty t^2/\mu$ drops very slowly after $\eta_{\rm eq}$, and at $\eta_{\rm rec}$ it is still about 1.5 times larger than its matter-era asymptotic value.

5.4 Cosmic string loops

The role of small loops produced by the string network has evolved from a proposed one-to-one correspondence between loops and cosmological objects [108] through to a completely subsidiary role relative to the wakes swept out by long strings [109]. This dethronement of loops was a result of numerical studies which showed that the average loop size $\bar{\ell} = \alpha t$ was much smaller than the horizon, $\bar{\ell} \ll d_{\rm H}$ [78, 103]; they may even be as small as the lengthscale set by gravitational backreaction, for which $\alpha \sim 10^{-4}$ is the value appropriate for GUT-scale strings [38]. Add the high ballistic loop velocities observed $\bar{v}_{\rm L} \approx c/\sqrt{2}$ and it was not surprising that these tiny loops have been assumed to be more or less uniformly distributed and hence a negligible source relative to the long-string network [102]. Nevertheless, small loops always make up a significant fraction of the total string energy density at any one time and, as we shall demonstrate in Chapter 6, loop-induced inhomogeneities are considerable if their lifetime is not much smaller than the Hubble time. By properly incorporating these loop perturbations, we have found their contribution relative to the long-string wakes to be almost comparable and also highly correlated with these wakes. We, therefore, first investigate the evolution of the loop distribution in this section.

In the simulations, we keep all the loops generated by the long-string network, and model those which are smaller than a fixed fraction of the horizon size as relativistic point masses. The effects of the evaporation of these loops into gravitational waves and the damping of the loop motion due to the expansion are also included. Note that here we assume the standard scenario with an evolving 'infinite' string network, but there are alternative models in which the initial string configuration consists entirely of loops (e.g. [110, 111, 112]). Our scheme should also be clearly distinguished from other recent work [113] and [114]. The former attempted to incorporate network decay products in the power spectrum of an additional fluid (with a variety of possible equations of state), while the latter endeavoured to argue that the primary energy loss mechanism for the network is through particle production rather than gravitational radiation from loops (refer to [115] on this point).

5.4.1 Energy transfer from long strings into loops

To begin with, let us consider some analytic properties of loops. As seen from equation (2.118), the Nambu equations of motion for cosmic strings in an expanding universe can be averaged to yield:

$$\frac{d\rho_\infty}{dt} + 2H(1+\langle v^2\rangle)\rho_\infty = -X_L, \tag{5.15}$$

where $\langle v^2\rangle$ is the mean square velocity of strings and $X_L \equiv c_l\rho_\infty/\xi$ is the transfer rate of energy density from long strings into loops. In the scaling regime we also know from equation (2.121) that the long-string energy density should scale with the background energy density as $d\rho_\infty/dt = -2\rho_\infty/t$. Substituting this into (5.15) to eliminate $d\rho_\infty/dt$ gives

$$\frac{tX_L}{\rho_\infty} = \begin{cases} (1-\langle v_{\rm r}^2\rangle) \sim 0.6 & \text{in radiation era,} \\ \frac{2}{3}(1-2\langle v_{\rm m}^2\rangle) \sim 0.2 & \text{in matter era,} \end{cases} \tag{5.16}$$

where $\langle v_{\rm r}^2\rangle \gtrsim \langle v_{\rm m}^2\rangle \sim 0.6$ [78, 103]. Both (2.121) and (5.16) provide a check for the scaling behaviour of long strings and loops in the cosmic string network simulations. Figure 5.6 shows the numerical evolution of X_L, starting from the radiation era through to the matter era. We can see that the expected amount of energy was converted into loops in our simulations so that X_L has the correct asymptotic behaviour given by (5.16). However, the typical loop-size (and consequently their lifetime) does not approach scaling so rapidly and is therefore larger than physically expected for most of the duration of our present simulations [78, 103].

5.4.2 Modelling gravitational decay and loop life time

To overcome this problem we rescale the loop lifetime in the following way. We know that the loops produced by a cosmic string network will decay into gravitational radiation, with a roughly constant decay rate $\Gamma G\mu^2$, where μ is the string linear energy density. Typically $\Gamma = 50-100$ with an average $\langle\Gamma\rangle \sim 65$ [80, 116]. In very high-resolution simulations, we observe a loop-production scale on which the loops are most effectively generated by the network, although this is only approached towards the end of the simulations used here. For both analytic and numerical simplicity, therefore, we make the reasonable assumption that the loop production is 'monochromatic', that is, all loops formed at the same time will have the same mass in the limiting scale-invariant distribution. We can write the initial rest mass of a loop formed at time t_* as

$$\bar{M}_L^* = \alpha\mu t_* \equiv f\Gamma G\mu^2 t_* \ . \tag{5.17}$$

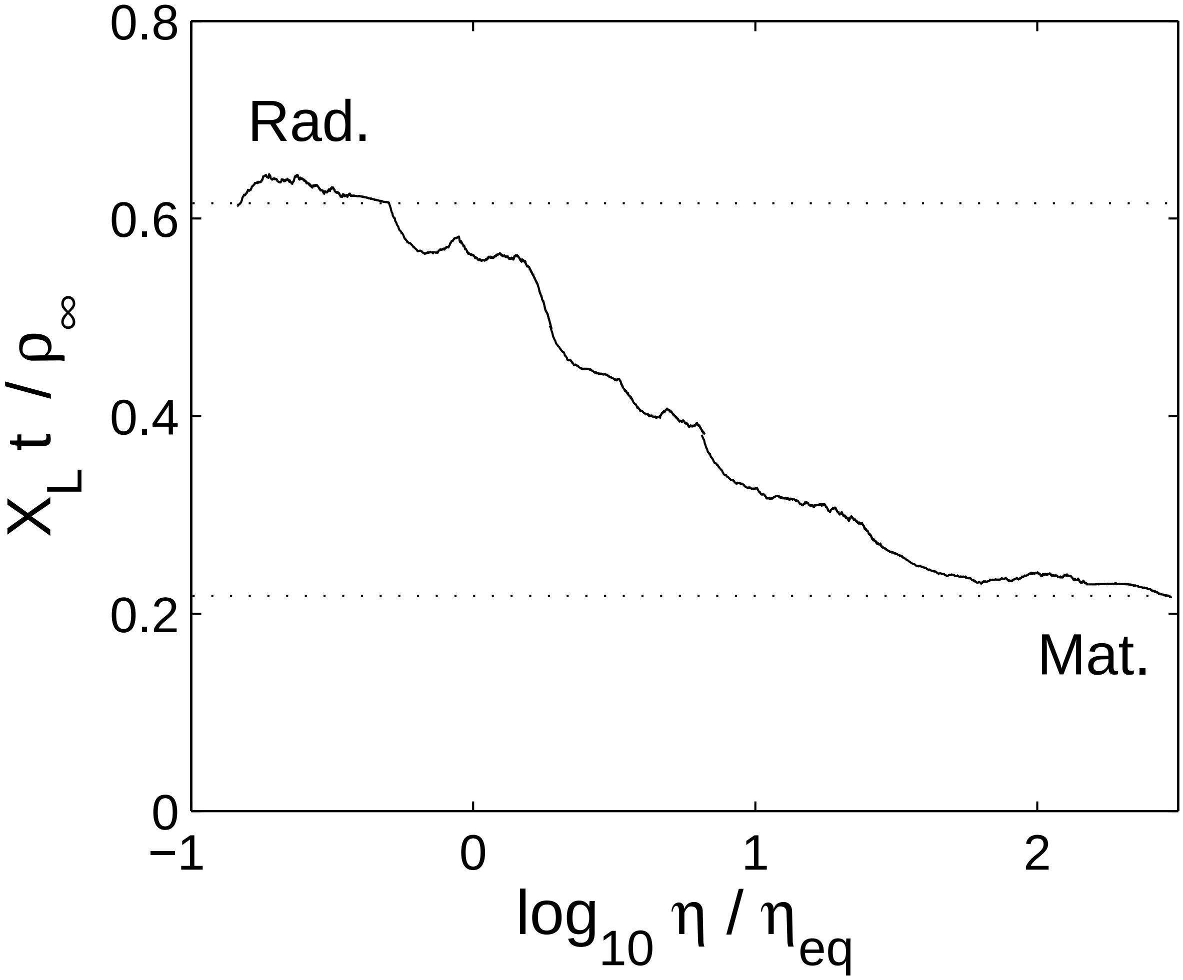

Figure 5.6: Evolution of $X_L(t)$. The dotted lines are the asymptotic values in the radiation and the matter eras in (5.16).

Here, the parameter $f = \alpha/\Gamma G\mu$ is expected to be of order unity if the size of the loops formed at the time t is determined by gravitational radiation back-reaction, which smoothes strings on scales smaller than $\Gamma G\mu t$. Thus the uncertainty in the average mass, and therefore the mean lifetime, of loops formed at a given time is quantified by the choice of the parameter f. We also find that the initial rms velocity of loops observed from the simulations is $\langle v_*^2\rangle^{1/2} \gtrsim 0.7c$ throughout all the regimes.

In the simulations we only impose the monochromatic assumption through f on the loop decay lifetimes, while keeping the original initial mass of each individual loop. Thus we have the rest mass of a loop formed at time t_* evolving as

$$M_L(t_*, t) = M_L^* W(t_*, t) \ , \tag{5.18}$$

where

$$W(t_*, t) = \begin{cases} 1 - \frac{t-t_*}{\tau(t_*)} & \text{for } t_* \leq t \leq t_* + \tau(t_*) \, , \\ 0 & \text{otherwise} \, . \end{cases} \tag{5.19}$$

Here $\tau(t_*) = f t_*$ is the modelled mean lifetime of loops produced at time t_* ($f = 2$, 3 implies the decay occurs in one horizon time in the radiation and matter eras respectively). The evolution of the loop energy density is then given by:

$$\begin{aligned} \rho_L(t) &= \int_0^t X_L(t') \left[\frac{a(t')}{a(t)}\right]^3 W(t', t) dt' \\ &\propto \begin{cases} f/t^2 & \text{for } f \ll 1, \\ \sqrt{f}/t^2 & \text{for } f \gg 1 \text{ (radiation era)}, \\ (\ln f)/t^2 & \text{for } f \gg 1 \text{ (matter era)}, \end{cases} \end{aligned} \tag{5.20}$$

where we have used the scaling behaviour (2.121) and (5.16). Consequently, the scaling of the power spectrum induced by loops in f should interpolate between f^2 and f (radiation era) or $(\ln f)^2$ (matter era). We notice in (5.20) that we have ignored the effect of loop velocity redshifting due to the expansion of the Universe, which causes a change in the effective mass. Because loops are formed with relativistic velocities, we expect this damping mechanism to have the strongest effect for $f \gg 1$, but to be negligible for $f \ll 1$.

5.4.3 Velocity damping and rocket effect

If a loop formed at time t_* has an initial physical velocity $\mathbf{v}_*$, then its trajectory in physical space accounting for the expansion of the Universe is then given by:

$$\mathbf{x}(t) = \mathbf{x}(t_*) + a(t) \int_{t_*}^{t} \frac{\mathbf{A}}{\sqrt{a^2(t') + A^2}} \frac{dt'}{a(t')} \, , \tag{5.21}$$

for $t \geq t_*$, where $\mathbf{A} = \gamma_* \mathbf{v}_* a_*$, $A = |\mathbf{A}|$ and $\gamma_* = (1 - |\mathbf{v}_*|^2)^{1/2}$. Here we have neglected the acceleration of loops due to the momentum carried away by the gravitational radiation, the so-called 'rocket effect'. A numerical calculation for several asymmetric loops shows that the rate of momentum radiation from an oscillating loop is

$$|\dot{\mathbf{P}}| = \Gamma_P G\mu^2 \ , \tag{5.22}$$

where $\Gamma_P \sim 10$ [117]. Combining with (5.18), one can show that this rocket effect will become important only when:

$$\frac{t}{t_*} \gtrsim 1 + \frac{f}{1 + \Gamma_P/(\Gamma\gamma v)} \, , \tag{5.23}$$

which affects only the final stages of the loop lifetime as long as $\Gamma_P/(\Gamma\gamma v) < 1$, or equivalently $v_{\rm crit} \gtrsim 0.15c$. For a typical $v_* \sim c/\sqrt{2}$, one requires a loop lifetime $\gtrsim 43t_*$ in the radiation era and $\gtrsim 17t_*$ in the matter era to redshift down to this critical velocity $v_{\rm crit}$ according to (5.21). Since the values of f we explore here are of order unity, it is a reasonable approximation to neglect the transfer of momentum due to gravitational radiation.

5.4.4 Strong correlation with long strings

With the treatment of (5.17) and the effects of (5.18) and (5.21) Figure 5.7 shows a loop spectrum with $f = 1$ in the scaling regime. As we can see, on large scales ($k\eta \lesssim 100$) the shape of the loop power spectrum is identical to that of long strings (see Figure 5.2). This is because loops trace the paths of long strings even after they are formed. We were able to verify this by calculating the correlation coefficient between long strings and loops in the scaling regime. This coefficient is always greater than 0.5 on large scales. This phenomenon has also been verified by observing animations of string evolution made from the simulations. On the other hand, on small

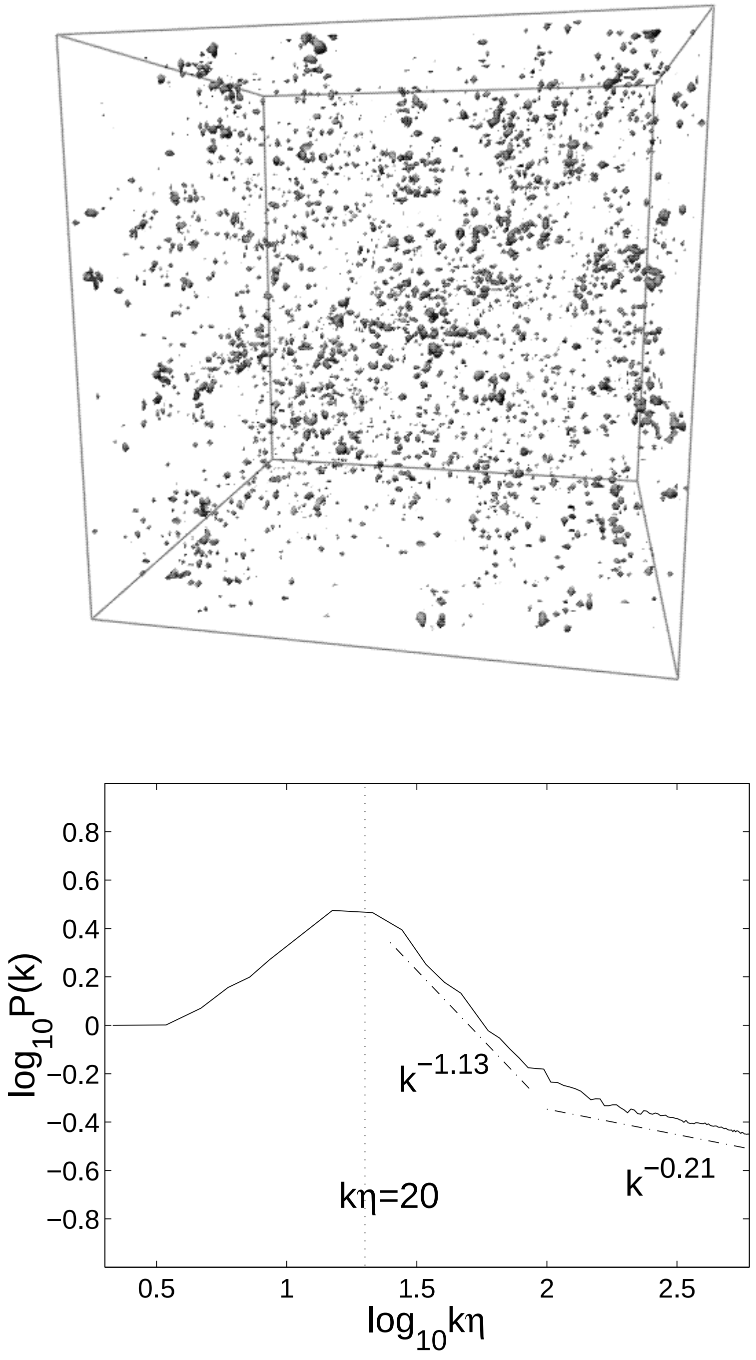

Figure 5.7: On the top is a snapshot of a loop distribution in the scaling regime. The size of each point reflects its mass. The box size length is equal to the horizon. On the bottom is their power spectrum with arbitrary normalization. The dot-dashed lines have slopes of -1.13 and -0.21. The dotted line indicates the correlation wavenumber $k_\xi = 20/\eta$ seen in the long-string network. On large scales ($k\eta \lesssim 100$), we see loops mimic long strings so that their spectral shape is identical to that of long strings (see Figure 5.2).

scales the loop spectrum is not exactly white noise. It has a slope of -0.21. This is because even on small scales loops are still not exactly point-like, but are correlated in some clumpy structures, so that their spectral slope is steepened as discussed in section 5.2. In section 6.4.6, we shall further explore the effect of loops on cosmic structure formation, in particular the manner in which the strong correlation between the long-string and loop distributions on large scales can significantly boost the resulting power spectrum of matter perturbations.

5.5 Conclusion

In this chapter, we began with a topological analysis using power spectra, and then investigated the properties of string networks, with separate studies of long strings and small loops. For long-string networks, we calculated the correlation length using power spectra $\xi \approx \eta/3$, the correlation time $\eta_c \approx \eta$, and the thickness of long string paths $\approx \xi/10 \approx \eta/30$. We also revealed the slow relaxation of long-string energy density during the radiation-matter transition. For loops, we investigated the evolution of their network by modelling them as relativistic point masses. Their gravitational decay and velocity damping due to the expansion of the universe were both taken into account. This is an improvement to the existing AS string simulation. In particular, we revealed the strong correlation between long-string and loop distributions, the cosmological implications of which for structure formation will be examined in the next chapter.

Chapter 6

Large-Scale Structure Seeded by Cosmic Strings

6.1 Introduction

The cosmic string scenario [38] predated inflation as a realistic structure formation model, but it has proved computationally much more challenging to make robust predictions with which to confront observations. Only recently has significant progress in understanding cosmic strings as seeds for LSS and CMBR anisotropies been achieved [1, 2, 5, 56, 106, 107, 113, 118, 119]. In Ref. [56], one specific quantitative scheme was proposed to study the power spectra induced by global defects. The authors employed the UETC's of defect sources as seeds for the Einstein-Boltzmann solver 'CMBFAST' [85]. Their results were later confirmed independently [120]. This seems to indicate that the use of UETC's is a promising way of calculating the power spectra for both LSS and CMB, although for cosmic strings the slow relaxation of the long-string energy density over the radiation-matter transition (see the previous chapter) may complicate the proposed treatment. In this chapter, we shall concentrate on local cosmic strings and describe a self-consistent method to study string-seeded matter perturbations with either a CDM or an HDM background. This method is based on the fluid approximation developed in Chapters 3 and 4. It will allow us not only to calculate the LSS power spectra but also to investigate the distribution of matter density perturbations. The framework developed here is suitable not only for studying cosmic strings, but also for investigating any other matter sources behaving stiffly, that is, whose evolution is largely independent of the background matter and any inhomogeneities in the universe. This direct numerical approach is not without some weaknesses which we shall discuss, but it marks a considerable quantitative advance by incorporating important aspects of the relevant physics not included in previous treatments.

The present work relies on high-resolution AS simulations of a cosmic string network [103] with a dynamic range extending from well before the matter-radiation transition through to deep into the matter era. The resulting power spectrum of linear matter perturbations, $\mathcal{P}(k)$, and the mass fluctuation amplitude at $8h^{-1}$Mpc, σ_8, are calculated. We shall propose a semi-analytic model, which can reproduce the $\mathcal{P}(k)$ of our high-resolution simulations. We shall also investigate the dependence of $\mathcal{P}(k)$ on the curvature K and the cosmological constant Λ. In particular, for string+CDM models with the shape parameter $\Gamma = 0.74 \times 10^{\pm 0.1}$ (95% confidence level), we show that the overall shape of $\mathcal{P}(k)$ is consistent within uncertainties with those currently inferred from galaxy surveys. With $\Omega_{\rm B0} = 0.05$, a likelihood analysis for finding the best fit of the string+CDM spectrum shape to the observation gives $\Omega_{\rm m0}h = 0.14 \times 10^{\pm 0.1}$ (95% confidence level). On the other hand, the HDM scenario with cosmic strings seems to require a large scale-dependent biasing in order to be consistent with observations. The effects from small loops will be also investigated. We find that if the loop lifetime is not much smaller than the Hubble time, the string-induced CDM power spectrum may be further boosted by a factor of three.

The structure of the chapter is as follows: in section 6.2, we recall the formalism of causal perturbation theories developed previously, and describe the numerical scheme we shall invoke, including an approximation for the compensation mechanism. Section 6.3 illustrates the approximation schemes used to cope with an HDM background, the limited dynamic range of our simulations, and the effect of a non-zero curvature or a non-zero cosmological constant. In section 6.4, the semi-analytic model proposed in sections 6.3.2 and 6.3.3 is first verified to high accuracy in its ability to reproduce the simulation results, and is thus shown to be capable of compensating for the lost dynamic range in the simulations. The unimportance of the late-time non-scaling behaviour of strings in open or Λ-models is addressed. We then use this semi-analytic model to investigate the effect of radiation, the uncertainty resulting from our compensation approximation, the string+CDM power spectrum and its dependence on the curvature and the cosmological constant, the effect of loops, and the effect of baryons. Empirical formulae describing our main results will be presented here, as well as a discussion of the main uncertainties which affect these results. Finally, a conclusion is given in section 6.5.

6.2 Perturbation equations and the numerical method

We consider density perturbations about a flat FRW model which are causally sourced by an evolving external-source network with energy-momentum tensor $\Theta_{\mu\nu}(\mathbf{x},\eta)$. In the synchronous gauge with a cosmological constant, the linear evolution of the radiation and CDM perturbations, δ_r and δ_c respectively, are given by equations (4.1) and (4.2):

$$\ddot{\delta}_\mathrm{c} + \frac{\dot{a}}{a}\dot{\delta}_\mathrm{c} - \frac{3}{2}\left(\frac{\dot{a}}{a}\right)^2 [\Omega_\mathrm{c}\delta_\mathrm{c} + (2+R)\Omega_\mathrm{r}\delta_\mathrm{r}] = 4\pi G\Theta_+, \tag{6.1}$$

$$\ddot{\delta}_\mathrm{r} - \frac{4}{3}\ddot{\delta}_\mathrm{c} + \frac{\dot{R}}{1+R}(\dot{\delta}_\mathrm{r} - \frac{4}{3}\dot{\delta}_\mathrm{c}) - \frac{1}{3(1+R)}\nabla^2\delta_\mathrm{r} = 0. \tag{6.2}$$

In Chapter 4, it proved useful to split the linear matter perturbations into the initial (I) and subsequent (S) parts [81],

$$\delta_N(\mathbf{x},\eta) = \delta_N^\mathrm{I}(\mathbf{x},\eta) + \delta_N^\mathrm{S}(\mathbf{x},\eta)\ , \quad N = \mathrm{c},\,\mathrm{r}, \tag{6.3}$$

with initial conditions $\delta_N^\mathrm{I} = \delta_N$, $\dot{\delta}_N^\mathrm{I} = \dot{\delta}_N$ at η_i, and $\delta_N^\mathrm{S} = \dot{\delta}_N^\mathrm{S} = 0$ at $\hat{\eta}$. In the case of topological defects, the initial perturbations $\delta^\mathrm{I}(\mathbf{x},\eta)$ depend on the defect configuration at the defect formation time η_i, because ultimately the formation of defects creates under-densities in the initially homogeneous background out of which they are carved. The subsequent perturbations $\delta^\mathrm{S}(\mathbf{x},\eta)$ are those actively generated by the defects themselves at $\hat{\eta} > \eta_\mathrm{i}$. Because defects induce isocurvature perturbations, $\delta^\mathrm{I}(\mathbf{x},\eta)$ must compensate $\delta^\mathrm{S}(\mathbf{x},\eta)$ on comoving scales $|\mathbf{x}-\mathbf{x}'| > \eta$ to prevent acausal fluctuation growth on superhorizon scales. We have discussed this in some detail in Chapter 4. As seen from equation (4.9), the system of equations (6.1) and (6.2) can be solved for the subsequent perturbations $\delta^\mathrm{S}(\mathbf{x},\eta)$ by using a discretized version of the integral equation with Green functions:

$$\widetilde{\delta}_\mathrm{c}^\mathrm{S}(\mathbf{k},\eta_0) = 4\pi G\int_{\eta_\mathrm{i}}^{\eta_0} \widetilde{\mathcal{G}}^\mathrm{cs}(k;\eta_0,\hat{\eta})\widetilde{\Theta}_+(\mathbf{k},\hat{\eta})d\hat{\eta}. \tag{6.4}$$

Here the Green function $\widetilde{\mathcal{G}}^\mathrm{cs}(k;\eta_0,\hat{\eta}) = \widetilde{\mathcal{G}}_4^\mathrm{c}(k;\eta_0,\hat{\eta})$ can be either (4.45) combined with (4.52), or numerically obtained from equations (6.1) and (6.2).

However, it is a very substantial numerical challenge to evolve the initial and subsequent perturbations induced by cosmic strings in (6.3) such that they accurately cancel on superhorizon scales by the present day η_0. Long-string network scaling [78, 79, 103] entails the copious production of enormous numbers of tiny loops whose evolving distribution and decay must be carefully followed as described in section 5.4. For the large dynamic range required for the present study, however, we have as a first step adopted the compensation factor approximation suggested in a semi-analytic context in Ref. [121]. To implement this, we accurately evolved the string network numerically, and then multiplied the Fourier transform of the resulting stress energy $\widetilde{\Theta}_+(\mathbf{k},\eta)$ by a cut-off function $\widetilde{F}(k,\eta)$ given by

$$\widetilde{F}(k,\eta) = \left[1 + \left(\frac{k_\mathrm{c}(\eta)}{k}\right)^2\right]^{-1}, \tag{6.5}$$

where $k_\mathrm{c}(\eta)$ is the compensation wave number. This results in the correct k^4 fall-off in the power spectrum at large wavelengths above the compensation scale $k_\mathrm{c}^{-1} \sim \eta$, although the compensation scale may not be precisely represented as we shall discuss. Thus we obtain

$$\widetilde{\delta}_\mathrm{c}(\mathbf{k},\eta_0) = \widetilde{\delta}_\mathrm{c}^\mathrm{I}(\mathbf{k},\eta_0) + \widetilde{\delta}_\mathrm{c}^\mathrm{S}(\mathbf{k},\eta_0) \approx 4\pi G\int_{\eta_\mathrm{i}}^{\eta} \widetilde{\mathcal{G}}^\mathrm{cs}(k;\eta_0,\eta')\widetilde{\Theta}_+(\mathbf{k},\eta')\widetilde{F}(k,\eta')d\eta'\,. \tag{6.6}$$

Several proposals for $k_\mathrm{c}(\eta)$ have been discussed in the literature. In Ref. [121], $k_\mathrm{c} = 2\pi\eta^{-1}$ was suggested as physically plausible on causality grounds, but not seriously justified. More recently, however, in Ref. [100] the efficacy of the approximation (6.6) has been demonstrated by studying multi-fluid compensation back-reaction effects in greater detail. It is claimed that the compensation scale arises naturally and uniquely from an algebraic identity in the perturbation analysis, although the discussion in section 4.5.3 may cast some doubt on this claim.

As argued previously, to obtain an accurate $k_c(\eta)$, in principle one will also need to study the detailed behaviour of the source near the horizon scale. However, as an approximation for the present study, we have adopted the analytic fit for $k_c(\eta)$ presented in Ref. [100]:

$$k_{\mathrm{c}}(\eta) = \frac{\sqrt{6(3A^2\eta^2 + 12A\eta + 16)}}{(A\eta + 4)\eta}, \tag{6.7}$$

where A is defined by (3.7). This $k_c(\eta)$ smoothly interpolates from $k_c = \sqrt{6}\eta^{-1}$ in the radiation era to $k_c = \sqrt{18}\eta^{-1}$ in the matter era. Of course, there is some dependence of the perturbation amplitude on this choice, and we shall discuss this later.

With the formalism (6.6), we can now calculate the matter density perturbations on a cubic lattice, using the string source which will be simulated in advance. The resulting power spectrum of the linear matter perturbations, $\mathcal{P}(k)$ or $S(k)$, and the mass fluctuation amplitude at $8h^{-1}$Mpc, σ_8, can then be obtained from equations (1.39), (1.41) and (1.40) respectively.

6.3 Approximation schemes

6.3.1 Hot dark matter

In order to study the formation of structures with cosmic strings in HDM models we use a straightforward modification to the perturbation source similar to that employed in Ref. [122]. This method is a reasonably accurate alternative to much more elaborate calculations using the collisionless Boltzmann equation with defect sources. We simply multiply the Fourier transform of the string source term $\widetilde{\Theta}_+(\mathbf{k},\eta)$ by a damping factor $\widetilde{G}(k,\eta)$ given by

$$\widetilde{G}(k,\eta) = \left[\frac{1}{1 + (0.435kD(\eta))^{2.03}}\right]^{4.43}, \tag{6.8}$$

where $D(\eta)$ is the comoving distance travelled by a neutrino with momentum T_ν/m_ν from time η to ∞. The factor $\widetilde{G}(k,\eta)$ is a fit to numerical calculations of the transfer function of a Fermi-Dirac distribution of non-relativistic neutrinos, and accounts for the damping of small-scale perturbations due to neutrino free-streaming [122]. We calculated $D(\eta)$ numerically and found an excellent fit to our results for $T_{\nu 0} = 1.6914 \times 10^{-13}$ GeV and $m_\nu = 91.5\,\Omega_{\mathrm{h}0}h^2$ eV:

$$D(\eta) = \frac{1}{20}\log\big(\frac{5\eta_{\mathrm{eq}}}{\eta} + 1\big). \tag{6.9}$$

We verified that this fit is accurate within 2% error in the time domain of our interest. Figure 6.1 shows the fit, together with the numerical result.

6.3.2 A semi-analytic model

The other key difficulty facing defect simulations is their limited dynamic range. At any one time, an evolving string network sources significant power over a length-scale range which exceeds an order of magnitude. Hence, even for simulations with a dynamic range of two orders of magnitude in conformal time, the power spectrum will only be reliable over one order, even before taking box-size limitations into account. Fortunately, however, we can invoke a semi-analytic model to compensate for this missing power [102, 121], which proves to be fairly accurate in the scaling regimes away from the radiation-matter transition. The procedure is essentially to square the expression (6.6) and then average over directions to obtain the power spectrum $\mathcal{P}(k)$. This becomes a Green function integral over the UETC $\langle\widetilde{\Theta}_+(\mathbf{k},\eta)\widetilde{\Theta}_+(\mathbf{k}',\eta')\rangle$ [121]:

$$\begin{aligned}(2\pi)^3\mathcal{P}(k)\delta^{(3)}(\mathbf{k}-\mathbf{k}') \;&=\; 16\pi^2G^2\int_{\eta_{\mathrm{i}}}^{\eta_0}\int_{\eta_{\mathrm{i}}}^{\eta_0}\widetilde{\mathcal{G}}^{\mathrm{cs}}(k;\eta_0,\eta)\widetilde{\mathcal{G}}^{\mathrm{cs}}(k';\eta_0,\eta') \\ &\quad \langle\widetilde{\Theta}_+(\mathbf{k},\eta)\widetilde{\Theta}_+(\mathbf{k}',\eta')\rangle\widetilde{F}(k,\eta)\widetilde{F}(k',\eta')d\eta d\eta'\,.\end{aligned} \tag{6.10}$$

This can be simplified by noting that although there are two time integrals, the only significant contributions come from times when η and η' are reasonably close. That is because the strings' configurations are uncorrelated when separated by a sufficiently long time, which is the correlation time η_c as defined in (5.12). Thus (6.10) can be simplified as

$$\mathcal{P}(k) = 16\pi^2G^2\mu^2\int_{\eta_{\mathrm{i}}}^{\eta_0}|\widetilde{\mathcal{G}}^{\mathrm{cs}}(k;\eta_0,\eta)|^2\mathcal{F}(k,\eta)\widetilde{F}^2(k,\eta)d\eta\,, \tag{6.11}$$

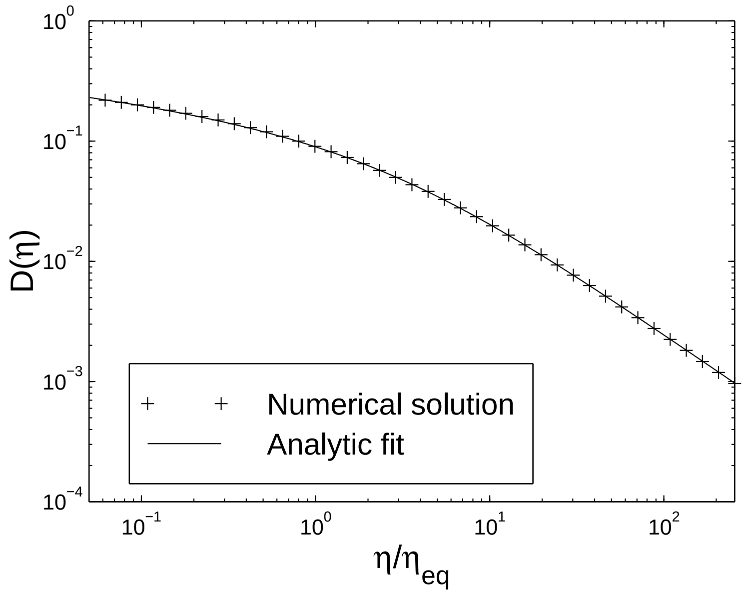

Figure 6.1: Neutrino free-streaming length $D(\eta)$ with $T_{\nu 0} = 1.6914 \times 10^{-13}\,\mathrm{GeV}$ and $m_\nu = 91.5\,\Omega_{\mathrm{h}0} h^2\,\mathrm{eV}$. The analytic fit (6.9) agrees with the numerical solution within a maximum of 2% error.

where

$$(2\pi)^3 \mu^2 \mathcal{F}(k,\eta)\delta^{(3)}(\mathbf{k}-\mathbf{k}') = \frac{\int_{\eta_{\mathrm{i}}}^{\eta_0} \langle \widetilde{\Theta}_+(\mathbf{k},\eta)\widetilde{\Theta}_+(\mathbf{k}',\eta')\rangle \widetilde{\mathcal{G}}^{\mathrm{cs}}(k';\eta_0,\eta')\widetilde{F}(k',\eta')d\eta'}{\widetilde{\mathcal{G}}^{\mathrm{cs}}(k;\eta_0,\eta)\widetilde{F}(k,\eta)} \tag{6.12}$$

$$\approx \int_{-\eta_{\mathrm{c}}}^{\eta_{\mathrm{c}}} \langle \widetilde{\Theta}_+(\mathbf{k},\eta)\widetilde{\Theta}_+(\mathbf{k}',\eta+\eta'')\rangle d\eta''\,. \tag{6.13}$$

The second step, here, is approximate for the following reasons. First, the correlation time η_{c} scales with η and is at most as large as η, while the main contribution to the UETC is from a dynamic range of $(1 \pm 0.2)\eta$ as mentioned in section 5.3.3 (see the contour '0.5' in Figure 5.3). Second, around η_{eq} when strings have the main contribution on the scales of interest (see section 6.3.3 later), $\widetilde{\mathcal{G}}^{\mathrm{cs}}(k';\eta_0,\eta')$ changes only within a factor of 2 when going from $0.05\eta_{\mathrm{eq}}$ to $5\eta_{\mathrm{eq}}$ on subhorizon scales, where strings can seed power into matter perturbations (see equation (4.29), and Figures 4.1 and 4.5). Thirdly, in the matter era, $\widetilde{\mathcal{G}}^{\mathrm{cs}}(k';\eta_0,\eta') \equiv \widetilde{\mathcal{G}}^{\mathrm{cs}}(\eta_0,\eta')$ decays linearly with η', so the scaling of $\mathcal{F}(k,\eta)$ in η is maintained, although a mild rescaling in the amplitude will be required to maintain accuracy. Finally, although another mild rescaling in k is also required for the approximation (6.13), the scaling behaviour of $\mathcal{F}(k,\eta)$ in $k\eta$ should hold.

The structure function $\mathcal{F}(k,\eta)$ can then be obtained as discussed below:

1. The form of $\mathcal{F}(k,\eta)$: We know cosmic strings are line-like objects which move relativistically, so the trajectories they sweep out will be two-dimensional. According to the study in section 5.2, we know sheet-like objects have power spectra $\mathcal{P}(k) \propto k^{-2}$ at small scales and $\mathcal{P}(k) \propto k^0$ at large scales. For cosmic string wakes, we obtain $\mathcal{P}(k) \propto k^{-2.25}$ on small scales because of their wiggliness (see section 5.3.4). In (6.12), because both $\widetilde{F}(k',\eta')$ and $\widetilde{\mathcal{G}}^{\mathrm{cs}}(k';\eta_0,\eta')$ are scale-independent on subhorizon scales (see equation (4.63) and Figure 4.7), the slope of $\mathcal{F}(k,\eta)$ on these scales should be identical to that of (6.13). We have observed this slope to be -2.25 in Figure 5.4. Although in the radiation-dominated regime, strings are more wiggly and therefore have a steeper small-scale slope than -2.25, we found it adequate to use this value throughout for the current purposes. In addition, because cosmic strings scale with the horizon size, the turnover scale from $\mathcal{P}(k) \propto k^{-2.25}$ to $\mathcal{P}(k) \propto k^0$ should also scale with conformal time. Adding all these factors together, we can schematically write down the structure function in the form:

$$\mathcal{F}(k,\eta) = \frac{\mathcal{E}(\eta)}{[1+(k/k_{\mathrm{to}})^n]^{2.25/n}} = \frac{\mathcal{E}(\eta)}{[1+(Ik\eta)^n]^{2.25/n}}\,, \tag{6.14}$$

where $\mathcal{E}(\eta)$ is the overall normalization against time, I is a constant, $k_{\mathrm{to}} = (I\eta)^{-1}$ is the turnover wave number, and the power n controls the sharpness of the turnover behaviour. We notice that the broad peak seen in Figure 5.4 is not modelled in (6.14). Nevertheless, we shall see in section 6.4.1 that such a simplification has a negligible effect on the final power spectrum once the phenomenological structure function $\mathcal{F}(k,\eta)$ is accurately calibrated by high-resolution simulations over a large dynamic range.

2. The time dependence: A convenient way to investigate the time dependence $\mathcal{E}(\eta)$ is to look at the horizon scale $k \approx \eta^{-1}$, at which $\mathcal{F}(k,\eta)|_{k\eta=1} \approx \mathcal{E}(\eta)$. From the scale-invariance property of defects as a consequence of their scaling behaviour, we know $S(k) = 4\pi k^3 \mathcal{P}(k)|_{\rm hor} = $ constant. Adding the fact that $\mathcal{P}(k) \propto (\frac{\delta\rho_{\rm s}}{\rho_{\rm bg}})^2$ where $\rho_{\rm bg} = \rho_{\rm m} + \rho_{\rm r}$ and $\delta\rho_{\rm s}$ is the string energy density behaving like a first order perturbation in the background, we get $\delta\rho_{\rm s}^2 \propto \rho_{\rm bg}^2 \eta^3$ at the horizon scale $k^{-1} \sim \eta$. Since the background energy density scales as $\rho_{\rm bg} \propto t^{-2} \propto (a\eta)^{-2}$, we obtain

$$\widetilde{\Theta}_+\Big|_{\rm hor} \equiv a^2 \delta\rho_{\rm s} \propto a^2 (\eta^3 \rho_{\rm bg}^2)^{1/2} \propto \eta^{-1/2} \,. \tag{6.15}$$

Thus by (6.13) the structure function is

$$\begin{aligned} \mathcal{F}(k,\eta)|_{\rm hor} &\approx \frac{1}{(2\pi)^3\mu^2} \int_{-\eta_{\rm c}}^{\eta_{\rm c}} \langle \widetilde{\Theta}_+(\mathbf{k},\eta) \widetilde{\Theta}_+(\mathbf{k},\eta+\eta'')\rangle \Big|_{\rm hor} d\eta'' \\ &\propto \int_{-\eta_{\rm c}}^{\eta_{\rm c}} \eta^{-1/2} (\eta+\eta'')^{-1/2} d\eta'' = \frac{2\eta_{\rm c}}{\eta} + {\rm O}^2(\tfrac{\eta_{\rm c}}{\eta}) \,. \end{aligned} \tag{6.16}$$

Since cosmic strings scale with the horizon size in the scaling regime, the ratio $\eta_{\rm c}/\eta$ is a constant. It then follows from equation (6.16) that the normalization of the structure function against time at large scales is a constant when deep in the radiation or matter era, i.e.

$$\mathcal{E}(\eta) \approx \mathcal{F}(k,\eta)|_{\rm hor} \approx \text{constant} \,. \tag{6.17}$$

This result is consistent with that presented in Ref. [121], but here we give a clear physical reason for the behaviour of the structure function and obtain an explicit expression for the relationship between $\mathcal{F}(k,\eta)$ and the correlation time. The longer the correlation time, the larger the normalization of $\mathcal{F}(k,\eta)$. Therefore an insufficient dynamic range when calculating the UETC's or $\mathcal{F}(k,\eta)$ will result in a severe under-estimate of the final power spectrum.

As illustrated in the discussion following equation (6.13), although the approximation may not be very accurate, the scaling behaviour of $\mathcal{F}(k,\eta)$ in $k\eta$ should still be maintained. Hence instead of taking the approximation (6.13), we can use simulations to calibrate $\mathcal{F}(k,\eta)$. By fitting the shape and amplitude of the power spectrum calculated from (6.11) with simulations of limited dynamic range deep in the matter and radiation eras, we can estimate $\mathcal{F}(k,\eta)$ in these two regimes. We were then able to use an interpolation based on the actual behaviour of the string density during the transition era (see Figure 5.5) to provide the normalization of $\mathcal{F}(k,\eta)$ which interpolates smoothly from the radiation to the matter era. This turns out to fit well the shape of the actual evolution of $\mathcal{E}(\eta)$ in (6.14). We found $I = 1/20$ and $n = 8$ in (6.14), and the resulting structure function is:

$$\mathcal{F}(k,\eta) = \frac{\mathcal{E}(\eta)}{[1+(k\eta/20)^8]^{0.28125}} \quad \text{with} \quad \mathcal{E}(\eta) = 0.028 \left[1 - \frac{0.37322}{1+(4\eta_{\rm eq}/\eta)^{1.1}}\right]^2 . \tag{6.18}$$

In (6.18), we note first that the turnover scale is the same for both the radiation and the matter eras, i.e. $k_{\rm to} = 20/\eta$. This corresponds to exactly the correlation length of long strings which we have seen in section 5.3.2. Second, $\mathcal{E}(\eta)$ in (6.18) interpolates smoothly from 0.028 deep in the radiation era to 0.011 deep in the matter era. Now with $\widetilde{\mathcal{G}}^{\rm cs}(k;\eta_0,\eta)$ numerically calculated from equations (6.1) and (6.2) or analytically obtained from equations (4.45) and (4.52), we can use the approach (6.11) and (6.18) to extrapolate the lost power due to the limited dynamical range in our numerical simulation. We notice that we have not considered the effect of loops here, which will be discussed in section 6.4.6.

6.3.3 The dependence of $S(k)$ on $\Omega_{\rm m0}$, Λ, and h

Since the existence of a cosmological constant seems to be favoured by recent observations, and an open universe is studied in many inflationary models (e.g. [123, 124]), it is natural to explore the string-induced power spectrum in these two regimes. To do this, we first introduce a simple rescaling scheme to extrapolate the simulated $\Omega_{\rm m0} = 1$ and $\Lambda = 0$ spectrum to open and Λ-models, and then verify their accuracy. As an analogue to equation (4.52), the rescaling scheme is (adapted from [118]):

$$S(k,h,\Omega_{\rm m0},\Omega_{\Lambda0}) = S(k,1,1,0) \times \Omega_{\rm m0}^2 h^4 \times f^2(\Omega_{\rm m0},\Omega_{\Lambda0}) \times g^2(\Omega_{\rm m0},\Omega_{\Lambda0}), \tag{6.19}$$

where k is in units of $\Omega_{\rm m0} h^2\,{\rm Mpc}^{-1}$, $g(\Omega_{\rm m0},\Omega_{\Lambda0})$ is given by equation (4.53), and

$$f(\Omega_{\rm m0},\Omega_{\Lambda0}) = \begin{cases} \Omega_{\rm m0}^{-0.3} & \text{for } \Omega_{\rm m0} \le 1,\ \Omega_{\Lambda0} = 0 \\ \Omega_{\rm m0}^{-0.05} & \text{for } \Omega_{\rm m0} + \Omega_{\Lambda0} = 1 \end{cases} . \tag{6.20}$$

In (6.19), the leading factor $\Omega_{\rm m0}^2 h^4$ results from the fact that the ratio of scale factors $a_0/a_{\rm eq}$ is proportional to $\Omega_{\rm m0}h^2$ and that the Green function is proportional to this ratio (see equation (4.28) and context). The middle factor $f(\Omega_{\rm m0},\Omega_{\Lambda 0})$ reflects the dependence of the COBE normalisation of $G\mu$ on the cosmological parameters $\Omega_{\rm m0}$ and $\Omega_{\Lambda 0}$. It changes upwards as we decrease the matter density in an open or Λ-universe. The last factor $g(\Omega_{\rm m0},\Omega_{\Lambda 0})$ takes into account the fact that in an open or Λ-universe the linear growth of density perturbations is suppressed relative to an $\Omega_{\rm m0}=1$ and $\Omega_{\Lambda 0}=0$ universe [98] (also verified in Ref. [99] for primordial perturbations). There is also a rescaling of k implicit in (6.19) (the unit of k here is in $\Omega_{\rm m0}h^2\,{\rm Mpc}^{-1}$ rather than $h\,{\rm Mpc}^{-1}$). This rescaling is due to the fact that the horizon size at radiation-matter equality $\eta_{\rm eq}$ is proportional to $(\Omega_{\rm m0}h^2)^{-1}$ (see section 3.2). Since the lengthscale in the theoretical framework is in units of $\eta_{\rm eq}$, the physical grid spacing in our simulation should be rescaled by this factor. We note that a similar scheme was also taken in Ref. [123] to study inflationary cosmology in an open universe.

To verify the accuracy of this scheme, we use equations (6.11), (6.18), and the Green functions obtained from solving (6.1) and (6.2) numerically, to get the spectrum $S_{\rm num}(k)$ for various choices of $\Omega_{\rm m0}$ and $\Omega_{\Lambda 0}$. We then compare this with $S_{\rm ap}(k)$ which is extrapolated from a $K=\Lambda=0$ model using (6.19). The initial time of the integral is set to $\eta_{\rm i}=10^{-3}\eta_{\rm eq}$, at which both the curvature and Λ terms are negligible, while the final time is today. Figure 6.2 shows that on the scales of interest, the rescaling scheme (6.19) is accurate within a few percent error for reasonable choices of $\Omega_{\rm m0}$ and $\Omega_{\Lambda 0}$. Similar independent work concerning the accuracy of (6.19) in generic defect models is also done in Ref. [125].

Figure 6.3 shows the relative contribution from different epochs of the total power spectrum. We can see that most power on the scales of interest is seeded around $\eta_{\rm eq}$. This is because at early times in the radiation-dominated regime, the growth of small-scale perturbations is suppressed by the radiation pressure, while at late times in the matter-dominated regime, the large-scale perturbations have less time to grow (see the $\widetilde{T}_\infty^{\rm cs}$ in Figure 4.1). In $K\neq 0$ or $\Lambda\neq 0$ models, the normalization of $S(k)$, which is a dimensionless quantity, scales in the way described in (6.19), while $\mathcal{P}(k)=S(k)/4\pi k^3$ is further changed by the rescaling of k if we are to use $h{\rm Mpc}^{-1}$ as the unit for k when comparing theoretical results to observations.

As to the non-scaling behaviour of cosmic strings at late times in $\Lambda\neq 0$ or $K\neq 0$ models, Figure 6.3 combined with Figure 3.1 implies the unimportance of this effect in the resulting CDM power spectrum. We know that in open or Λ models, the motion of cosmic strings will be damped by the late-time rapid expansion of the universe, so that the string network has a decreasing RMS velocity and eventually has no effect in seeding matter perturbations. Taking an extreme case $\Omega_{\rm m0}=0.15$ with $h=0.7$, the largest observable scale $k\sim 0.01h{\rm Mpc}^{-1}$ corresponds to $k\sim 0.1\Omega_{\rm m0}h^2{\rm Mpc}^{-1}$ in Figure 6.3. On this scale, the late time strings (the dotted line, whose initial time is $20\eta_{\rm eq}$) contribute only 3% to the total power spectrum (the solid line). In addition, according to Figure 3.1, by the time $\eta=20\eta_{\rm eq}$, the scale factor departs from the $K=\Lambda=0$ model by only 9% in the open model, and 0.3% in the flat Λ model. Such a small departure in the background dynamics will have little effect on the evolution of cosmic strings. Therefore, we expect the late time non-scaling behaviour of cosmic strings to be negligible in affecting the resulting matter density perturbations at the scales of interest.

6.4 Results and discussion

To obtain our numerical results, we first performed string simulations with a string sampling comoving spacing of 10^{-3} for the simulation box sizes. The dynamic ranges in these simulations covered from $0.05\eta_{\rm eq}$ to $300\eta_{\rm eq}$, with each single run having a dynamic range of at least 10 in conformal time. We then used the formalism (6.6) to perform the structure formation simulations with comoving box sizes ranging from 5–$130h^{-1}$Mpc today, and a resolution of 128^3–512^3. Our simulations were carried out on the UK Computational Cosmology Consortium supercomputer COSMOS, a Silicon Graphics Origin2000 with 20 Gbytes main memory. The code was parallelised to enhance performance. We also used the SGI math libraries to implement the parallelised Fast Fourier Transform.

6.4.1 The semi-analytic fit

In Figure 6.4 we plot the CMBR normalized linear power spectrum ($G\mu_6=G\mu\times 10^6=1.7$, the most recent COBE normalization for strings [119]) induced by long strings in an $\Omega_{\rm c0}=1$ and $\Omega_{\rm B0}=0$ CDM cosmology with $h=0.7$. The central set of numerical points was sourced by string network simulations beginning at $\eta=0.4\eta_{\rm eq}$ which were continued for 1318 expansion times. Given the dynamic range limitations we have also plotted the semi-analytic fit (6.11) combined with (6.18) over the full range of wave-numbers. The good agreement with the semi-analytic model is illustrated by the dashed line fits to the central points, as well as to the short normalization runs in the matter and radiation eras (also shown in Figure 6.4). We can see the small discrepancy between the semi-analytic fits and simulations at both the large-scale and the small-scale ends. This results from the fact that towards the beginning and the end of the simulations, the UETC does not fully contribute to the

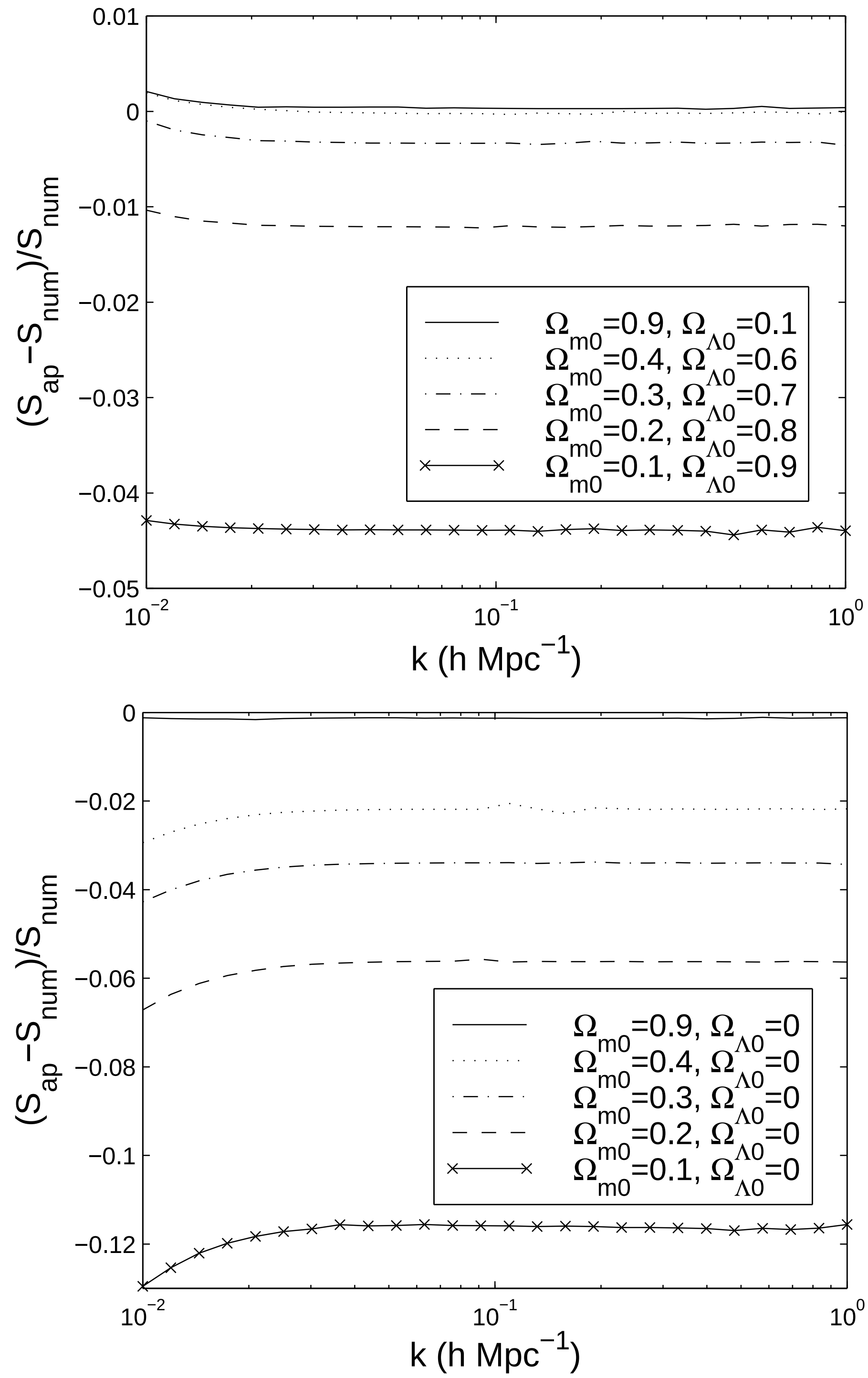

Figure 6.2: The accuracy of the rescaling scheme (6.19) for cosmic string models. $S_{\rm ap} \equiv 4\pi k^3 \mathcal{P}(k)$ is obtained from a flat $\Lambda = 0$ model using (6.19), while $S_{\rm num}$ is the numerical result obtained by solving perturbation equations with various choices of $\Omega_{\rm m0}$ and $\Omega_{\Lambda 0}$.

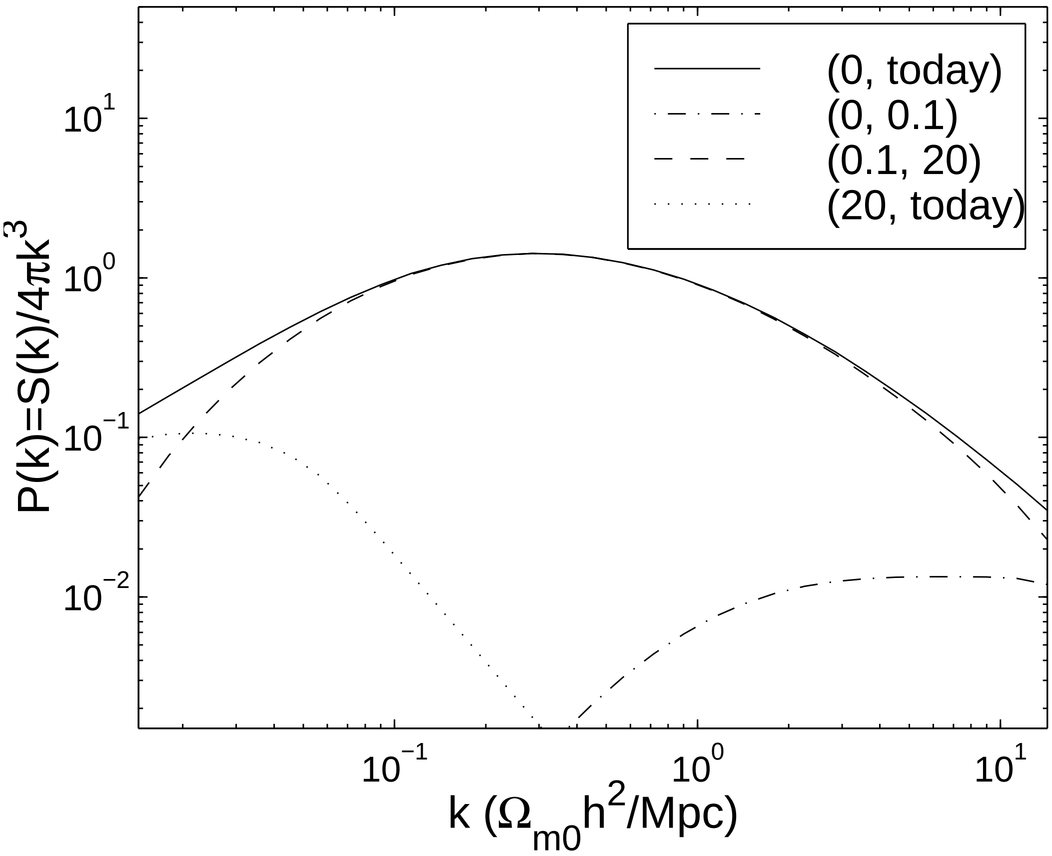

Figure 6.3: The relative contribution in the power spectrum from strings at different epochs: the total (solid line), the early times (dot-dashed line), around $\eta_{\rm eq}$ (dashed line), and the late times (dotted line). The numbers in brackets specify in $\eta_{\rm eq}$ the initial and final time of each epoch.

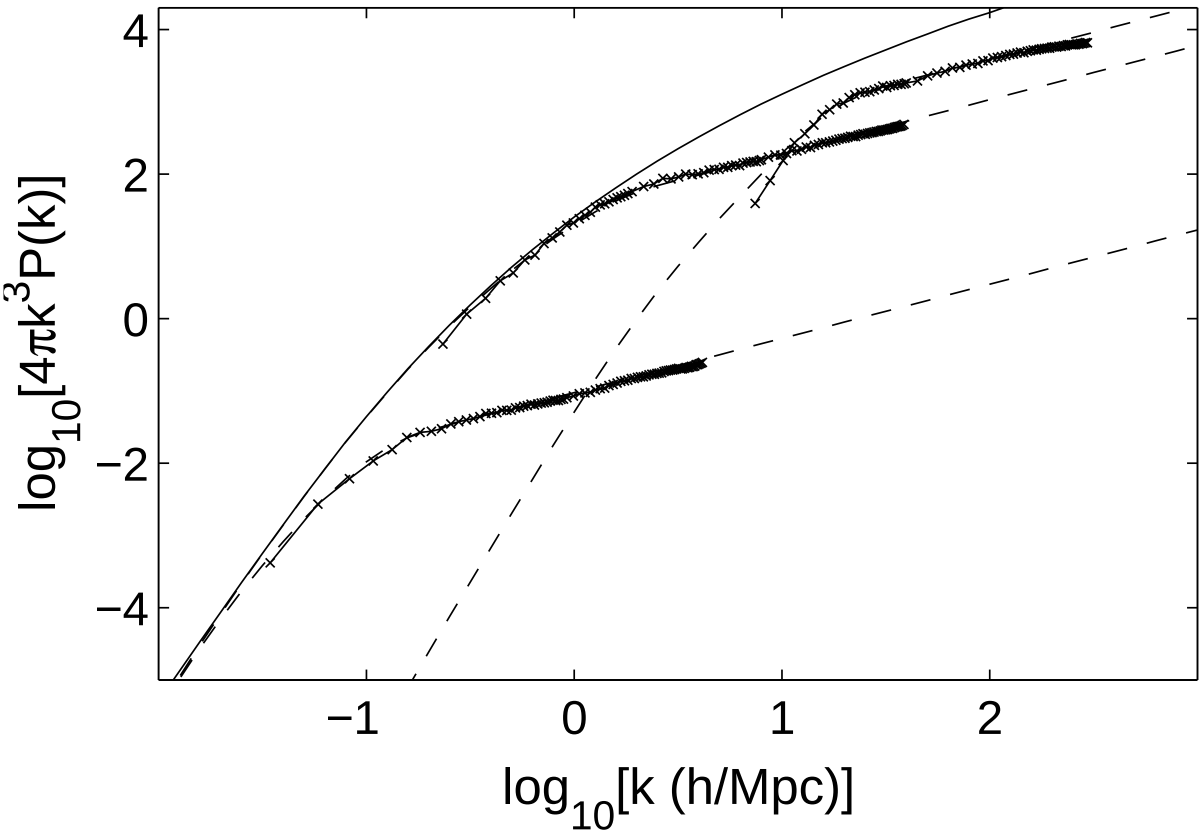

Figure 6.4: The comparison of our semi-analytic fit and our simulation result for the CDM model excluding loops. The top-right, central, and bottom-left solid lines with crosses are the simulation results in the deep radiation, transition through to deep into the matter, and deep matter eras respectively. The dashed lines are our semi-analytic fits corresponding to the same dynamical ranges of those simulations. They show a good agreement with each other. The solid line is the semi-analytic fit with a full dynamic range from $\eta_\mathrm{i} = 0$ to today.

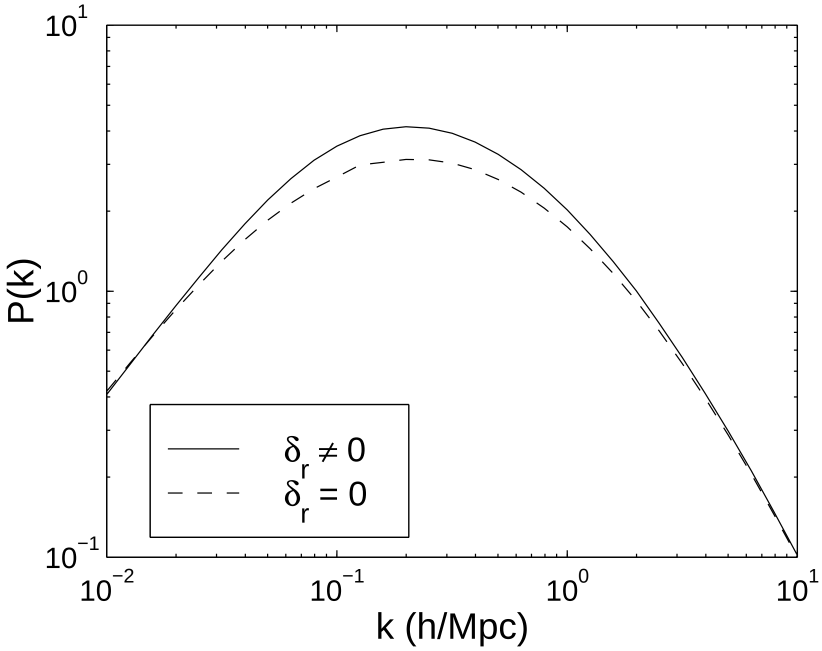

Figure 6.5: The effect of radiation perturbations in the cosmic string-seeded CDM power spectrum. The dynamic range goes from $\eta = 0$ to today.

simulation power spectrum (see equation (6.13)). Nevertheless, at large scales ($k \sim 0.03$) the full-dynamic-range run of the semi-analytic fit (the solid line) is well constrained by the deep-matter-era simulation. Given this close correspondence, which was also exhibited in HDM simulations, we have considerable confidence that this approximation can reproduce the correct shape and amplitude of the string simulation power spectrum.

6.4.2 The effect of radiation

Using this semi-analytic fit, Figure 6.5 shows how the radiation perturbations affect the matter perturbation power spectrum. The solid line is obtained from (6.1) and (6.2) with $\Lambda = K = 0$, while the dashed line is obtained by setting $\delta_{\rm r} = 0$ in (6.1) and ignoring (6.2). We can see that the radiation perturbations boost the power spectrum by about 30% on scales $k \sim 0.2h\mathrm{Mpc}^{-1}$, but have little effect towards large or small scales. This is because at very small scales, radiation will oscillate many times per expansion time and will have little net effect on the matter, while at very large scales, most power is contributed from the matter era in which the radiation is negligible. Although ignoring $\delta_{\rm r}$ is convenient for numerical purposes (e.g. [126], where only the matter-dominated era and no radiation component were considered), the assumption of a smooth radiation background would have ignored significant intermediate-scale power coming from the epoch around the radiation-matter transition. We also note that the result with $\delta_{\rm r} = 0$ is equivalent to setting the condition $k\eta \gg 1$ when solving (4.10) and (4.11), so that the Green functions $\widetilde{\mathcal{G}}^{\rm c}_i$ in (4.45) decay to only $\widetilde{\mathcal{G}}^{\rm c}_{i\infty}$.

6.4.3 The compensation scale

As mentioned previously, there is a major uncertainty in our framework for calculating the matter perturbation power spectrum—the compensation mechanism. Figure 6.6 shows the effect of different compensation scales on the spectrum. Here we have used the semi-analytic model which is described in section 6.3.2 and verified in 6.4.1. Evidently, a larger compensation scale (at the same conformal time) will give significantly more power at larger scales. For example, with the larger $k_{\rm c}$ of Ref. [121], the amplitude is approximately 40% smaller around length-scales of $60h^{-1}$Mpc and larger.

Clearly, the quantitative implementation of compensation effects is one of the key uncertainties in all previously published work on gauged cosmic strings. In this case, one needs not only to conserve the total energy of the universe as considered in section 4.5.3, but also to properly evolve all the byproducts of the string network in order to account for extra possible mechanisms for seeding density perturbations [3, 113, 127]. Therefore apart from the treatment presented in equation (4.61) or (4.62), the compensation problem ultimately requires a full-scale numerical treatment in which all strings and background fluid components are accurately evolved through to the present day [10].

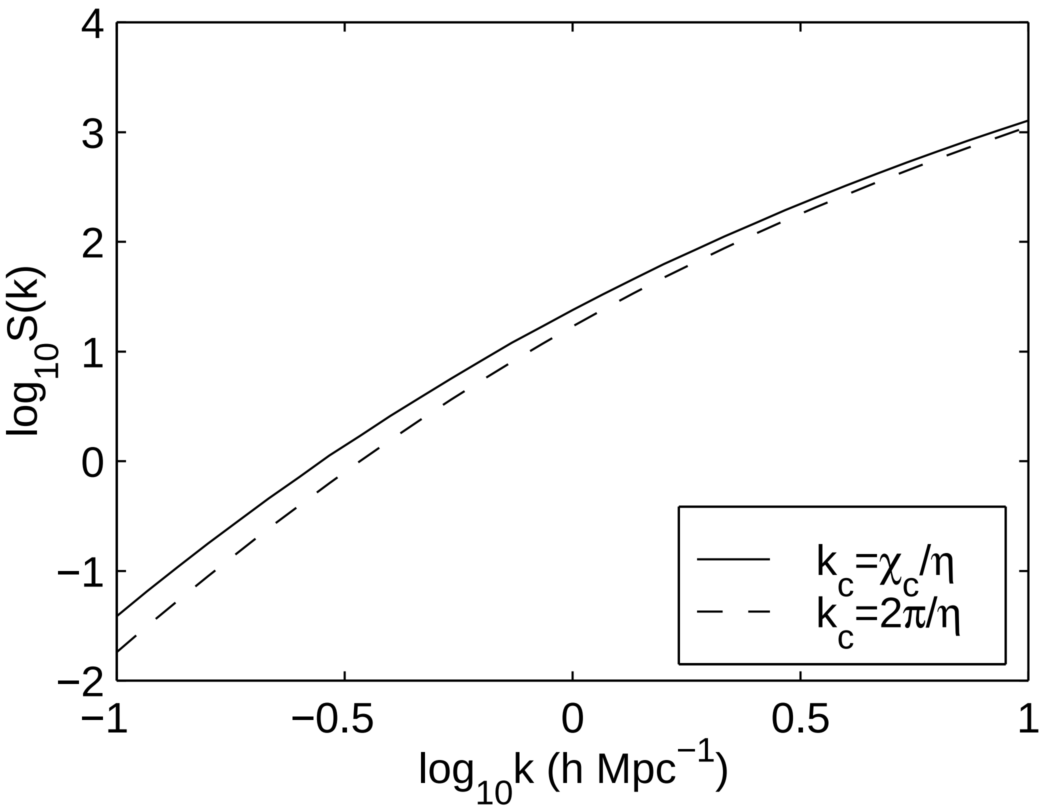

Figure 6.6: A comparison of spectra using different compensation scales. The solid line used the definition in [100], where χ_c varies from $\sqrt{6}$ in the radiation epoch to $\sqrt{18}$ in the matter epoch; the dashed line used the compensation scale $k_c = 2\pi/\eta$ suggested in [121]. The dynamical range is $\eta = (0.01, 50)\eta_{eq}$.

6.4.4 The $\Omega_{\rm m0} = 1$ and $\Lambda = 0$ model

In Figure 6.7, we compare our CDM and HDM string power spectra (excluding loops and baryons) with the observational results inferred from galaxy surveys, in each of five different background cosmologies: (I) $\Omega_{\rm m0} = 1$, $\Omega_{\Lambda 0} = 0$, (II) $\Omega_{\rm m0} = 0.3$, $\Omega_{\Lambda 0} = 0.7$, (III) $\Omega_{\rm m0} = 0.3$, $\Omega_{\Lambda 0} = 0$, (IV) $\Omega_{\rm m0} = 0.15$, $\Omega_{\Lambda 0} = 0.85$, and (V) $\Omega_{\rm m0} = 0.15$, $\Omega_{\Lambda 0} = 0$. The results presented here are obtained by integrating (6.11), with $\mathcal{F}(k,\eta)$ given in equation (6.18) and $\widetilde{\mathcal{G}}^{\rm cs}(k;\eta_0,\eta)$ numerically obtained from equations (6.1) and (6.2).

Consider first the $\Omega_{\rm c0} = 1$ CDM model (I). An empirical formulae which can reproduce our result (excluding loops and baryons) within a maximum of 10% error for $k = 0.01$–$100h{\rm Mpc}^{-1}$ is obtained as:

$$S_{\rm CDM}(k) = (G\mu_6)^2 10 (0.7q)^{p(q)}\,, \tag{6.21}$$

where $q = k/\Gamma$ and

$$p(q) = 3.9 - \frac{2.7}{1 + (2.8q)^{-0.44}}\,. \tag{6.22}$$

The Γ here is called the 'shape parameter', which accounts for the rescaling in k due to different choices of cosmological parameters. In the $\Omega_{\rm B0} = 0$ cosmologies, Γ is simply $\Omega_{\rm m0}h$, which reflects the scaling in the horizon size at $\eta_{\rm eq}$. As for the $\Omega_{\rm B0} \neq 0$ models, the shape parameter Γ will be also a function of $\Omega_{\rm B0}$ as we shall discuss later.

We calculated the standard deviation σ_8 by convolving our CDM perturbations with a spherical window of radius $8\,h^{-1}\,$Mpc to find $\sigma_{8({\rm sim})}(h = 0.5) = 0.32G\mu_6$, $\sigma_{8({\rm sim})}(h = 0.7) = 0.39G\mu_6$ and $\sigma_{8({\rm sim})}(h = 1.0) = 0.47G\mu_6$. A comparison with the observational data points is shown in the top panel of Figure 6.7. It indicates that strings appear to induce an excess of small-scale power and a shortage of large-scale power, that is, the $\Omega_{\rm c0} = 1$ string model excluding loops and baryons requires a strongly scale-dependent bias. This is not necessarily a fatal flaw for the model on small scales because, as the corresponding HDM spectrum indicates, such excess power can be readily eliminated in a mixed dark matter model. However, the problem is less tractable on large scales where biases up to $\sigma_{100({\rm obs})}/\sigma_{100({\rm sim})} \approx 3.9$ around $100h^{-1}$Mpc might be inferred from the data points (using $G\mu_6 = 1.7$ and $h = 0.7$). Should we, therefore, rule out string models on this basis [106]?[1] While acknowledging that the model (I) linear spectrum looks unattractive, there are still three important mitigating factors. First, the determination of the power spectrum on large scales around $100h^{-1}$Mpc remains uncertain (e.g. the observational data in Figure 6.7 has at least a 40% normalization uncertainty), and these will be superseded by much larger more reliable data sets in the near future. Secondly, the immediate nonlinearity of string wakes indicates that strong biasing mechanisms might operate, as illustrated on large scales in the post-recombination hydrodynamic simulations of Ref. [128]. Thirdly, cosmic string loops have not been included here, and we shall see how they may circumvent this problem later. Finally, there are other cosmologies which have not been considered.

As for the HDM results, the small-scale power ($k \gtrsim 0.5h{\rm Mpc}^{-1}$) is removed by the neutrino free-streaming, while on large scales it remains the same as that of the CDM model. An empirical formulae (excluding the effect of loops and baryons) within a maximum of 10% error for $k = 0.01$–$100h{\rm Mpc}^{-1}$ is given by:

$$S_{\rm HDM}(k) = (G\mu_6)^2 \frac{170}{\left[(0.7q)^{-1.2} + 5.3q\right]^{3.3}}\,, \tag{6.23}$$

where again $q = k/\Gamma$ and Γ is the shape parameter. Figure 6.8 shows a comparison between the mass distributions of CDM and HDM models. They are slices of a simulation box of size $(128h^{-1}{\rm Mpc})^3$, with exactly the same string source. They look the same on larger scales, but the HDM model has no fine structures on smaller scales due to the neutrino free-streaming. We can also see non-Gaussian features in the distribution, which we shall investigate further in Chapter 7.

6.4.5 The cases $\Omega_{\rm m0} < 1$ and $\Lambda \neq 0$

We can observe from Figure 6.7, that for open or Λ-cosmologies with $\Omega_{\rm m0} \approx 0.1$–$0.3$, the string+CDM power spectrum is much more encouraging. We find that the bias on large scales is much less scale-dependent. For example, in model (IV) ($\Omega_{\rm m0} = 0.15$, $\Omega_{\Lambda 0} = 0.85$), the relative bias remains $\sigma_{100({\rm obs})}/\sigma_{100({\rm sim})} \approx 1.4 \pm 0.2$ across all the observational scales.

[1] In [106], the authors used a numerical fit for the observational data to obtain $\sigma_{100({\rm obs})} = 3.5 \times 10^{-2}$, and compared this with their simulated spectrum to yield $b_{100} = 5.4$. They concluded that cosmic string theory is ruled out because of this high bias factor. However, according to the observational data [33] (see Figure 6.7), there are only few data points (with big error bars) below the scale of $100h^{-1}$Mpc, so this seems a potentially unreliable way to estimate $\sigma_{100({\rm obs})}$. Nevertheless, the discrepancy between the observation and the string-predicted result is still apparent.

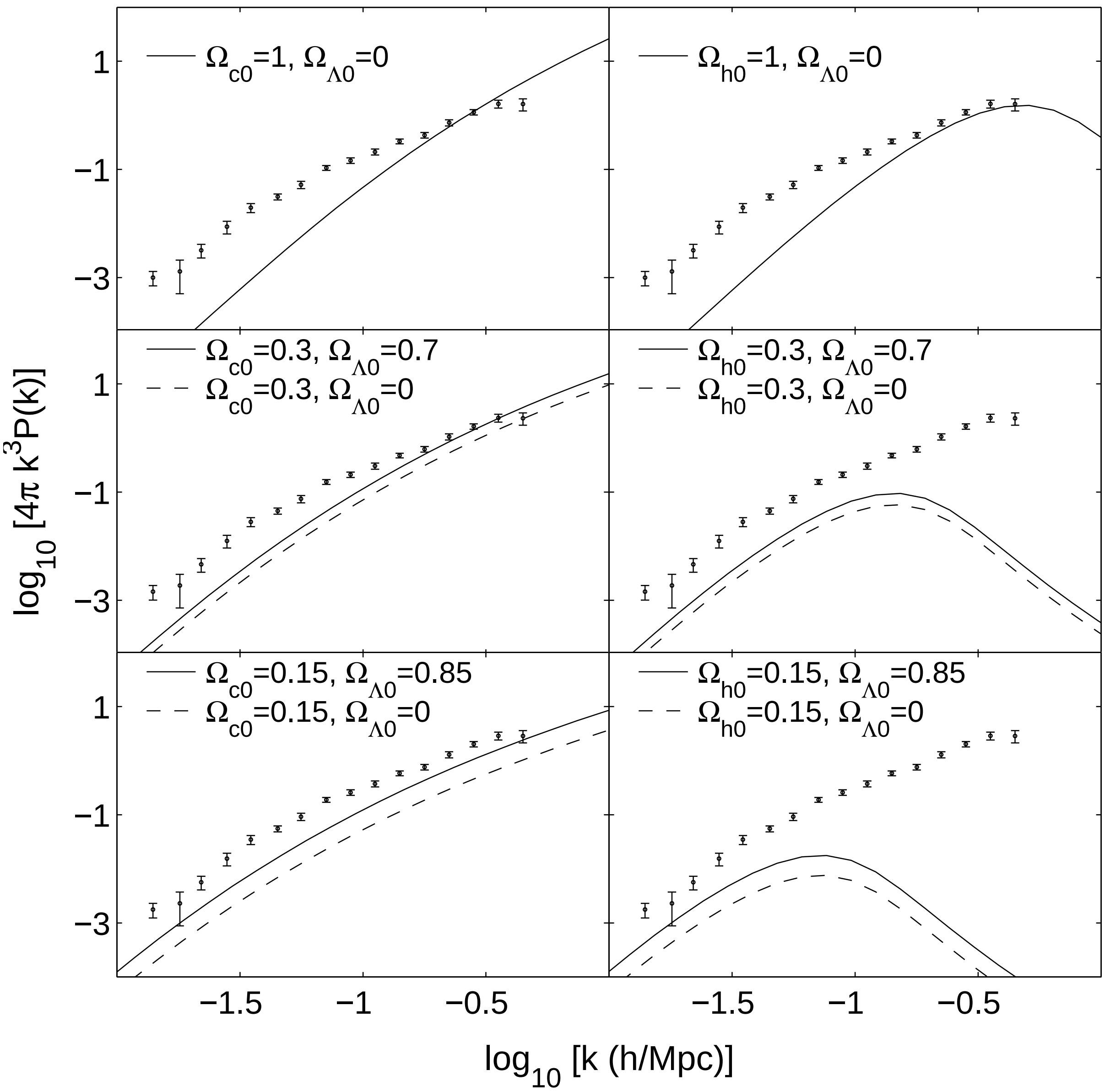

Figure 6.7: The spectra of cosmic strings (excluding loops) with CDM (left) and HDM (right) models in different background cosmologies. The solid lines are the flat models with $\Omega_{\rm m0} + \Omega_{\Lambda 0} = 1$; the dashed lines are the open models with $\Omega_{\Lambda 0} = 0$. In an open universe with $\Lambda = 0$, the spectrum has more large-scale power and less small-scale power as $\Omega_{\rm m0}$ decreases. In a flat universe made by Λ, the spectrum has a scale-independent boost over the previous case with the same $\Omega_{\rm m0}$. Here we use $h = 0.7$ and $G\mu_6 = 1.7$ [119], which is the most recent COBE normalisation for cosmic strings. The data points with error bars are the reconstructed linear spectrum by Peacock and Dodds [33].

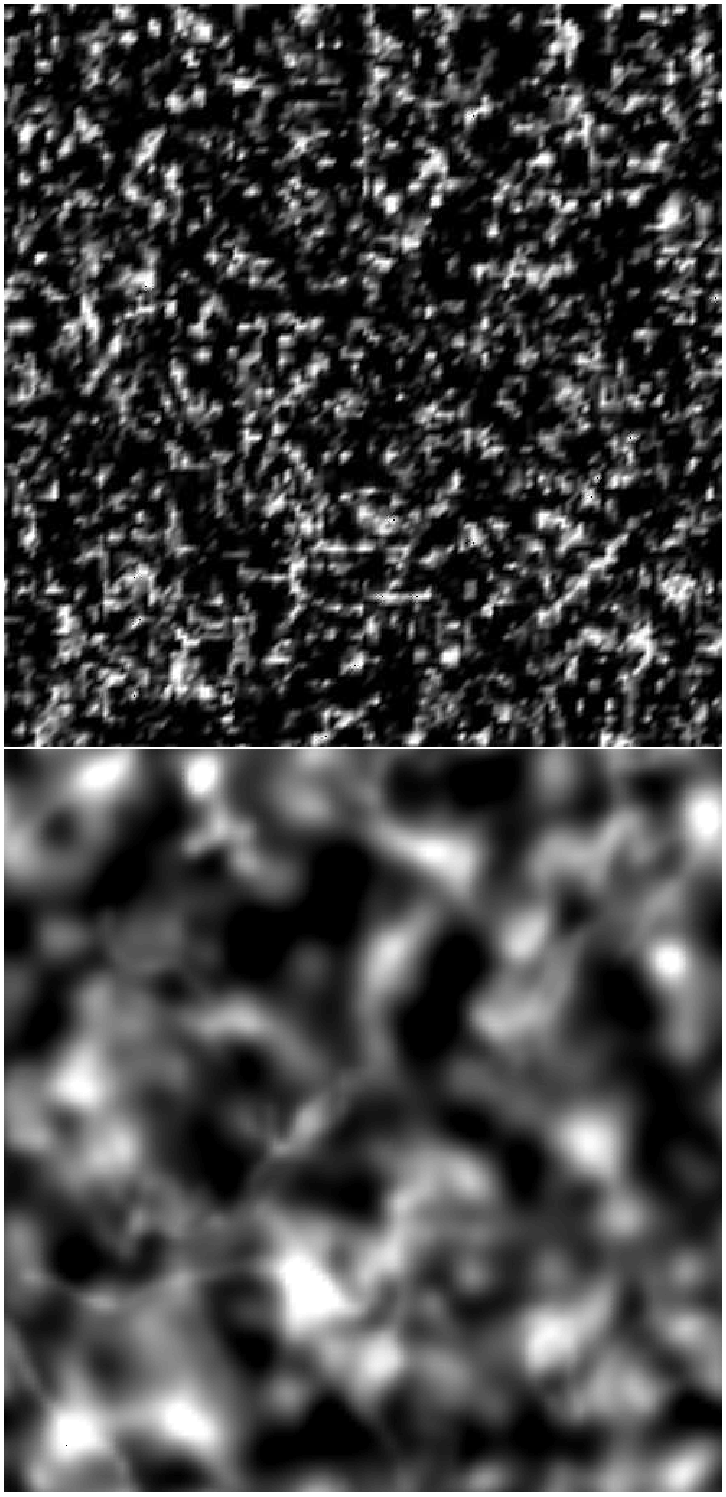

Figure 6.8: A comparison between CDM (top) and HDM (bottom) results. They are taken from slices in a simulation box of size $(128h^{-1}\mathrm{Mpc})^3$, with exactly the same string source. They look the same on large scales, but the HDM model has no fine structure. The colour scheme is chosen so that the non-Gaussian feature can be clearly distinguished.

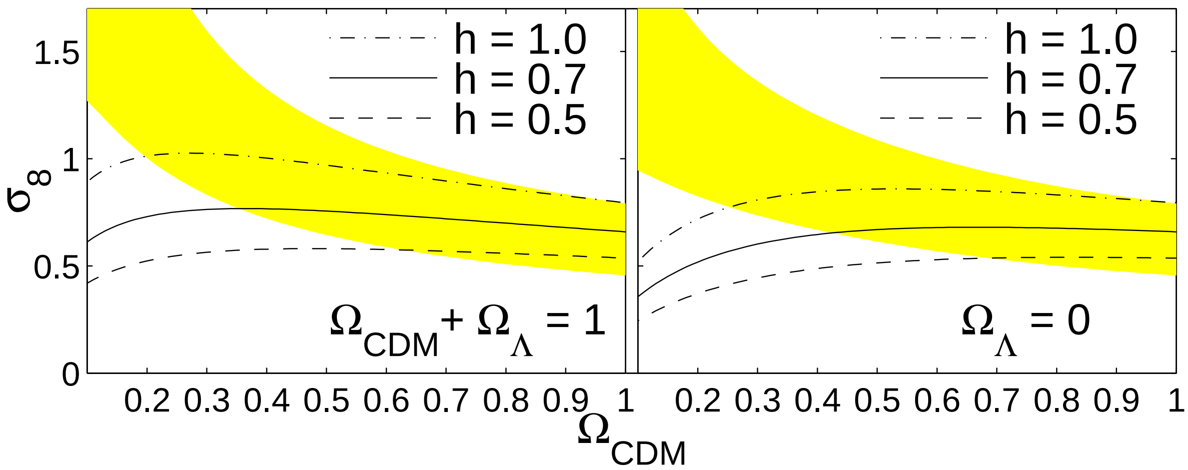

Figure 6.9: The comparison of the observationally inferred standard deviation at the scale 8 h^{-1}Mpc, $\sigma_{8(\mathrm{obs})}$, and that induced from our simulation, $\sigma_{8(\mathrm{sim})}$. We have used the COBE normalization $G\mu_6(\Omega_{\mathrm{c}0} = 1, \Omega_{\Lambda 0} = 0) = 1.7$ [118, 119]. $\sigma_{8(\mathrm{obs})}$ is shown as the shaded area [129]; $\sigma_{8(\mathrm{sim})}$ is plotted as dot-dashed ($h = 1.0$), solid ($h = 0.7$) and dashed($h = 0.5$) lines.

In Figure 6.9, we plot σ_8 for the full gamut of open and Λ models. The $\sigma_{8(\mathrm{sim})}$ deduced from our simulations with $G\mu_6$ normalized by COBE [118, 119] is represented by the dot-dashed ($h = 1.0$), the solid ($h = 0.7$) and the dashed ($h = 0.5$) lines; $\sigma_{8(\mathrm{obs})}$ is the shaded area, computed by Viana & Liddle [129], and which is also consistent with Ref. [28, 130], but more conservative. We can see from Figure 6.9 that $\sigma_{8(\mathrm{obs})}/\sigma_{8(\mathrm{sim})}(\Omega_{\mathrm{m}0} = 1) \approx 0.79 \pm 0.21, 0.95 \pm 0.25, 1.17 \pm 0.31$ for $h = 1.0, 0.7, 0.5$ respectively. When $h = 0.7$, $\sigma_{8(\mathrm{sim})}$ matches $\sigma_{8(\mathrm{obs})}$ within the uncertainties for flat Λ-models when $\Omega_{\mathrm{m}0} \gtrsim 0.35$ and for open models when $\Omega_{\mathrm{m}0} \gtrsim 0.4$, while for both cases the ratio $\sigma_{8(\mathrm{obs})}/\sigma_{8(\mathrm{sim})} \lesssim 2$ for all $\Omega_{\mathrm{m}0} \gtrsim 0.1$. However, the observational result used here is based on the standard CDM model [129, 131], so a direct comparison of it with our results [1] does not provide reliable information regarding the bias of the string model. We shall further investigate this aspect in Chapter 8.

Hence in terms of the spectrum shape, an open or Λ-cosmology in the context of string+CDM model seems to show remarkable agreement between observations and the results of our current numerical simulations. Indeed, similar comparisons made in an inflationary context with models like I–V appear to require a more strongly scale-dependent bias [124]. To find the best fit of the string-induced spectrum shape to the observation, we proceed to determine the 'shape parameter' Γ. With a χ^2 likelihood analysis using the observational power spectrum reconstructed by Peacock and Dodds [33], we find a best fit for the shape of the long string power spectrum (6.21) with

$$\Gamma = 0.074 \times 10^{\pm 0.1}, \tag{6.24}$$

at the 95 per cent confidence level. This value is about three times smaller than the shape parameter in the standard CDM model $\Gamma \approx 0.23$ [131].

As for the HDM results, the comparison with observation seems to require a strongly scale-dependent bias for any choice of the cosmological parameters (models I–V). Further investigation using a hydrodynamical code will be required to investigate if galaxies form early enough. We also notice that the results presented here are in qualitative agreement with semi-analytic results [118, 121, 122, 132], with unpublished matter era simulations [133], and with those based on a phenomenological string model [107].

6.4.6 Loop effects

Now let us include the effect of loops. Figure 6.10 shows the power spectrum of density perturbations induced by long strings and by cosmic string loops for $f = 1$ (see section 5.4) with a small dynamic range from 2.5 to $5\eta_{\mathrm{eq}}$. We can see that when compared with the spectrum induced by static loops (dot dashed), the amplitude of small-scale perturbations induced by moving loops (dashed) is clearly reduced by their motion. However, their large-scale power is higher because of the dependence of the gravitational interaction on the loop velocities, especially when they are relativistic. We also see that the gravitational decay of the loop energy (thin solid) damps the overall amplitude of the power spectrum (dashed) by about a factor of 3. We notice that, between the long-string correlation scale $k_\xi \approx 20/\eta$ (see section 5.3.2) and the scale $k_{\mathrm{L}} \approx 10k_\xi$, the slope of the loop spectrum (thin solid) is exactly the same as that of the long-string spectrum $n \approx -2.25$. We have shown in section 5.4 that this close correspondence is due to copious loop production being strongly correlated with

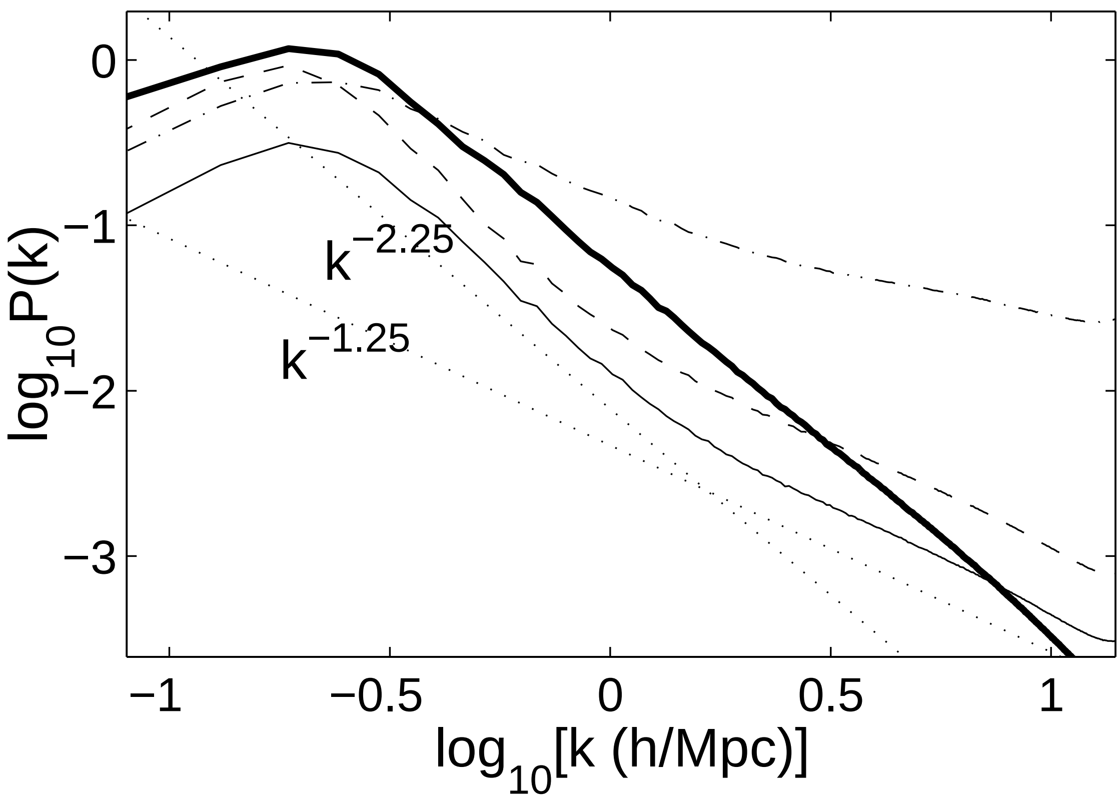

Figure 6.10: The effect of loop free-streaming and gravitational decay in the resulting CDM power spectrum. With a dynamic range of two, these power spectra of density perturbations are seeded by long strings (thick solid), by loops with initial velocities v_* switched to zero (dot-dashed), and by loops with v_* determined by string network evolution (dashed and thin solid). The thin solid line includes the effect of gravitational decay of the loop energy, while the other two loop lines don't but with loops removed after a period of time $\tau_* = t_*$.

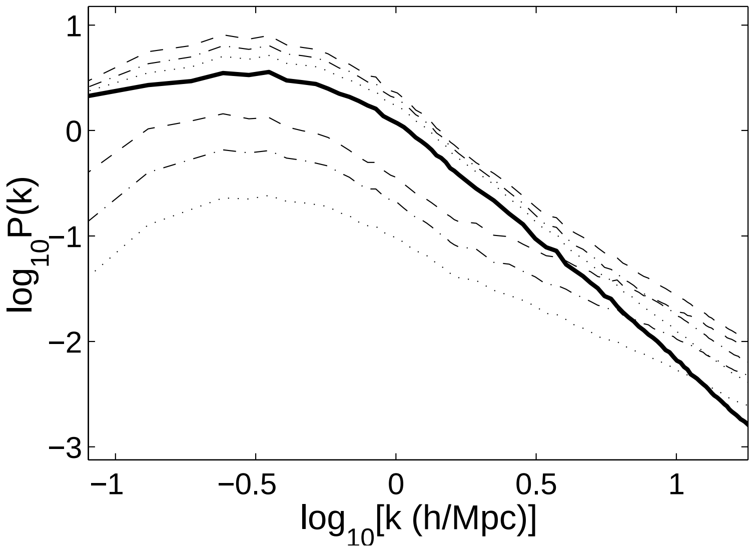

Figure 6.11: CDM power spectra seeded by loops, by long strings, and by both. The lower set of 3 lines are $\mathcal{P}_{\rm L}(k)$ for $f = 0.5$ (dotted), 1 (dot-dashed) and 2 (dashed). $\mathcal{P}_\infty(k)$ is plotted as a solid line. The upper set of lines are $\mathcal{P}_{\rm tot}(k)$ with corresponding line styles and f values to the lower set of lines.

long string intercommuting events and the collapse of highly curved long-string regions [103], that is, near the strongest long-string perturbations. Moreover, these correlations persist in time with the subsequent motion of loops and long strings lying preferentially in the same directions, a phenomenon which has been verified by observing animations of string network evolution. These correlations between loops and long strings, however, have a lower cutoff represented by the mean loop spacing $d_{\rm L} \sim k_{\rm L}^{-1}$. Below $d_{\rm L}$, the effects of individual filaments swept out by moving loops can be identified. In terms of the power spectrum, for $k < k_{\rm L}$ the loops are strongly correlated with the long strings and therefore reinforce the wake-like perturbations, while for $k > k_{\rm L}$ their filamentary perturbations increase the spectral index from about one to $n \approx -1.25$. This change is expected on geometrical grounds.

In Figure 6.11 we plot the power spectra of density perturbations seeded by long strings $\mathcal{P}_\infty(k)$, by small loops $\mathcal{P}_{\rm L}(k)$, and by both loops and long strings $\mathcal{P}_{\rm tot}(k)$. The dynamic range here extends from 0.6 to $7.5\eta_{\rm eq}$. As expected $\mathcal{P}_{\rm L}(k)$ scales more moderately than f^2 but more strongly than f (see (5.20)). It is also apparent that the perturbations induced by long strings and by loops are positively correlated with $\mathcal{P}_{\rm tot}(k) > \mathcal{P}_{\rm L}(k) + \mathcal{P}_\infty(k)$ throughout the whole scale range. This positive correlation between loops and long strings boosts the large-scale $\mathcal{P}_\infty(k)$ by a factor of 1.5, 1.8 and 2.2 to reach $\mathcal{P}_{\rm tot}(k)$ for $f = 0.5, 1$ and 2 respectively, even if $\mathcal{P}_{\rm L}(k)$ is a relatively small fraction of $\mathcal{P}_\infty(k)$ on these scales.

Figure 6.12 shows the correlation coefficient between the long-string- and loop-induced perturbations:

$$\mathcal{K} = \frac{\langle \delta_\infty \delta_{\rm L} \rangle}{\langle \delta_\infty^2 \rangle^{1/2} \langle \delta_{\rm L}^2 \rangle^{1/2}} \ . \tag{6.25}$$

We see that long strings and loops are strongly positively correlated on large scales, but weakly correlated on small scales where the loops dominate the perturbations (also see Figure 6.11). The threshold $k_{\rm t}$ between these two regimes must be significantly larger than $k_{\rm L}$ because, for $k < k_{\rm L}$, $\mathcal{P}_{\rm L}(k)$ is well below and roughly parallel to $\mathcal{P}_\infty(k)$ (see Figures 6.10 and 6.11). We also verify that $\mathcal{P}_{\rm tot}(k)/\mathcal{P}_\infty(k)$ is approximately constant for $k < k_{\rm L} \lesssim k_{\rm t}$, which again provides strong evidence for the fact that loops behave as part of the long-string network on large scales.

Given these properties of the string power spectra, one can easily construct a semi-analytic model for $\mathcal{P}_{\rm tot}(k)$ as for $\mathcal{P}_\infty(k)$. We first multiply the structure function $\mathcal{F}(k,\eta)$ of $\mathcal{P}_\infty(k)$ by $\mathcal{J}(\eta, f)$ to account for the boost $\mathcal{P}_{\rm tot}(k)/\mathcal{P}_\infty(k)$ on large scales ($k < k_{\rm t}$):

$$\mathcal{J}(\eta, f) = \bar{\mathcal{J}}(f) \left[1.4 + \frac{0.6}{1 + \left(\frac{\eta}{7\eta_{\rm eq}}\right)^{1.5}} \right] , \tag{6.26}$$

where $\bar{\mathcal{J}}(f) = 0.806$, 1, 1.22, 1.44 and 1.56 for $f = 0.5$, 1, 2, 4 and 6 respectively. We then multiply it again by

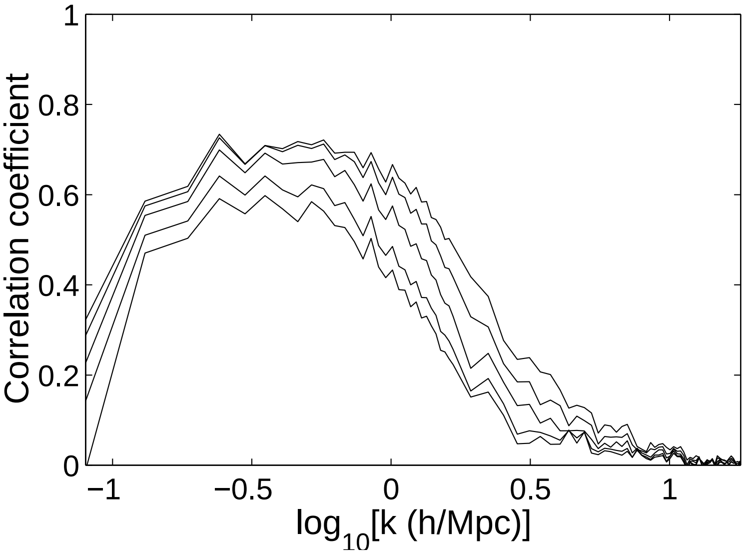

Figure 6.12: The correlation coefficients between the long-string- and loop-induced CDM perturbations, with $f = 0.5, 1, 2, 4, 6$ (downwards).

a numerically verified factor

$$\mathcal{H}(k, \eta, f) = \left[1 + \left(\frac{k}{k_{\mathrm{t}}}\right)^4\right]^{1/4}, \tag{6.27}$$

to account for the turnover for $k > k_{\mathrm{t}}(\eta, f)$. $\mathcal{J}(\eta, f)$ is calibrated phenomenologically from simulations deep in the radiation era through to those deep in the matter era. In the pseudo-scaling regime for the loop size, k_{t} is revealed to be at least $10k_\xi \approx 200/\eta$ depending on f. Thus we can carry out a full-dynamic-range integration to obtain $\mathcal{P}_{\mathrm{tot}}(k)$. In Figure 6.13 we compare this $\mathcal{P}_{\mathrm{tot}}(k)$ and $\mathcal{P}_\infty(k)$ with observations [33]. The background cosmology is $\Omega_{\mathrm{c}} = 0.15$, $\Omega_\Lambda = 0.85$ and $h = 0.7$, and we have used the COBE normalization $G\mu = 1.7 \times 10^{-6}$ [119] throughout. Since loops are point-like, they have little impact through the Kaiser-Stebbins effect on COBE-scale CMB anisotropies, so we expect this normalization to be weakly dependent on the value of f; indeed, loops were found to be negligible in Ref. [134]. Thus we see from Figure 6.13 that for $f \gtrsim 0.5$, loops can contribute significantly to the total power spectrum, easing the large-scale bias problem seen previously [1, 5, 106]. Definite conclusions, therefore, about biasing in cosmic string models will need further advances in determining the magnitude of the parameter f, while all future large-scale structure simulations will now require the inclusion of loops.

These additional complications in modelling cosmic string structure formation are most obvious on small scales, where even higher resolution and large dynamic range simulations will be required. However, we expect the power spectrum on large scales to be only weakly dependent on the details of loop formation. Within the present pseudo-scaling regime for loop size, we know that $k_{\mathrm{t}} \gtrsim k_{\mathrm{L}} \gtrsim 10k_\xi$ as shown in Figure 6.10 and discussed previously. Taking this extreme minimum $k_{\mathrm{t}} = 10k_\xi$, then, we find that the semi-analytic model over the full dynamic range gives at most a 2% difference in $\mathcal{P}_{\mathrm{tot}}(k)$ for $k < 1h\mathrm{Mpc}^{-1}$ when the filament term $\mathcal{H}(k, \eta, f)$ is excluded from $\mathcal{F}(k, \eta)$ (for $\Omega_{\mathrm{c}} = 0.15$, $\Omega_\Lambda = 0.85$ and $h = 0.7$). This means that although the simulations described in this chapter are already on the verge of present computer capabilities, a further detailed study on small scales will improve only the overall normalization of $\mathcal{P}_{\mathrm{tot}}(k)$ but not the shape revealed here, if f is constant throughout the whole history of the string evolution. We notice that although f should be constant in both deep radiation and deep matter eras due to the scaling of string network, it may vary over the transition era. This time dependence in f could slightly alter the shape of $\mathcal{P}_{\mathrm{tot}}(k)$ given here, and this motivates further investigation in cosmic string evolution. We also note that advances in understanding loop formation mechanisms will also be crucial in quantifying the importance of the gravitational radiation background emitted by a cosmic string network and its effect on large-scale structure and CMB anisotropies [127].

6.4.7 The effect of baryons

In all the analysis carried out so far, we have not included baryons. To this end, we use the formalism (6.11) and (6.18) with the Green function $\widetilde{\mathcal{G}}^{\mathrm{cs}}(k; \eta_0, \eta)$ numerically obtained from equations (6.1) and (6.2), in which

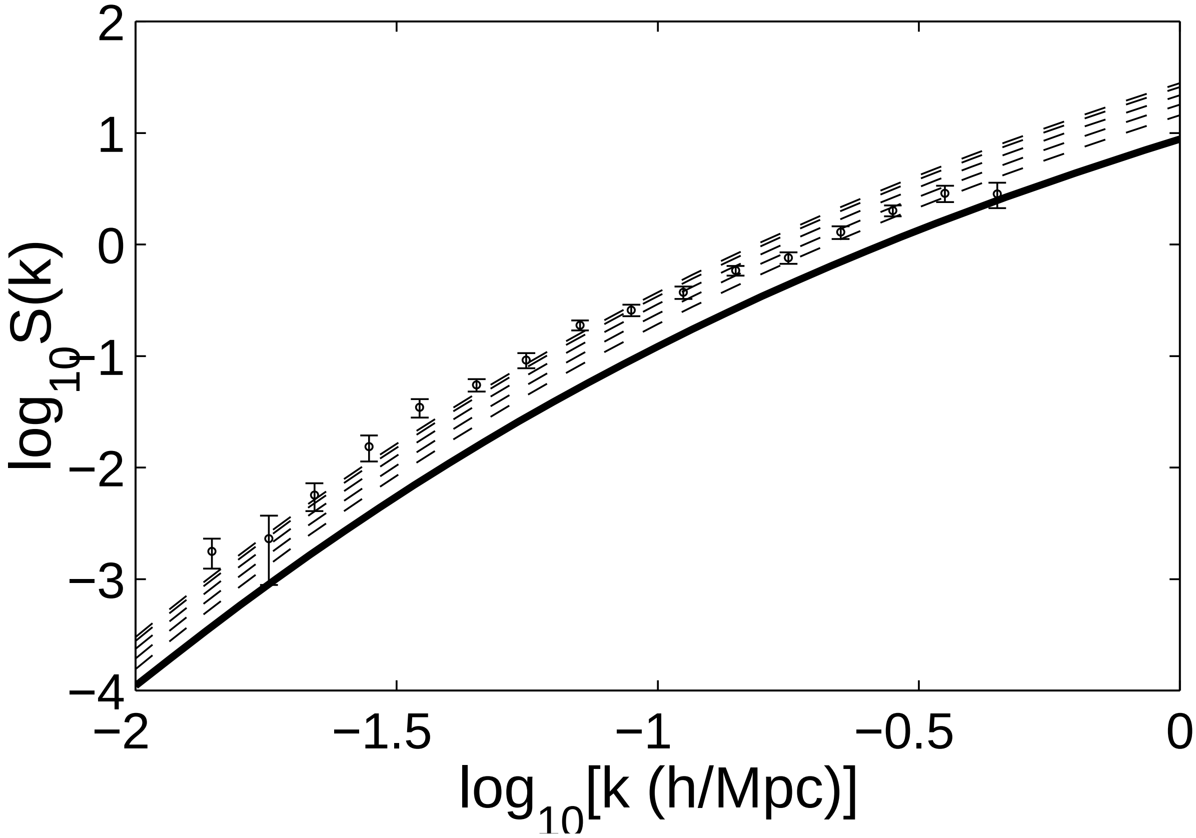

Figure 6.13: Comparison of the observational power spectrum [33] with $\mathcal{P}_\infty(k)$ (solid), and $\mathcal{P}_{\rm tot}(k)$ for $f = 0.5, 1, 2, 4, 6$ (dashed, upwards), with a full dynamic range. The background cosmology is $\Omega_{\rm c} = 0.15$, $\Omega_\Lambda = 0.85$ and $h = 0.7$, and we have used the COBE normalization $G\mu = 1.7 \times 10^{-6}$ [119].

we now take $\Omega_{\rm B0} \neq 0$ and thus $R \neq 0$. In Figure 6.14, we plot the resulting string-induced CDM power spectra, with different choices of baryon energy density. The solid lines in the upper panel are flat $\Lambda \neq 0$ models, with $\Omega_{\rm m0} = \Omega_{\rm c0} + \Omega_{\rm B0} = 0.2$ and $\Omega_{\rm B0} = 0,\ 0.05,\ 0.1$ (right to left); the solid lines in the lower panel are open models, with $\Omega_{\rm m0} = \Omega_{\rm c0} + \Omega_{\rm B0} = 0.2$ and $\Omega_{\rm B0} = 0,\ 0.05,\ 0.1$ (right to left). The dotted lines in both plots are flat $\Lambda = 0$ models, with $\Omega_{\rm B0} = 0,\ 0.05,\ 0.1$ (right to left). Here we have used $G\mu_6 = 1.7$. As one can see, the inclusion of baryons results in a nearly scale-independent suppression in the power spectrum at small scales (large k). This confirms the point we illustrated in Chapter 4 that the effect of baryons in CDM perturbations is to nearly uniformly reduce the power inside the sound horizon of the decoupling epoch. From Figure 6.14, we see that with the same choice of $\Omega_{\rm m0}$, the more baryons, the lower the small-scale power. By comparing the solid and dotted lines, we also see that baryonic effects are more severe in cosmologies with lower $\Omega_{\rm m0}$. Since the baryonic suppression in the power spectrum is nearly scale-independent on small scales, we can combine the above two observations to yield an effective shape parameter of the form

$$\Gamma = \Omega_{\rm m0} h \exp(-2\Omega_{\rm B0} - \frac{\Omega_{\rm B0}}{\Omega_{\rm m0}}). \tag{6.28}$$

This fit to our numerical results has an accuracy of better than 10% in the range $k = 0.01$–$100h{\rm Mpc}^{-1}$, provided that the baryon to dark matter ratio is relatively small ($2\Omega_{\rm B0} + \Omega_{\rm B0}/\Omega_{\rm m0} \lesssim 0.3$). For a baryon energy density $\Omega_{\rm B0} = 0.05$, a χ^2 likelihood analysis using the observational data in Ref. [33] further yields

$$\Omega_{\rm m0} h = 0.14 \times 10^{\pm 0.1}, \tag{6.29}$$

at the 95% confidence level. This appears to be within an acceptable cosmological parameter range. We note that equation (6.28) is slightly different from the form $\Gamma = \Omega_{\rm m0} h \exp(-\Omega_{\rm B0} - \Omega_{\rm B0}/\Omega_{\rm m0})$ [135], which is commonly used in inflationary models.

6.4.8 Implications in CMB anisotropies

A key feature of all these string-induced power spectra is the influence of the slow relaxation to the matter era string density from the much higher radiation string density, which has an effective structure function $\mathcal{F}(k,\eta)$ in (6.18) with approximately 2.5 times more power than the matter era version. Even by recombination in an $\Omega_{\rm m0} = 1$ ($h = 0.7$) cosmology, the string density $\zeta(\eta)$ is more than twice its asymptotic matter era value to which we normalize on COBE scales (see Figure 5.5). This implies that the string model provides higher than expected large-scale power around $100h^{-1}$Mpc and below. Interestingly, this can also be expected to produce a significant Doppler-like peak on small angle CMBR scales, an effect noted in Ref. [119] but not observed because matter era strings were employed. Recent work in Ref. [107] using a phenomenological semi-analytic approach confirms that such Doppler-like features can result from significant non-scaling effects during the transition era.

6.4.9 Uncertainties

Finally, it is appropriate to comment on the key uncertainties affecting these calculations. The most recent COBE string normalization is $G\mu_6 \approx 1.7$ [119], which is at variance with a previous COBE normalization $G\mu_6 \approx 1$ using the same string simulation [134]. This remains to be satisfactorily resolved, but systematic relativistic effects seem likely to have affected the earlier result. Next there is uncertainty implementing the effective compensation scale at which the perturbations are cut-off on large scales, which we have discussed already. Thirdly, there are uncertainties in the long-string energy density as we have addressed in section 5.3.5. This causes an uncertainty of about 10% in the resulting power spectrum. All these uncertainties add up to a factor of about 2 in the normalisation of the final power spectrum, while the shape of the spectrum might also be altered once a proper treatment of the compensation is included. Finally, there are loop effects which remain uncertain. As we have seen, the high correlation between loops and long strings may boost the power spectrum up to a factor of 3, if the loop lifetime is not much smaller than the Hubble time. In addition, if the loop lifetime varies significantly over the radiation-matter transition, the spectrum shape will also be changed. We also note that the exact contribution of the background of gravitational radiation emitted by cosmic string loops still remains a significant uncertainty [127]. Therefore, to include the effect from loops properly has become an important challenge to be addressed in future work. This consideration will be necessary for any further development in studying structure formation by cosmic strings.

6.5 Conclusion

In this chapter we described the results of high-resolution numerical simulations of structure formation seeded by a local cosmic string network, with a large dynamical range spanning from the radiation era through to deep

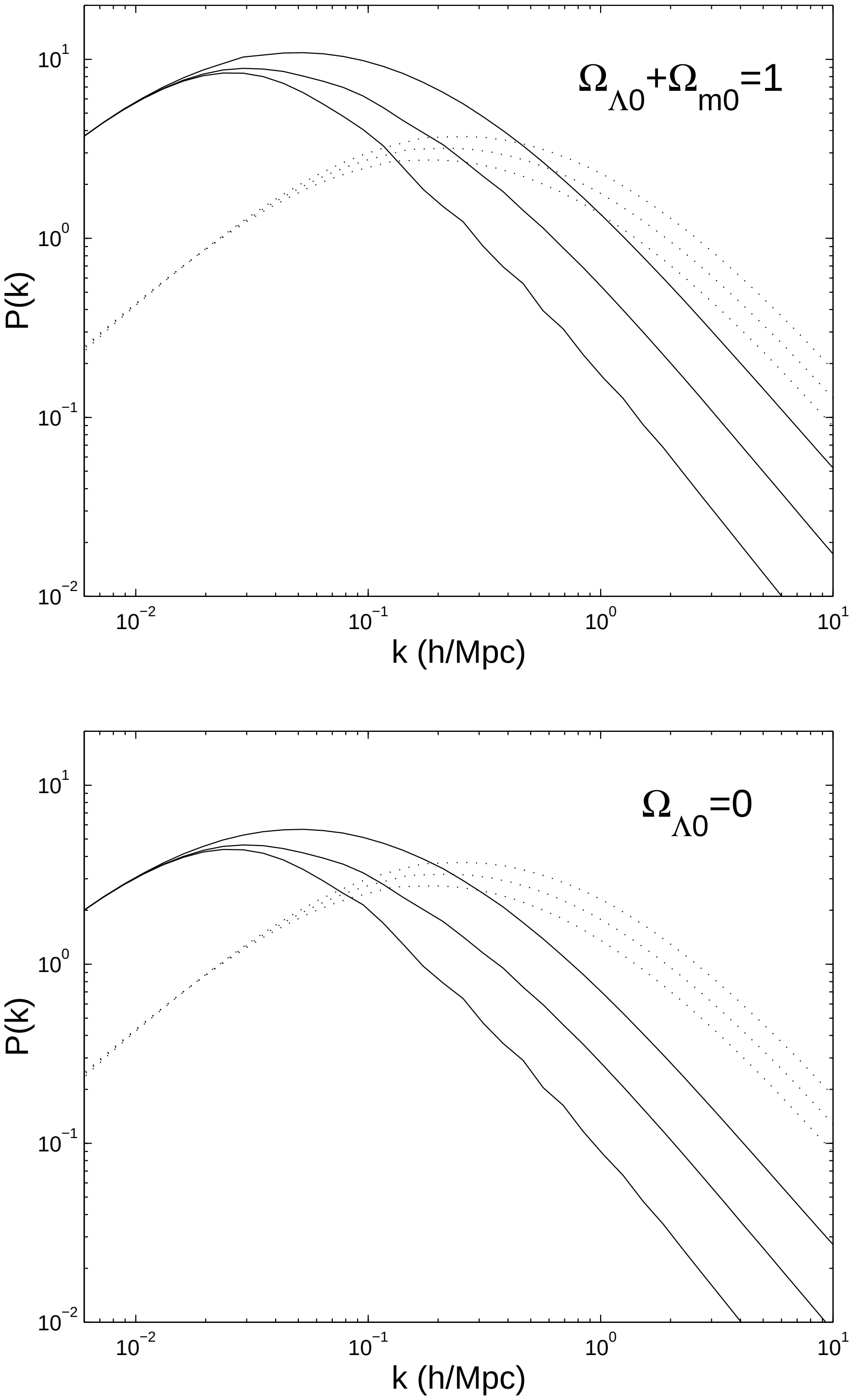

Figure 6.14: Baryonic effects in the string-induced CDM power spectra. The solid lines in the upper panel are flat $\Lambda \neq 0$ models, with $\Omega_{\rm m0} = \Omega_{\rm c0} + \Omega_{\rm B0} = 0.2$ and $\Omega_{\rm B0} = 0$, 0.05, 0.1 (right to left); the solid lines in the lower panel are open models, with $\Omega_{\rm m0} = \Omega_{\rm c0} + \Omega_{\rm B0} = 0.2$ and $\Omega_{\rm B0} = 0$, 0.05, 0.1 (right to left). The dotted lines in both plots are flat $\Lambda = 0$ models, with $\Omega_{\rm B0} = 0$, 0.05, 0.1 (right to left). We have used $G\mu_6 = 1.7$.

into the matter era. We also presented a semi-analytic model which can accurately reproduce our simulation results. In the regime of large-scale structure formation, the most serious problem with both cosmic strings and the standard CDM model is that they produce too much small-scale power and insufficient large-scale power. In the analysis presented, we saw that for cosmic strings this problem can be relaxed a great deal by considering an open universe or the existence of a cosmological constant. In particular, we showed that the best fit to the observations [33] yields a shape parameter $\Gamma = 0.074 \times 10^{\pm 0.1}$ (95% confidence level). With $\Omega_{\mathrm{B0}} = 0.05$, we further obtained $\Omega_{\mathrm{m0}} h = 0.14 \times 10^{\pm 0.1}$ for a best fit (95% confidence level). Baryonic effects here induce a nearly scale-independent suppression at small scales in the matter power spectrum. Therefore when the baryon fraction is small ($2\Omega_{\mathrm{B0}} + \Omega_{\mathrm{B0}}/\Omega_{\mathrm{m0}} \lesssim 0.3$), there is an effective shape parameter $\Gamma = \Omega_{\mathrm{m0}} h \exp(-2\Omega_{\mathrm{B0}} - \Omega_{\mathrm{B0}}/\Omega_{\mathrm{m0}})$.

As to the shortage of the overall amplitude in the string power spectrum when compared with the observations, we saw that this can be overcome by including cosmic string loops whose lifetime is comparable to the Hubble time or greater. This is because on large scales the loops behave like part of the long-string network and therefore contribute significantly to the total power spectrum of density perturbations. At present, the typical size and lifetime of loops formed by a string network remains to be studied in more detail, and this problem is both computationally and analytically challenging. We note that a longer loop lifetime will result in more power in the matter perturbations, while a significantly varying loop lifetime over the radiation-matter transition will potentially alter the power spectrum shape calculated here.

We conclude that although more work needs to be done, notably in improving the implementation of compensation and in the study of loops, the picture which emerges from the work here has advanced our knowledge of structure formation with cosmic strings. In the next two chapters, we shall proceed to apply these results so as to further our understanding of string-induced matter perturbations.

Chapter 7

Non-Gaussian Tests of LSS for Cosmic Strings

7.1 Introduction

As reviewed in Chapter 2, the two main candidates for the origin of cosmic structure are inflation and topological defects. Although, as previously seen, both scenarios may produce a power spectrum of CDM perturbations consistent with observations, they have very different predictions regarding the statistical properties of the density field. While most inflationary models produce Gaussian random-phase initial conditions, defect models produce non-Gaussian perturbations, particularly on small scales. Therefore the power spectrum alone is insufficient to describe all the statistical properties of a string-induced CDM density field. This is even more important in open or Λ-models because in those models the characteristic scales of the density field are shifted to larger scales relative to a flat model with $\Lambda = 0$. Thus in this chapter, we shall investigate the non-Gaussian properties of cosmic-string-seeded CDM and HDM density perturbations.

In the following sections, we shall employ the one-point probability distribution function (PDF), its skewness, kurtosis, and genus curves for different smoothing scales, to examine the non-Gaussianity of the string-induced CDM density field resulting from our previous high-resolution simulations. The non-Gaussian properties we reveal on small lengthscales provide a significant observational signature for cosmic string-seeded structure formation models. We note that previous analytic work has investigated the string-induced velocity field on scales above several h^{-1}Mpc, which was inferred to be Gaussian [136, 137], and that some of the features we study here were also observed in global defect models, notably for textures [138]. Past work on genus statistics in the context of topological defects was made using toy models which incorporated some important features of the models studied [139, 140, 141].

The structure of the chapter is as follows. In section 7.2, we define the one-point PDF of an underlying density field, its skewness, kurtosis, and genus curves, all as functions of the smoothing scale. Section 7.3 computes these statistics using the string-induced CDM and HDM density fields simulated previously, and the non-Gaussian features of the string scenario are clearly revealed. A semi-analytic model is then used in section 7.4 to provide an appropriate physical interpretation of our results. Finally a conclusion is given in section 7.5.

7.2 Non-Gaussian test statistics

7.2.1 One-point PDF, skewness and kurtosis

With the high-resolution simulations of string-induced structure formation obtained previously, we first convolved the resulting matter density field with a top-hat window at different smoothing scales R, and then calculated its one-point PDF, $p(\nu)$, where ν is the number of standard deviations from the mean. The chosen smoothing scales were always at least three times larger than the grid spacing to guarantee sufficient resolution for the smoothing, and at least twelve times smaller than the box size to reduce sample variance effects. We then used the skewness and the kurtosis to test their Gaussianity. The skewness is defined as

$$\text{skewness} = \frac{M_3}{{M_2}^{3/2}}\,, \tag{7.1}$$

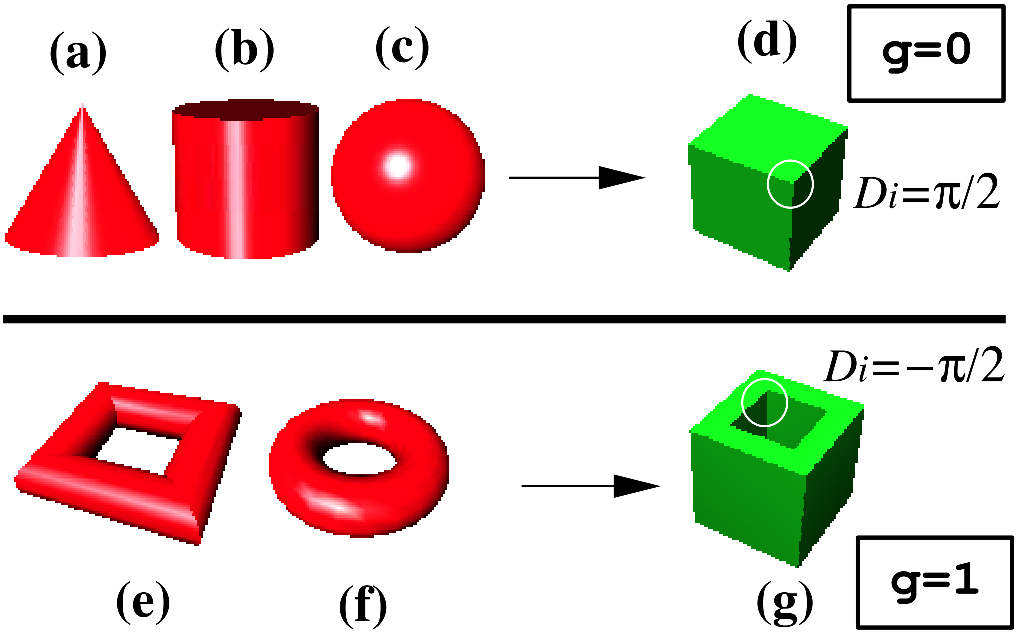

Figure 7.1: The genus of objects with simple topologies. The objects in the upper part ((a)–(d)) have genus zero, while the objects in the lower part ((e)–(g)) have genus unity.

where M_i is the i-th central moment of a PDF. Its sign indicates the skewed direction against a normal distribution. The kurtosis is defined as

$$\text{kurtosis} = \frac{M_4}{{M_2}^2} - 3\,, \tag{7.2}$$

and it measures the size of the side tails against those of a normal distribution. Both statistics are zero for a Gaussian distribution, but will depart from zero when the distribution is non-Gaussian. We can therefore estimate a 'non-Gaussian scale' R_{ng}, below which both the skewness and the kurtosis depart significantly from zero.

7.2.2 Genus statistics

Another useful statistic sensitive to the non-Gaussian features of a density field is the genus, which is employed to measure the topology of a continuous field. Conceptually, the genus of a given surface can be defined as

$$g = \text{number of holes} - \text{number of isolated regions} + 1\,. \tag{7.3}$$

A more useful definition of genus is given by the Gauss-Bonnet theorem which relates the integrated Gaussian curvature (a local property) of a surface with the genus (a global property) of that surface:

$$\oint K dA = 4\pi(1-g), \tag{7.4}$$

where dA is a differential two-dimensional surface patch, and $K = 1/a_1a_2$ is the Gaussian curvature of the patch with a_1 and a_2 the two principal radii of curvature. To calculate the genus of these isodensity surfaces, we applied the method proposed by Gott, Mellot and Dickinson [142]. Their method basically employs the definition of the genus in discrete models (c.f. equation(7.3)):

$$\sum_i D_i = 4\pi(1-g), \tag{7.5}$$

where D_i is the deficit angle of the ith element on the surface of a specified threshold. For example, the genus a cube is zero, because the only elements with non-zero D_i are on its eight vertices, each with $D_i = \pi/2$ ($i = 1$–8). This is illustrated as object (d) in Figure 7.1. Objects (a)–(c) also have the same genus, because they have the same topology—one isolated region with no holes (see equation (7.3)). If there is a hole in a cube (see object (g) in Figure 7.1), then its genus is unity, because the hole creates extra eight vertices, each with $D_i = -\pi/2$ ($i = 9$–16). Object (e) and (f) in Figure 7.1 also have the same genus due to the same topology—one isolated

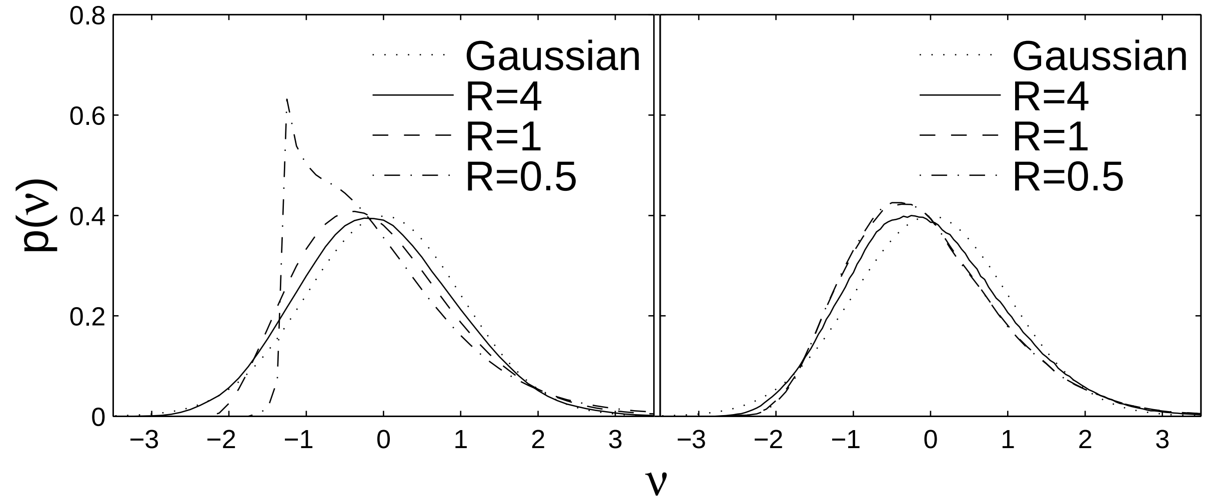

Figure 7.2: The one-point PDF $p(\nu)$ of the density field in the linear $\mathcal{S}$CDM (left) and $\mathcal{S}$HDM (right) models for different R (in $(\Omega_{\rm m0}h^2)^{-1}$Mpc). Here $\sigma_{\rm 4c} : \sigma_{\rm 1c} : \sigma_{\rm 0.5c} : \sigma_{\rm 4h} : \sigma_{\rm 1h} : \sigma_{\rm 0.5h} \approx 1 : 5.2 : 8.9 : 0.69 : 1.2 : 1.3$. All simulation curves are averaged from 6 realizations, with standard deviation less then 5% of the amplitudes throughout.

region with one hole (see equation (7.3)). For detailed numerical treatments of the genus calculation, please refer to Ref. [142].

We also employed the commonly used volume fraction parametrisation of the genus curve, that is, we define a new variable ν_f so that

$$f = \frac{1}{(2\pi)^{1/2}} \int_{\nu_f}^{\infty} e^{-u^2/2}\, du\,, \tag{7.6}$$

where f is a previously calculated volume fraction of the whole simulation box within which the density is above a given threshold. This parametrisation is to find the threshold ν_f in a Gaussian PDF above which the area of the PDF is the same as f. Having done this, the genus curve becomes less-sensitive to the PDF which we studied separately.

7.3 Results

To examine the non-Gaussianity of the string scenario, our first step in the present analysis was to perform high-resolution numerical simulations of cosmic string networks in an expanding universe from which we subsequently computed the causally-sourced density field with either a CDM or HDM background. The details of this procedure were described in the previous chapter. The cosmic string simulations had a dynamic range extending from before the radiation-matter transition at $0.4\eta_{\rm eq}$ through to deep into the matter era ($8.4\eta_{\rm eq}$). The structure formation simulation boxes contained 256^3 grid-points and their physical volume was in the range $(4\text{–}100h^{-1}\text{Mpc})^3$.

In Figure 7.2 we plot the one-point PDF $p(\nu)$ of the density field in the linear string CDM and HDM models (denoted as $\mathcal{S}$CDM and $\mathcal{S}$HDM respectively) for different smoothing scales R. The skewness and kurtosis of $p(\nu)$ as functions of R were plotted in Figure 7.3. Although string perturbations are almost indistinguishable from Gaussian random-phase perturbations on large scales, on scales smaller than $1.5(\Omega_{\rm m0}h^2)^{-1}$Mpc the non-Gaussian character is remarkably distinct, especially in a CDM background.

In Figure 7.4 we plot the genus curves $g(\nu_f)$ for $R = 0.4(\Omega_{\rm m0}h^2)^{-1}$Mpc in both $\mathcal{S}$CDM and $\mathcal{S}$HDM models, as well as those of Gaussian random fields with correspondingly identical power spectra. These curves have been smoothed using three-point boxcar smoothing (see for example [143]) and were parametrised by the volume fraction (7.6). We can observe that, on these small scales, the genus curves of the cosmic-string-seeded density fields deviate in a significant fashion from Gaussian random fields. As expected, the correlations among the phases of different Fourier modes imply a smaller number of independent regions/holes in real space and therefore a smaller genus amplitude when compared with that of a Gaussian random field. This can be seen in the comparison between Figures 7.5 and 7.6. Figure 7.5 plots the isodensity surfaces of $\nu_f = 0$ and 2 in the $\mathcal{S}$CDM model, while Figure 7.6 plots the isodensity surfaces of the same thresholds but in the field Gaussianised from the previous case. The smoothing scale is $R = 0.4(\Omega_{\rm m0}h^2)^{-1}$Mpc. We have also verified that

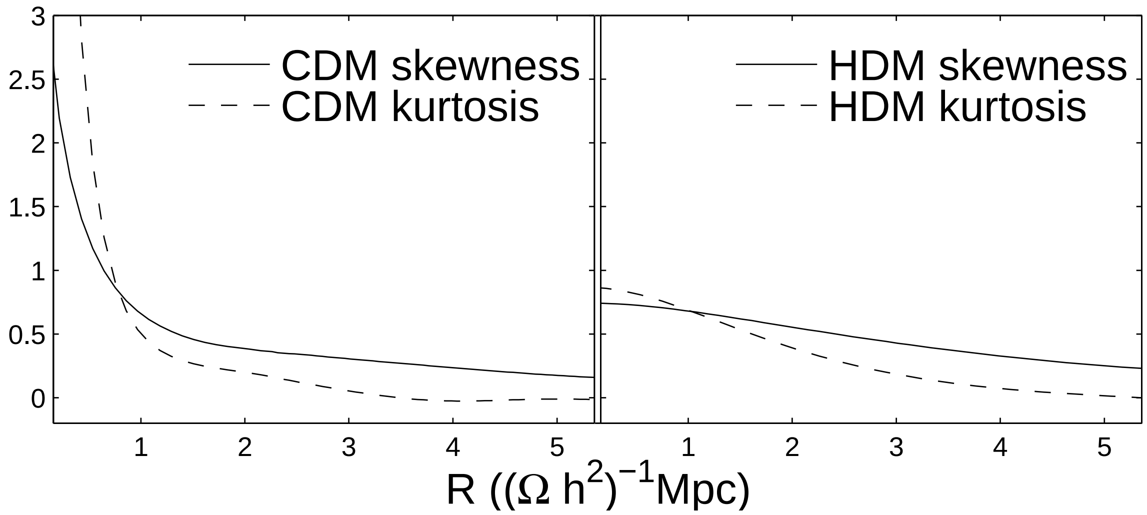

Figure 7.3: The skewness and kurtosis of $p(\nu)$ induced by cosmic strings.

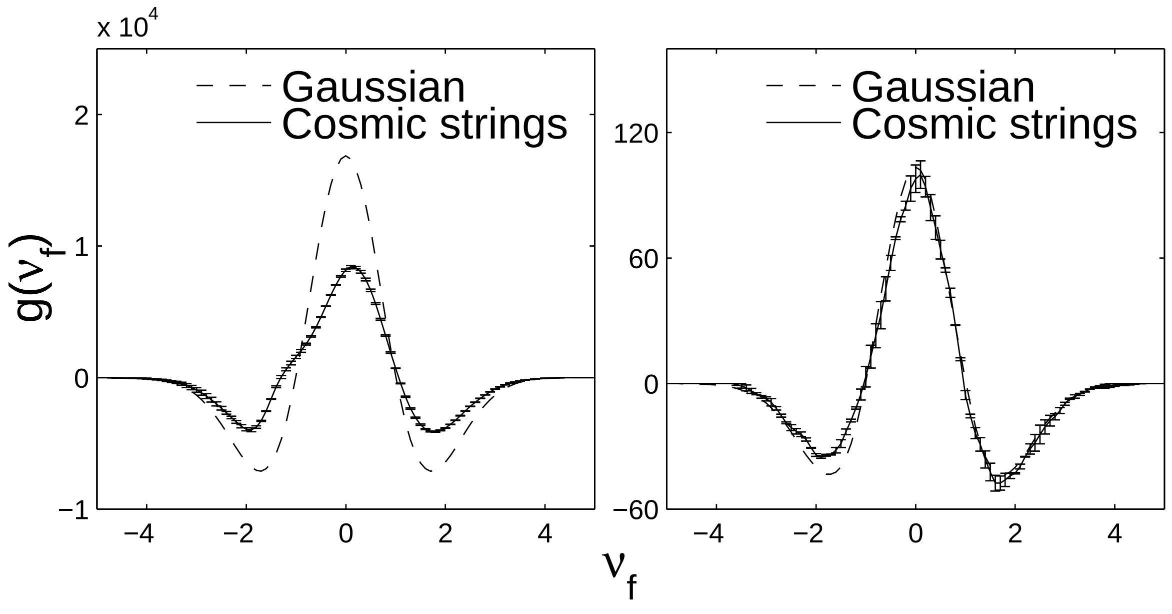

Figure 7.4: The genus curves in the CDM (left) and HDM (right) models. Here $R = 0.4(\Omega_{\rm m0}h^2)^{-1}$Mpc and the box size is $(50h^{-1}{\rm Mpc})^3$. Error bars are taken from 6 realizations for each curve.

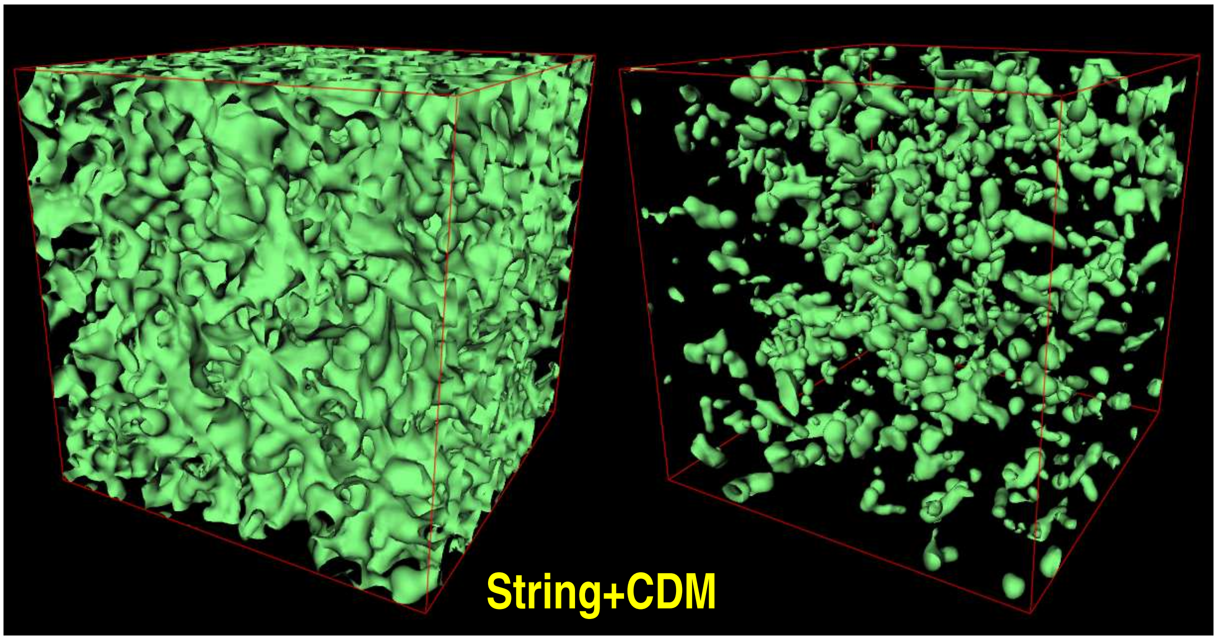

Figure 7.5: The isodensity surfaces with $\nu_f = 0$ (left) and $\nu_f = 2$ (right) for the $\mathcal{S}$CDM model. We have used the smoothing scale $R = 0.4(\Omega_{\rm m0}h^2)^{-1}$Mpc, and the box size is $(25h^{-1}{\rm Mpc})^3$.

when smoothed on larger scales $R \gtrsim 3(\Omega_{\rm m0}h^2)^{-1}$Mpc, cosmic-string and Gaussian-random-phase genus curves are very close to each other in both amplitude and shape.

Finally, we note that the departure of the $\mathcal{S}$HDM model from a Gaussian distribution is less apparent than that of the $\mathcal{S}$CDM model throughout all the analysis above. In addition, Figure 7.7 plots the isodensity surfaces of $\nu_f = 0$ and 2 in the $\mathcal{S}$HDM model. As one can see, when compared with the $\mathcal{S}$CDM model (see Figure 7.5), the much smaller genus amplitude of the $\mathcal{S}$HDM model in Figure 7.4 is reflected as the much smaller number of isolated regions/holes in Figure 7.7. We shall discuss consequences of these observations later.

7.4 A semi-analytic model and discussion

In order to interpret our results, we employ the semi-analytic model described in Chapter 6. This model was shown to be very accurate in reproducing the simulation power spectrum in both matter-era and radiation-era scaling regimes, as well as during the radiation-matter-era transition. The power spectrum of string-induced density perturbations can be written as:

$$\mathcal{P}(k) = 16\pi^2 G^2\mu^2 \int_{\eta_{\rm i}}^{\eta_0} |\widetilde{\mathcal{G}}^{\rm cs}(k;\eta_0,\eta)|^2 \mathcal{F}'(k,\eta) d\eta \,, \tag{7.7}$$

where μ is the string energy density per unit length, $\widetilde{\mathcal{G}}^{\rm cs}(k;\eta_0,\eta)$ is the Fourier transform of the appropriate Green functions, and $\mathcal{F}'(k,\eta)$ is the structure function where we have absorbed the compensation factor $\widetilde{F}(k,\eta)$ (c.f. equation (6.11)). At a given time, $\mathcal{F}'(k,\eta)$ has a turn-over scale at $k_\xi \approx 20/\eta$ which reflects the correlation length of cosmic strings $\xi \approx d_{\rm H}/3$. At a particular time, the perturbations induced on scales larger than the string correlation length are generated by many string elements and, therefore, they are nearly Gaussian according to the central limit theorem (CLT). On the other hand, perturbations induced on smaller scales are very non-Gaussian because they can be either very large within the string-induced wakes or else very small in regions outside these wakes. In consequence, we can roughly divide the power spectrum of cosmic-string-seeded density perturbations into two parts: a nearly Gaussian part generated when the string correlation length was smaller than the scale under consideration, and a strongly skewed non-Gaussian part generated when the string correlation length was larger. We shall abuse these strict definitions and refer simply to the 'Gaussian' and 'non-Gaussian' contributions respectively. In terms of the structure function in (7.7), we can make the split

$$\mathcal{F}'_{\rm g}(k,\eta) = \begin{cases} \mathcal{F}'(k,\eta), & k \le k_\xi\,, \\ 0, & k > k_\xi\,, \end{cases} \tag{7.8}$$

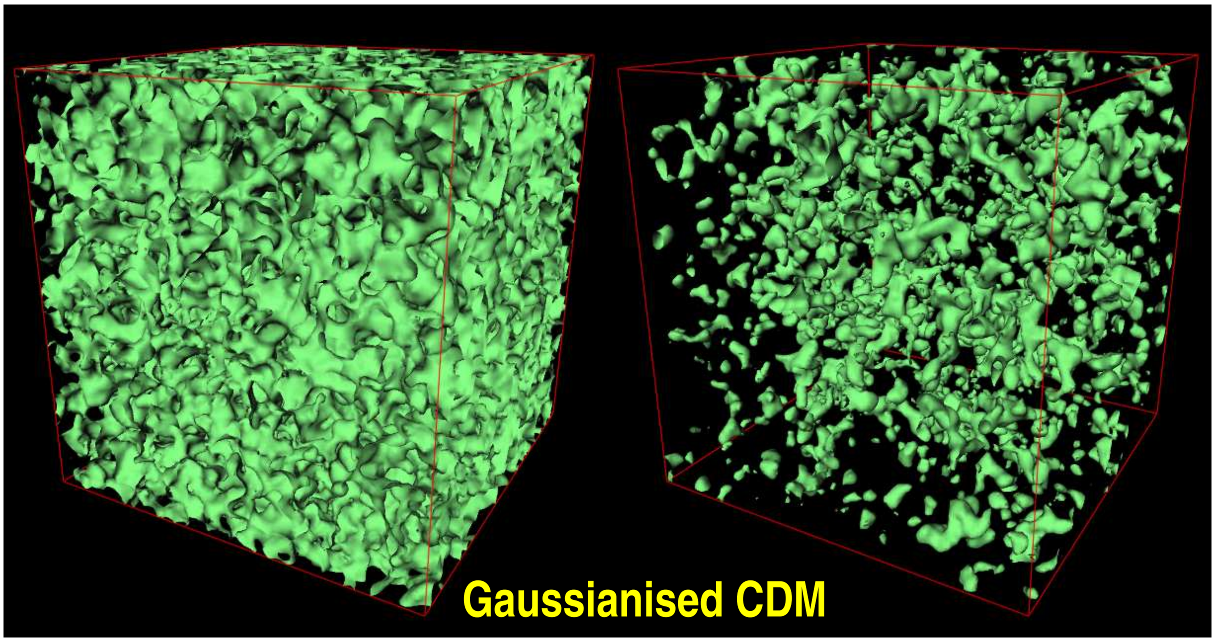

Figure 7.6: Same as Figure 7.5, but the underlying density field is that Gaussianised from the $\mathcal{S}$CDM model.

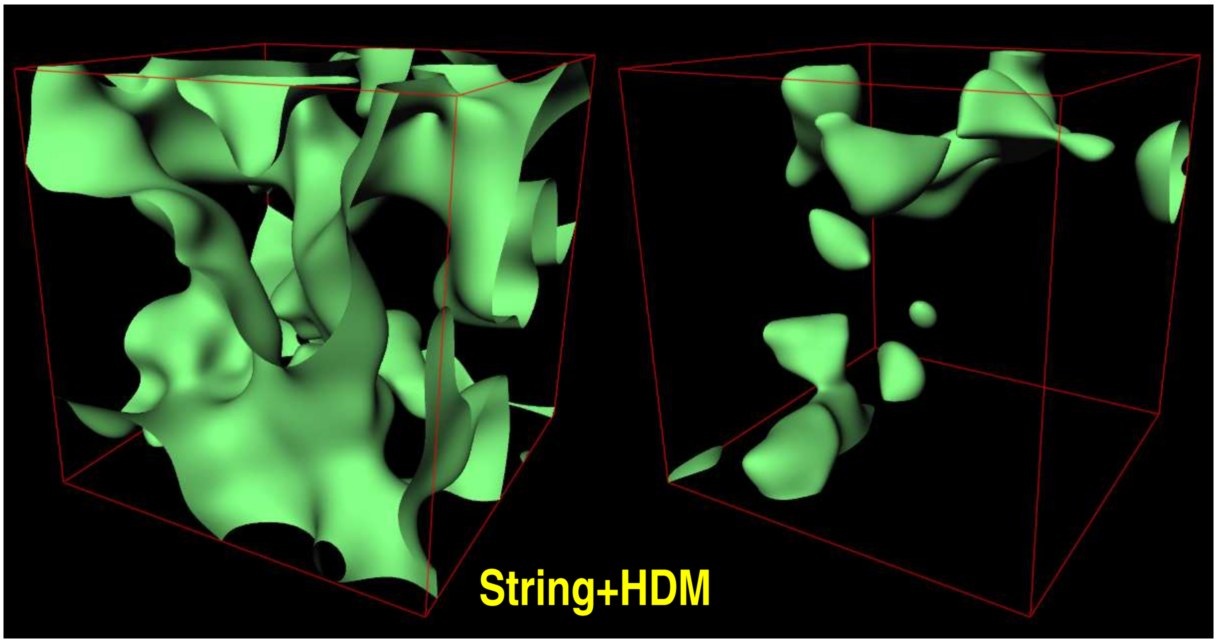

Figure 7.7: Same as Figure 7.5, but the underlying density field is from the $\mathcal{S}$HDM model.

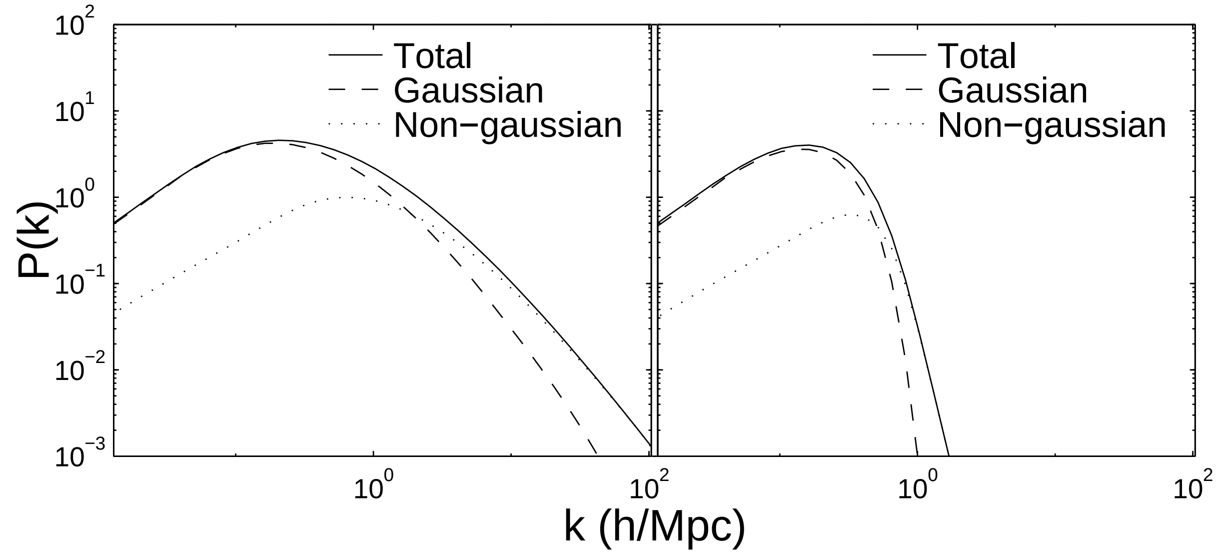

Figure 7.8: The estimated Gaussian and non-Gaussian contributions as a comparison with the total power spectrum in the $\mathcal{S}$CDM (left) and $\mathcal{S}$HDM (right) models.

and similarly

$$\mathcal{F}'_{\rm ng}(k,\eta) = \begin{cases} 0, & k < k_\xi, \\ \mathcal{F}'(k,\eta), & k \geq k_\xi. \end{cases} \tag{7.9}$$

$\mathcal{F}'_{\rm g}$ corresponds to the left-hand side of the structure function from its peak to the causal compensation cut-off where $\mathcal{F}'_{\rm g} \propto k^4$ for $k \ll k_\xi$. $\mathcal{F}'_{\rm ng}$ can be largely identified with the imprint of string wakes and behaves as $\mathcal{F}'_{\rm ng} \propto k^{-2}$ for $k > k_\xi$.

Integrating (7.7) with this Gaussian/non-Gaussian split, we can compare the relative contributions to the total power spectrum in CDM and HDM models, as illustrated in Figure 7.8. This plot demonstrates that these two contributions become comparable at $k \sim 2h\,{\rm Mpc}^{-1}$ for the CDM model and $k \sim 0.5h\,{\rm Mpc}^{-1}$ for the HDM model, taking a Hubble parameter $h = 0.7$. We further calculated the variance of the density fields from these non-Gaussian and Gaussian contributions when smoothed with a top-hat window of radius R:

$$\sigma_R^2 = 4\pi \int_0^\infty |\widetilde{w}(kR)|^2 \mathcal{P}(k) k^2 dk\,, \tag{7.10}$$

where $\widetilde{w}(x) = 3j_1(x)/x$ is the Fourier transform of a top-hat window and $j_1(x)$ the spherical Bessel function of order one. Figure 7.9 plots the ratio between the standard deviations of the non-Gaussian and Gaussian density perturbations, $\sigma_{R({\rm ng})}/\sigma_{R({\rm g})}$, against R for both CDM and HDM models. The criterion $\sigma_{R({\rm ng})}/\sigma_{R({\rm g})} \gtrsim 1$, gives an estimate for the non-Gaussian scale $R_{\rm ng}$. In Figure 7.9, we also consider the effect of the limited dynamic range of our simulations, by comparing with the full semi-analytic integral (7.7). The simulations slightly enhance non-Gaussianity essentially because there is more 'Gaussian' power missing from early times $\eta < 0.4\eta_{\rm eq}$. We have checked that the power lost $(\sigma_{\rm (full)} - \sigma_{\rm (sim)})/\sigma_{\rm (full)}$ on scales $R \geq 1(\Omega_{\rm m0}h^2)^{-1}$Mpc is always $\leq 10\%$ for the non-Gaussian contribution, $\leq 35\%$ for the Gaussian contribution, and $\leq 25\%$ for the total power. Nevertheless, as we can see in Figure 7.9, this power loss in our simulations does not significantly affect the non-Gaussian scales which we are investigating.

The consistency between Figure 7.3 and Figure 7.9 should be evident, notably for the skewness. Both plots show that the the non-Gaussian scale is $R_{\rm ng} \lesssim 1.5(\Omega_{\rm m0}h^2)^{-1}$Mpc. The physical reason for the much stronger non-Gaussian features on small scales can now be interpreted as the following. Before the radiation-matter equality $\eta_{\rm eq}$, the growth of matter density perturbations was suppressed on subhorizon scales, where strings can seed power. (see the $\widetilde{T}^{\rm cs}_{\infty}$ in Figure 4.1). This means that even on small scales the contribution to the power spectrum from the radiation-matter transition epoch and shortly afterwards is very important. Adding the fact that by the time of $\eta_{\rm eq}$ the string correlation length had already grown to several $(\Omega_{\rm m0}h^2)^{-1}$Mpc, the perturbations generated at $\eta \gtrsim \eta_{\rm eq}$ on scales smaller than this will add an important non-Gaussian contribution to the total power spectrum. This also explains the peak in the PDF on small scales (see Figure 7.2). The limited dynamic range of our simulations tends to produce a sharp peak in the PDF for small smoothing scales because of the presence of voids, that is, regions left largely unperturbed by the transition era strings. However, even for the full dynamic range a similar peak, though slightly smoother, should still emerge. This is the origin

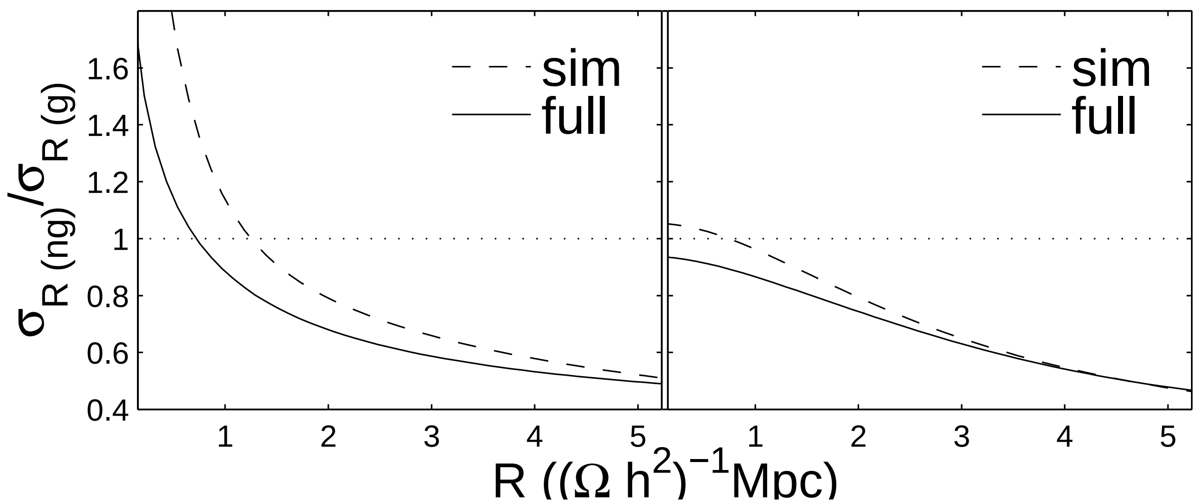

Figure 7.9: The ratio $\sigma_{R(\mathrm{ng})}/\sigma_{R(\mathrm{g})}$ against R for both $\mathcal{S}$CDM (left) and $\mathcal{S}$HDM (right) models. The solid lines have a full dynamic range going from $\eta = 0$ to today, while the dashed lines have the same dynamic range of our simulations.

of the non-Gaussianity on smaller scales, the existence of which is not affected by the limited dynamic range as we demonstrate in Figure 7.9. The reason why the HDM model is more Gaussian than the CDM model is that the neutrino free-streaming prevents small-scale perturbations from emerging until late times. This significantly reduces the HDM power at large k, precisely where the non-Gaussian contribution is dominant in the CDM model (see Figure 7.8). This also accounts for the smaller genus amplitude for the HDM model (see Figure 7.4).

We also note that our results were obtained in the linear regime. When the density field evolves into the non-linear regime at later times, the skewness increases as a direct result of the constraint that the density contrast $\delta_N \geq -1$. However, when compared with observations, the skewness of our linear cosmic-string-seeded result for any reasonable value of Γ is well below that of the QDOT survey [144]. Our results are also consistent with those in Ref. [145], as we find that cosmic-string and Gaussian-random-phase genus curves are very similar in the scale range where no strong phase correlations were found by the genus analysis of the PSCz survey. For Gaussian models, it is well known that in the non-linear regime the skewness is equal to its initial value plus a contribution due to the non-linear collapse. This contribution scales proportionally to the RMS density fluctuation σ [146]. Similarly, the non-linear contribution to the kurtosis also has a simple scaling dependence on σ, depending on whether or not the initial density field is Gaussian [147]. This indicates that non-linear clustering preserves the non-Gaussianity inherent in the linear evolution, but further investigation using N-body simulations will be required to establish if the non-Gaussian signature of cosmic string models is potentially measurable.

Finally we should comment on the baryonic effects and the manner in which the uncertainties in the string-induced power spectrum affect our results for the non-Gaussian scale R_{ng}. In Chapter 6, we saw that the inclusion of baryons would result in a nearly scale-independent suppression in the power spectrum on large k. If the baryon density fraction $\Omega_{\mathrm{B0}}/\Omega_{\mathrm{m0}}$ is not too high, this effect is equivalent to shifting the shape of the power spectrum to the smaller k. Therefore we expect the non-Gaussian scale R_{ng} to become larger once baryons are included. We should also bear in mind that the current uncertainties in the string-induced power spectrum could always shift the non-Gaussian scale predicted here. Nevertheless, the order of R_{ng} revealed here should not be significantly changed since it reflects more or less the correlation length of strings at the epoch of radiation-matter equality. In addition, together with a better understanding of the compensation scale and thus a more accurate string-induced power spectrum which will be obtained in the future, we can readily employ the semi-analytic model presented here to reliably locate the non-Gaussian scale.

7.5 Conclusion

We conclude that on lengthscales smaller than approximately $1.5(\Omega_{\mathrm{m0}}h^2)^{-1}$Mpc, the perturbations seeded by cosmic strings are very non-Gaussian, especially in the context of a CDM model. In an open or Λ-universe with $\Gamma = \Omega_{\mathrm{m0}}h \sim 0.15$, this scale will be shifted to $10h^{-1}$Mpc, which may still be in the linear or mildly non-linear regime, thus potentially providing a strong empirical test for cosmic string models. Although the current uncertainties in the estimated string-induced power spectrum may shift the non-Gaussian scale predicted here, the generic non-Gaussian features of the string scenario revealed in this chapter should not be changed. These include the order of the non-Gaussian scale which reflects the string correlation length at $\eta_{\rm eq}$, the much stronger non-Gaussianity towards smaller scales (the strong skewness and kurtosis, the smaller genus amplitude when compared with a Gaussian case, etc.), and the nearly Gaussian distribution on sufficiently large scales. Once a better understanding of the compensation scale and thus a more accurate string-induced power spectrum are obtained, the semi-analytic scheme proposed in this chapter should readily provide an accurate way to locate the non-Gaussian scale.

Chapter 8

Cluster Abundance in Cosmic String Models

8.1 Introduction

One of the most important constraints on models of structure formation is the observed abundance of galaxy clusters. Although the cluster abundance has been widely used to constrain cosmological models with primordial Gaussian fluctuations [28, 129, 131, 148, 149, 150] there have been few studies in the context of non-Gaussian perturbations such as those generated by topological defects. This is due to the difficulty of making robust predictions in topological defect scenarios, owing to their non-linear effects which are difficult to model and require large-scale numerical simulations. Nevertheless, in this chapter, we aim to use the present observed number density of large X-ray clusters to constrain the string-induced σ_8, in both open cosmologies and flat models with a non-zero cosmological constant.

In section 8.2, we first recall the string-induced CDM power spectrum estimated in Chapter 6 to determine the scale dependence of the RMS amplitude of the CDM perturbations $\sigma(R)$. We then use our previous results from high-resolution simulations to find a fit for the one point probability distribution functions (PDF) of the string-induced CDM density perturbations. In section 8.3, we then employ a simple generalization of the Press-Schechter formalism [151], which is suitable for non-Gaussian perturbations with a general one point PDF [152], in order to obtain an expression for the expected number density of collapsed objects with a mass greater than a given threshold. This generalised Press-Schechter formalism has been used to constrain the Gaussianity of the density fluctuations in the universe [153, 154], and has been verified for a simplified flat space cosmic string model [155]. In order to use the generalised Press-Schechter formalism, section 8.4 recalls the results in Ref. [129, 131] to relate the X-ray temperature of a cluster with its virial mass, as well as the observational results for the number density of galaxy clusters [129, 156, 157]. Finally in section 8.5, we estimate the amplitude of matter density perturbations induced by cosmic strings on the standard scale of $8h^{-1}$Mpc, using the presently observed number density of large X-ray clusters. We do this for cosmic string models in open universes without a cosmological constant ($\mathcal{S}$OCDM), and also in flat universes with a non-zero cosmological constant ($\mathcal{S}\Lambda$CDM). We compare the σ_8 constrained by cluster abundance with σ_8 constrained by COBE, and show that they are consistent within present uncertainties.

8.2 Power spectrum and PDF

We first recall some results from Chapter 6, where we described the high-resolution numerical simulations of string-induced structure formation. The power spectrum of CDM perturbations induced by long strings was found to be approximately described by

$$S_{\rm CDM}(q) \equiv 4\pi k^3 \mathcal{P}(k) \propto (0.7q)^{p(q)}, \tag{8.1}$$

where

$$p(q) = 3.9 - \frac{2.7}{1 + (2.8q)^{-0.44}}\,, \tag{8.2}$$

$q = k/\Gamma$ and the shape parameter Γ is defined as

$$\Gamma = \Omega_{\rm m0} h \exp(-2\Omega_{\rm B0} - \frac{\Omega_{\rm B0}}{\Omega_{\rm m0}}). \tag{8.3}$$

This fit to our numerical results has an accuracy of better than 10% in the range $k = 0.01$–$100h\mathrm{Mpc}^{-1}$, provided that the baryon to dark matter ratio is relatively small ($\Omega_{\mathrm{B0}}/\Omega_{\mathrm{m0}} \lesssim 0.3$). As noted in Chapter 6, however, there are specified uncertainties in the underlying numerical results (mainly the compensation), which are significantly larger. This result also does not include the contribution from small loops which could significantly enhance the overall power while leaving the overall shape unchanged. These uncertainties will affect our final results as we shall discuss later. A χ^2 analysis using the observational power spectrum reconstructed by Peacock and Dodds [33] gave a best fit to the long string power spectrum (8.1) with $\Gamma = 0.074 \times 10^{\pm 0.1}$ at the 95 per cent confidence level. For a baryon energy density $\Omega_{\mathrm{B0}} = 0.05$, it further yields that string models provide a consistent match to observations in the acceptable cosmological parameter range $\Omega_{\mathrm{m0}}h = 0.14 \times 10^{\pm 0.1}$.

By using equation (8.1), we can obtain a numerical fit for the standard deviation of matter density perturbations at the present time as a function of the smoothing scale R (in units of h^{-1}Mpc):

$$\sigma(R) = \sigma_8 \frac{\varrho(R\Gamma)}{\varrho(8\Gamma)}\,, \tag{8.4}$$

where

$$\varrho(r) \;=\; r^{-\gamma(r)}, \quad \gamma(r) = 0.5 + \frac{1.48}{1 + (6/r)^{0.3}}\,. \tag{8.5}$$

The red-shift dependence of σ for arbitrary Ω_{m}, Ω_Λ is described accurately by

$$\sigma(R,z) = \sigma(R)\frac{g(\Omega_{\mathrm{m}},\Omega_\Lambda)}{(1+z)g(\Omega_{\mathrm{m0}},\Omega_{\Lambda 0})}\,, \tag{8.6}$$

where

$$g(\Omega_{\mathrm{m}},\Omega_\Lambda) = \frac{2.5\Omega_{\mathrm{m}}}{\left[\Omega_{\mathrm{m}}^{4/7} - \Omega_\Lambda + (1+\Omega_{\mathrm{m}}/2)(1+\Omega_\Lambda/70)\right]}\,. \tag{8.7}$$

Here, $g(\Omega_{\mathrm{m}},\Omega_\Lambda)$ gives the suppression of the growth of density perturbations relative to that of a critical-density-universe [98] (see equation (4.52) and context), and we can describe the evolution of the Ω_{m} with red-shift as

$$\Omega_{\mathrm{m}} \equiv \Omega_{\mathrm{m}}(z) = \frac{\Omega_{\mathrm{m0}}(1+z)^3}{(1+z)^2(1+\Omega_{\mathrm{m0}}z - \Omega_{\Lambda 0}) + \Omega_{\Lambda 0}}\,. \tag{8.8}$$

Using the simulation results for string-induced matter perturbations, we also found an accurate fit to the positive side of the one point PDF as follows:

$$P_R(\nu) = g_n(\nu\sqrt{2n} + n), \qquad \nu \geq 0, \tag{8.9}$$

where R is the smoothing radius in units of h^{-1}Mpc, ν is the number of standard deviations from mean density, and $g_n(\chi^2)$ is the standard chi-square (χ^2) distribution, with the number of degrees of freedom n given as

$$n = 0.4\,(R\Gamma + 2)^{3.6}\,. \tag{8.10}$$

This fit is accurate within 10% error for $R \gtrsim 0.3h^{-1}$Mpc and $0 < \nu \lesssim 3$.

8.3 Modified Press Schechter formalism

The Press-Schechter formalism [151] relates the mass fraction of collapsed objects whose masses are larger than some given threshold M, with the fraction of space in which the evolved linear density field exceeds some threshold δ_{c}. It has been extensively tested with great success against N-body simulations in the context of primordial Gaussian fluctuations, allowing for the computation of the number density of clusters within a given background cosmology. The first attempt to extend the Press-Schechter formalism into the non-Gaussian regime was carried out by Colafrancesco, Lucchin and Matarrese [158]. Here we used the extension of the Press-Schechter formalism for non-Gaussian perturbations proposed by Chiu, Ostriker and Strauss [152]. It has been verified by Robinson et al. [155] for a particular set of non-Gaussian structure formation models. The fraction of the total mass within collapsed objects larger than a given mass M is given by:

$$\frac{\Omega_{\mathrm{m}}(> M(R,z), z)}{\Omega_{\mathrm{m}}(z)} = f_R \int_{\delta_*}^{\infty} P_R(\nu)d\nu, \tag{8.11}$$

where

$$M = \frac{4\pi R^3 \rho_{\mathrm{b}}}{3} \tag{8.12}$$

is the cluster mass, ρ_{b} is the background energy density of the universe,

$$f_R = \left[\int_0^\infty P_R(\nu) d\nu\right]^{-1}, \tag{8.13}$$

$$\delta_* = \frac{\delta_{\mathrm{c}}}{\sigma(R,z)}, \tag{8.14}$$

and $\delta_{\mathrm{c}} = 1.7 \pm 0.2$ (95% confidence interval) assuming spherical collapse [159].

To obtain the number density of clusters in a mass interval $\mathrm{d}M$ about M at a redshift z, we differentiate the Press-Schechter formula (8.11) to obtain (c.f. [129, 131])

$$n(M,z)\mathrm{d}M \approx -f_R \frac{\rho_{\mathrm{b}}}{M} \frac{\delta_{\mathrm{c}}}{\sigma^2(R,z)} \frac{\mathrm{d}\sigma(R,z)}{\mathrm{d}M} P_R\left[\frac{\delta_{\mathrm{c}}}{\sigma(R,z)}\right] \mathrm{d}M. \tag{8.15}$$

In the derivation of equation (8.15) we have ignored other terms resulting from the R dependence of P_R and f_R. This is a mathematically motivated approximation in the regime of mild non-Gaussianity, and we have verified in our case that it gives rise to at most a 2% error in the final estimate of σ_8. Substituting (8.6) into (8.15) gives

$$n(M,z)\mathrm{d}M = f_R \frac{\rho_{\mathrm{b}}}{M^2} \frac{\delta_{\mathrm{c}}}{3\sigma(R,z)} P_R\left[\frac{\delta_{\mathrm{c}}}{\sigma(R,z)}\right] \left\{\frac{0.76(R\Gamma)^{0.3}\log(R\Gamma)}{[(R\Gamma)^{0.3}+1.71]^2} + \gamma(R\Gamma)\right\} \mathrm{d}M. \tag{8.16}$$

Ref. [160, 161] constructed a merging history for dark matter halos based on the excursion set approach and obtained an analytic expression for the probability that a galaxy cluster with present virial mass M would have formed at some redshift z. The probability that a galaxy cluster with present virial mass M would have formed at a given red-shift z is given by:

$$p(z) = p(w(z)) \frac{\mathrm{d}w(z)}{\mathrm{d}z}, \tag{8.17}$$

where

$$p(w(z)) = 2w(z)(F^{-1}-1)\mathrm{erfc}\left(\frac{w(z)}{\sqrt{2}}\right) - \sqrt{\frac{2}{\pi}}(F^{-1}-2)\exp\left(-\frac{w^2(z)}{2}\right), \tag{8.18}$$

$$w(z) = \frac{\delta_{\mathrm{c}}(\sigma(M,0)/\sigma(M,z)-1)}{\sqrt{\sigma^2(FM,0)-\sigma^2(M,0)}}, \tag{8.19}$$

and $F = 0.75 \pm 0.15$ (95% confidence interval) is the fraction of the cluster mass assembled by a red-shift z [162]. We note that this result was derived in the context of primordial Gaussian fluctuations and must ultimately be verified using N-body simulations. Although it may not be valid for generic non-Gaussian models, we still expect this to be a good approximation in the context of the cosmic string scenario for structure formation as the departures from Gaussianity on clusters scales are relatively small (see Chapter 6).

8.4 The mass-temperature relation

In order to use the generalised Press-Schechter formalism to determine the abundance of X-ray clusters with a given temperature, we need to relate the X-ray temperature of a cluster to its virial mass. Here, we use the results in Ref. [129, 131] for the normalized virial mass-temperature relation (modified from [129]):

$$M_{\mathrm{v}} = (1.23 \pm 0.33) \times 10^{15} \left[\frac{\Omega_{\mathrm{mt}}^{b(\Omega_{\mathrm{mt}})}}{\Omega_{\mathrm{m0}}}\right]^{1/2} \left[\frac{1.67}{1+z_{\mathrm{t}}} \frac{2(2-\eta)(4-\eta)^2}{64-56\eta+24\eta^2-7\eta^3+\eta^4} \frac{k_{\mathrm{B}}T}{7.5\mathrm{keV}}\right]^{3/2} h^{-1}\mathrm{M}_\odot, \tag{8.20}$$

where z_{t} is the turnaround redshift, $\Omega_{\mathrm{mt}} \equiv \Omega_{\mathrm{m}}(z_{\mathrm{t}})$, and

$$\eta \equiv \eta(z_{\mathrm{t}}) = \frac{32}{9\pi^2} \frac{\Omega_{\Lambda 0}\Omega_{\mathrm{mt}}^{b(\Omega_{\mathrm{mt}})}}{\Omega_{\mathrm{m0}}(1+z_{\mathrm{t}})^3}, \tag{8.21}$$

$$b(\Omega) = \begin{cases} 0.76 - 0.25\Omega & (\mathrm{OCDM}), \\ 0.73 - 0.23\Omega & (\Lambda\mathrm{CDM}). \end{cases} \tag{8.22}$$

For a given red-shift of cluster collapse $\mathrm{z_c}$, the turnaround redshift $\mathrm{z_t}$ is easily obtained using the fact that $2t(\mathrm{z_t}) = t(\mathrm{z_c})$.

Hence we can now estimate the present comoving number density of galaxy clusters per temperature interval $\mathrm{d}(k_{\mathrm{B}}T)$ with a mean X-ray temperature $k_{\mathrm{B}}T$ which were formed at a given redshift z as:

$$n_T(k_{\mathrm{B}}T, z)\mathrm{d}(k_{\mathrm{B}}T)\mathrm{d}z = \frac{3M}{2k_{\mathrm{B}}T} n(M, 0)p(z)\mathrm{d}(k_{\mathrm{B}}T)\mathrm{d}z. \tag{8.23}$$

The present abundance of X-ray clusters with a temperature $k_{\mathrm{B}}T$ greater than 6.2keV can be estimated by integrating equation (8.23) from $z = z_{\mathrm{c}} = 0$ to infinity. A comparison between the observed cluster abundance and its theoretical prediction will give an estimate of σ_8.

We shall use the observation for the number density of galaxy clusters at $z = 0.05$ with X-ray temperature exceeding 6.2keV, given by Ref. [129], based on the dataset presented by Ref. [156], and updated by Ref. [157]:

$$N(> 6.2\mathrm{keV}, 0.05) = 1.53 \times 10^{-7 \pm 0.16} h^3 \mathrm{Mpc}^{-3}. \tag{8.24}$$

The uncertainty in (8.24) is the one-sigma interval, and these results have taken into account the effect of temperature measurement errors. The reasons for concentrating on galaxy clusters with temperature larger than 6.2keV has been extensively discussed by Viana & Liddle [129].

8.5 Results and discussion

By comparing (8.24) with the result integrated from (8.23), we obtain the observationally constrained σ_8 as plotted in Figure 8.1. The overall error in the value of σ_8 was estimated by Monte Carlo simulations over 10^4 realizations, treating the intrinsic uncertainties in Γ, $N(> 6.2\mathrm{keV}, 0.05)$ as lognormal, and those in δ_{c}, M_{v}, F as Gaussian. An accurate numerical fit to this result is

$$\sigma_8 = \begin{cases} 0.44\,\Omega_{\mathrm{m0}}{}^{-0.45+0.15\Omega_{\mathrm{m0}}} & (\mathcal{S}\mathrm{OCDM}), \\ 0.44\,\Omega_{\mathrm{m0}}{}^{-0.6+0.3\Omega_{\mathrm{m0}}} & (\mathcal{S}\Lambda\mathrm{CDM}). \end{cases} \tag{8.25}$$

The 95% confidence limits are $\pm 32\%$ in the $\mathcal{S}$OCDM case, and $+35\%$ and -32% in the $\mathcal{S}\Lambda$CDM case. We note that the overall shape of σ_8 in Figure 8.1 is similar to that obtained in the context of inflationary models with primordial Gaussian fluctuations [129, 131]. This is expected because both the string-induced and inflationary power spectra used in the calculation of cluster abundance are constrained by the same observations [33], and have roughly the same shape within the scales of interest (though with quite different choices of Γ). However, we also notice that the amplitude of σ_8 here is about 10–20% lower than that of inflationary perturbations. This is due to a slightly larger right-hand side tail of the PDF, which is obviously an intrinsic property in the cosmic string scenario. Therefore even using a further accurate power spectrum as the input of the calculation here, we expect the picture revealed above to be unchanged, i.e. when normalised to the cluster abundance, the string-induced σ_8 is smaller than that obtained from the standard CDM model.

In addition, we have verified that both the finite size and limited dynamic range of our numerical simulations do not result in significant uncertainties in our PDF. The chosen dynamic range around the matter-radiation transition includes the primary contributions to perturbations on cluster scales. The epochs missing from the simulations (both deep radiation and later matter eras) induce only weak perturbations which, in principle, should not affect the high density tail on the right-hand side of the PDF which is associated with clusters. The finite simulation size is again tailored to study cluster formation, with a grid resolution $0.25h^{-1}$Mpc well below cluster scales up to a box-size of $64h^{-1}$Mpc. Note that on the largest scales the PDF is very nearly Gaussian, so we do not expect to gain further information from larger simulations.

The value of σ_8 in the context of the cosmic string model for structure formation was also investigated by van de Bruck [163]. In his work he assumed the one point PDF to be independent of scale, by taking the distribution at the non-Gaussian scale ($\sim 1.5(\Omega_{\mathrm{m0}}h^2)^{-1}$Mpc) from our results in the previous chapter to be valid up to scales relevant for the cluster abundance calculation. Although this assumption can give the right qualitative results, for small values of $\Omega_{\mathrm{m0}}h$, the R dependence of the one point PDF needs to be taken into account in order to obtain more accurate results. This improvement has been incorporated in our work, which also took into consideration the merging history of dark matter halos.

To see the difference between the COBE normalized long-string induced σ_8 and the cluster-abundance constraint, we define a parameter

$$\mathcal{B}(\Omega_{\mathrm{m0}}, \Omega_{\Lambda 0}) = \frac{\sigma_8^{(\mathrm{cl})}(\Omega_{\mathrm{m0}}, \Omega_{\Lambda 0})}{\sigma_8^{(\mathrm{COBE})}(\Omega_{\mathrm{m0}}, \Omega_{\Lambda 0})}, \tag{8.26}$$

where $\sigma_8^{(\mathrm{cl})}$ is the cluster result (8.25), and $\sigma_8^{(\mathrm{COBE})}$ is inferred from the COBE normalised long-string induced power spectrum (6.21). In the calculation of $\sigma_8^{(\mathrm{COBE})}$, we have used $\Omega_{\mathrm{B0}} = 0.05$, $\Omega_{\mathrm{m0}}h = 0.14$, and $G\mu_6 =$

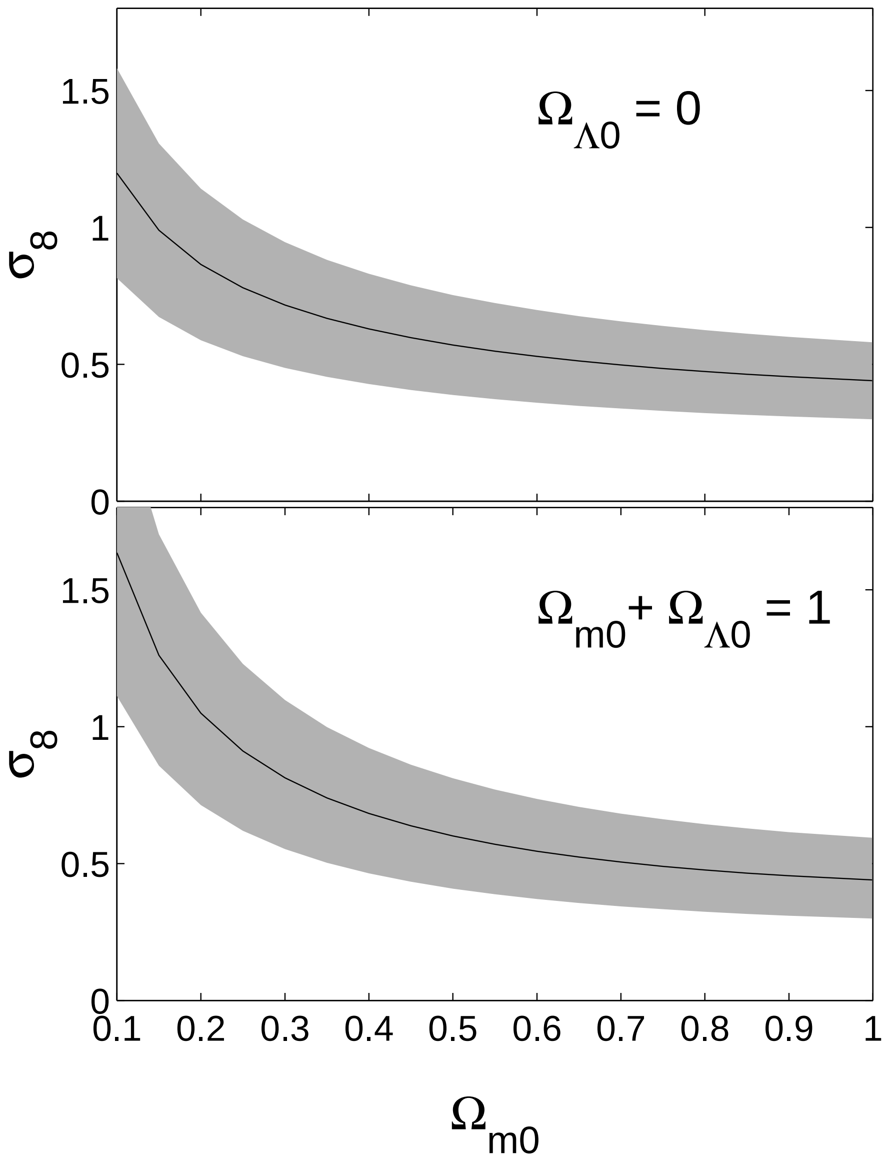

Figure 8.1: The σ_8 in cosmic string scenarios normalized to cluster abundance. The upper panel is the $\mathcal{S}$OCDM models, while the lower panel is the $\mathcal{S}\Lambda$CDM models.

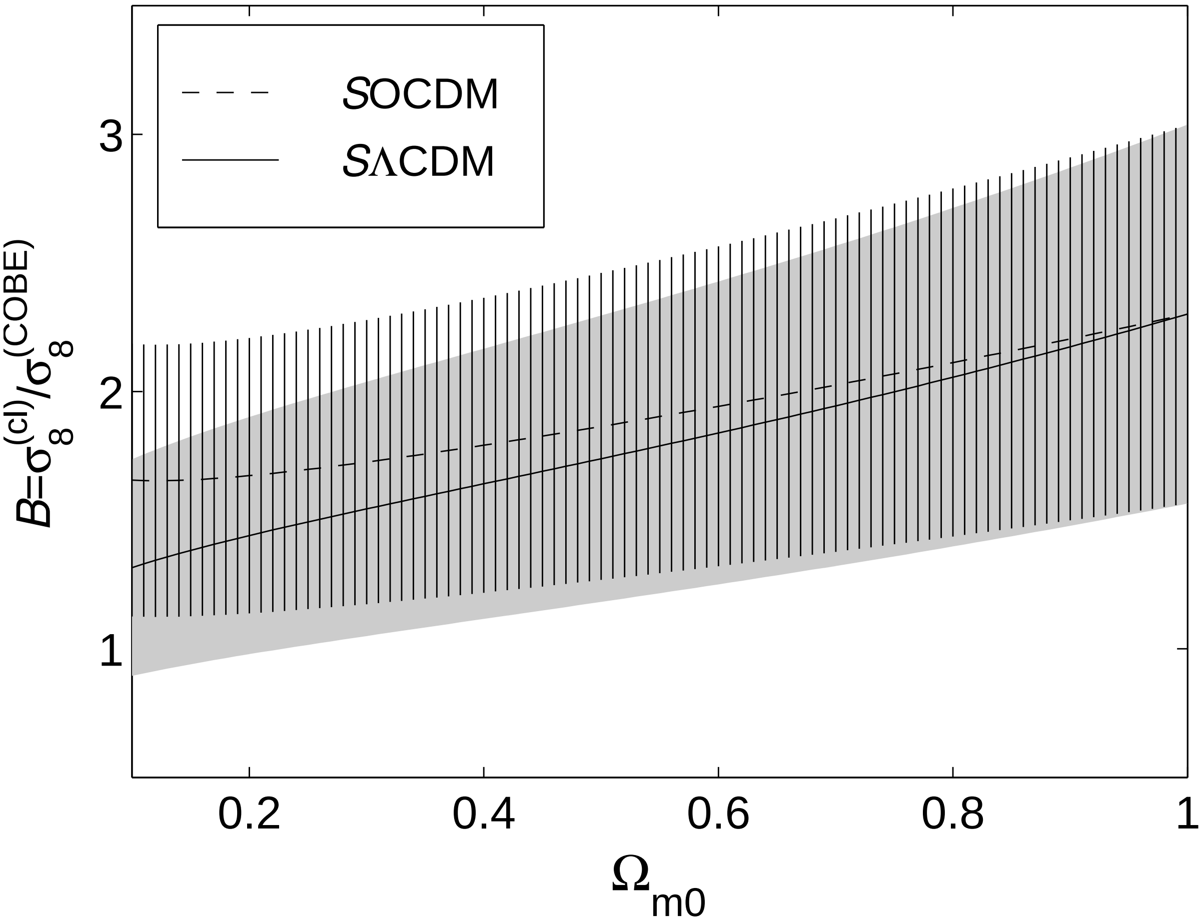

Figure 8.2: $\mathcal{B}$, the ratio of the σ_8 normalized by cluster abundance and by COBE, is plotted as a function of $\Omega_{\rm m0}$ and $\Omega_{\Lambda 0}$. The hatched ($\mathcal{S}$OCDM) and shaded ($\mathcal{S}\Lambda$CDM) areas are within a 95% confidence level.

$G\mu \times 10^6 = 1.7$ [119] for $\Omega_{\rm m0} = 1$ and $\Lambda = 0$, with the open and Λ-parameter dependence for $G\mu_6$ given by $G\mu_6 \propto {\Omega_{\rm m0}}^{-0.3}$ for $\mathcal{S}$OCDM and $G\mu_6 \propto {\Omega_{\rm m0}}^{-0.05}$ for $\mathcal{S}\Lambda$CDM [118].

We can observe in Figure 8.2 for the flat Λ-model that $\mathcal{B}(\Omega_{\rm m0}, \Omega_{\Lambda 0})$ ranges from 1.3 ± 0.5 to 2.3 ± 0.8 (95% confidence level) for different choices of $\Omega_{\rm m0}$ and $\Omega_{\Lambda 0}$. For small $\Omega_{\rm m0} \approx 0.1$–$0.2$, $\mathcal{B}$ is approximately unity, meaning that $\sigma_8^{\rm (cl)}$ and $\sigma_8^{\rm (COBE)}$ seem to be consistent. However, for large $\Omega_{\rm m0}$, $\mathcal{B}$ is higher than unity. Nevertheless, as shown in Chapter 6, the inclusion of perturbations induced by cosmic string loops can boost the string-induced power spectrum by a factor of about two under reasonable assumptions. This may help remove the small discrepancy between the COBE and cluster-abundance constraints on σ_8, which we see in Figure 8.2 even for high values of $\Omega_{\rm m0}$. We notice that the consistency between $\sigma_8^{\rm (cl)}$ and $\sigma_8^{\rm (COBE)}$ would also imply the consistency between the $G\mu$ normalised by cluster abundance and COBE, because σ_8 is proportional to the string linear energy density $G\mu$. We also note that the exact contribution of the background of gravitational radiation emitted by cosmic string loops remains a significant uncertainty [127].

8.6 Conclusion

We have constrained the amplitude of matter density perturbations induced by cosmic strings on the scale of $8h^{-1}$Mpc in both open cosmologies and flat models with a non-zero cosmological constant, using the currently observed number density of large X-ray clusters. Because string-seeded matter perturbations intrinsically have a slightly larger right-hand side tail in the PDF resulting in the mild non-Gaussianity on cluster scales, we obtained a slightly lower normalization of σ_8 than that found for cosmological models with primordial Gaussian fluctuations. We compared this σ_8 with that constrained by COBE and found that they are consistent when current uncertainties in the normalization of the power spectrum are taken into account. This also implies the consistency between the COBE normalised string linear energy density $G\mu$ and that normalised by cluster abundance.

Part IV

COSMIC MICROWAVE BACKGROUND

Chapter 9

Cosmic Microwave Background by Cosmic Strings

9.1 Introduction

Having investigated the LSS with cosmic strings in Part III, we now study another important aspect of cosmological perturbations seeded by cosmic strings—the CMB. As mentioned in Chapter 1, the CMB anisotropies provide us with currently the cleanest and deepest probe of the evolution of cosmological perturbations, on the largest scales accessible by observations. Thus the recent experiments of CMB anisotropy measurements [164] are clearly an advance in understanding the universe around us. Since different models for structure formation predict cosmological perturbations with different properties, we expect that a detailed study of CMB for these models will help advance our knowledge regarding the current status of the universe.

To this end, we shall first investigate the CMB anisotropies from cosmic strings in this chapter, and then propose two different methods in the next two chapters for distinguishing Gaussian models like the standard CDM model from non-Gaussian ones such as defects (e.g. [165, 166, 167]). In section 9.2, we review the various components in the analysis of CMB anisotropies, including two main categories: the primary perturbations and the integrated Sachs-Wolfe (ISW) effect. Section 9.3 briefly derives the small-field-limit (SFL) approximation, which is widely applied to CMB experiments [168]. Within the SFL, we then compute the string-seeded primary perturbations in section 9.4, and the string-induced ISW effect in section 9.5. Finally we give a conclusion in section 9.6.

9.2 Various components in the CMB

Before the decoupling epoch $\eta_{\rm d}$, the temperature of the universe is high enough to ionise hydrogen. Compton scattering tightly couples the photons to the electrons which are in turn coupled to the baryons by electromagnetic interactions. The dynamics of the system can thus be approximated as a photon-baryon fluid as we described in Chapter 3. Photon pressure resists gravitational compression of the fluid and sets up acoustic oscillations. At the decoupling $\eta_{\rm d}$, baryons recombine to form neutral hydrogen and the photons scatter for the last time. After the epoch $\eta_{\rm d}$, photons are free to propagate but there can be further gravitational effects on their way to the observer. Eventually, the resulting fluctuations appear to the observer today as CMB anisotropies on the sky. We shall refer to the fluctuations originating from the epoch $\eta_{\rm d}$ as 'primary anisotropies', while those further induced on the photon path to the observer following $\eta_{\rm d}$ will be called 'secondary anisotropies' [169]. On large scales, the physics of CMB anisotropies can be described by the geodesic equation in linearised gravity and was first calculated by Sachs and Wolfe in synchronous gauge perturbation theory [170]. For a perturbed FRW model ($g_{\mu\nu} = a^2(\eta)(\eta_{\mu\nu} + h_{\mu\nu})$), the general result can be written as

$$\frac{\Delta T}{T}(\hat{\mathbf{n}}) = \left[\frac{\Delta T}{T}\right]_{\rm i} + \frac{1}{2}\left[h_{00}\right]_{\rm i}^{\rm f} + \hat{\mathbf{n}} \cdot \left[\mathbf{v}_{\gamma}\right]_{\rm i}^{\rm f} - \frac{1}{2}\int_{\eta_{\rm i}}^{\eta_{\rm f}} d\eta\, \dot{h}_{\mu\nu}\left[x^{\alpha}_{(0)}(\eta)\right] n^{\mu} n^{\nu}, \tag{9.1}$$

where the subscript 'i' denotes the emission time of photons, 'f' denotes the observing time, the unit vector $\hat{\mathbf{n}}$ gives the direction in which one is measuring the temperature, $\mathbf{v}$ is the velocity with respect to the cosmic rest frame at a given time ($\eta_{\rm i}$ or $\eta_{\rm f}$), a dot represents the derivative with respect to the conformal time η, and $x^{\alpha}_{(0)}(\eta) = n^{\alpha}\eta$ is the unperturbed photon geodesic, with $n^{\alpha} = (1, -\hat{\mathbf{n}})$. These four terms have different physical interpretations: the first term gives the fractional deviation from the temperature anisotropy at the point of photon emission (or last scattering); the second term accounts for the gravitational potential redshift; the third term is the Doppler effect resulting from the local motion of the photon fluid; and the last term is an integral along the photon trajectory of the time rate of change of the gravitational field. This last term is normally referred to as the 'integrated Sache-Wolfe (ISW) effect'.

In the instantaneous decoupling approximation as illustrated in Chapter 3, photons are assumed to be tightly coupled to baryons before η_d, and suddenly become free to propagate after this epoch. Thus in the synchronous gauge and assuming no reionization after η_d, equation (9.1) can be simplified to describe the CMB anisotropies observed today:

$$\frac{\Delta T}{T}(\hat{\mathbf{n}}) = \left.\frac{\Delta T}{T}\right|_\mathrm{prim}(\hat{\mathbf{n}}) + \left.\frac{\Delta T}{T}\right|_\mathrm{ISW}(\hat{\mathbf{n}}), \tag{9.2}$$

where

$$\left.\frac{\Delta T}{T}\right|_\mathrm{prim}(\hat{\mathbf{n}}) = \frac{1}{4}\delta_\gamma(\hat{\mathbf{n}};\eta_\mathrm{d}) - \hat{\mathbf{n}}\cdot\mathbf{v}_\gamma(\eta_\mathrm{d}), \tag{9.3}$$

$$\left.\frac{\Delta T}{T}\right|_\mathrm{ISW}(\hat{\mathbf{n}}) = -\frac{1}{2}\int_{\eta_\mathrm{d}}^{\eta_0} d\eta \dot{h}_{ij}(n^\alpha\eta)n^i n^j. \tag{9.4}$$

Here we have used the fact that $\Delta T/T = \delta_\gamma/4$ (because $\rho_\gamma \propto a^{-4} \propto T^4$) for the first term in equation (9.3), and we have assumed that the present observer is in the cosmic rest frame. Equation (9.3) gives the primary anisotropies mentioned earlier, while equation (9.4) gives the secondary anisotropies, or equivalently the ISW effect. For the primary anisotropies (9.3), the first term is simply the photon density contrast at the decoupling epoch, which we shall refer to as the 'intrinsic temperature contrast'; the second term is the 'Doppler effect' originating from the velocity of the photon-baryon coupled fluid. These two terms are like a snapshot at the last scattering surface, and can be computed by evolving the density perturbations from the early universe to the epoch of decoupling η_d. On the other hand, the ISW effect (9.4) accounts for the further gravitational effects on the past light cone between the last scattering surface and the present observer. Thus the integral needs to be carried out from the epoch η_d to the present time η_0, with the time derivative of the spatial perturbations $\dot{h}_{ij}$ determined at each moment. We shall compute the string-seeded primary perturbations (9.3) in section 9.4 and the string-induced ISW effect (9.4) in section 9.5.

9.3 Small field approximation

One fundamental test of theories using the observed CMB anisotropies is to compare their angular power spectra. In the calculation of the CMB spectrum, one should in principle carry out the multipole transform (MT; the spherical harmonic decomposition) as demonstrated in section 1.5.2 (see also Figure 1.1). However, if we consider only a small patch in the CMB sky which is nearly flat, the MT can be approximated by a Fourier Transform (e.g. [168, 171]). This is called the small-field limit (SFL) or sometimes the small-angle limit. In the language of the MT, this limit implies the multipole index $l \gg 1$. Since the Fast Fourier Transform (FFT) algorithm has been well studied (e.g. [172]), the SFL approximation enhances the efficiency of calculating the CMB spectrum and its inverse. We shall now briefly review this approximation.

Within the SFL, consider a small patch in the CMB sky and place the spherical coordinates $\hat{\mathbf{n}} = (\theta,\phi) = (0,0)$ at its centre. The (θ,ϕ) coordinates can then be approximated with polar coordinates $\mathbf{y} = (r,\phi)$ on the patch, where $r = 2\sin(\theta/2) \approx \theta$. Similarly, we replace the indices (l,m) in the multipole space with $\mathbf{l} = (l,\phi_l)$, and define [168]

$$a_\mathbf{l} \equiv \sqrt{\frac{4\pi}{2l+1}}\sum_m i^{-m} a_{lm} e^{im\varphi_l}, \tag{9.5}$$

where we notice $a_{-\mathbf{l}} = a_\mathbf{l}^*$, and '$*$' denotes complex conjugation. Using this definition, we expand the temperature in terms of the multipole moments as

$$\frac{\Delta T}{T}(\hat{\mathbf{n}}) = \sum_{lm} a_{lm}Y_{lm}(\hat{\mathbf{n}}) = \frac{1}{2\pi}\int dl\, d\varphi_l\, a_\mathbf{l}\sqrt{\frac{2l+1}{4\pi}}\sum_m i^m e^{-im\varphi_l}Y_{lm}(\hat{\mathbf{n}}) \tag{9.6}$$

$$\approx \frac{1}{(2\pi)^2}\int l dl\, d\varphi_l\, a_\mathbf{l}\left[J_0(l\theta) + 2\sum_{m=1}^{\infty} i^m J_m(l\theta)\cos m(\phi-\varphi_l)\right], \tag{9.7}$$

where we have replaced the sum over l with an integral, and the second step is achieved by writing the sum over m in terms of $m<0$ and replacing the associated Legendre polynomials with Bessel functions (equation

8722(2) in Ref. [173]). The term with square brackets is the Rayleigh expansion of a plane wave, so we finally obtain

$$\frac{\Delta T}{T}(\hat{\mathbf{n}}) \equiv \frac{\Delta T}{T}(\mathbf{y}) = \sum_{lm} a_{lm} Y_{lm}(\mathbf{y}) \approx \frac{1}{(2\pi)^2} \int d^2 l \; a_{\mathbf{l}} \, e^{i\mathbf{l}\cdot\mathbf{y}}, \tag{9.8}$$

and

$$\langle a^*_{\mathbf{u}} a_{\mathbf{w}} \rangle = C_l \delta^{(2)}(\mathbf{u} - \mathbf{w}). \tag{9.9}$$

As one can see, equation (9.8) is simply a Fourier transform. Thus for a small patch of angular area A, the CMB spectrum can be obtained as

$$l(l+1)C_l \approx \frac{k^2}{A} \langle |a_{\mathbf{k}}|^2 \rangle, \tag{9.10}$$

where

$$a_{\mathbf{k}} = \int dx^2 \frac{\Delta T}{T}(\mathbf{y}) e^{-i\mathbf{k}\cdot\mathbf{x}}. \tag{9.11}$$

This SFL approximation is convenient not only for computing the C_l curve from small-field data in the CMB sky, but also for simulating small-angular Gaussian maps from a underlying C_l curve [171].

9.4 Primary perturbations

The primary perturbations (9.3) consists of two terms: the intrinsic temperature contrast and the Doppler effect. To compute them for a full-sky map, in principle one needs to evolve the density perturbations from well before radiation-matter equality to $\eta_{\rm d}$, in a box of the size equal to the comoving distance that photons can travel since last scattering (i.e. of the size $\eta_0 - \eta_{\rm d} \approx \eta_0$). One then takes a snapshot of δ_γ and $\mathbf{v}_\gamma$ at $\eta_{\rm d}$ to obtain the primary anisotropies using equation (9.3). However, due to the limitation of computational resources, we shall work in the SFL, with a finite dynamic range starting from $0.75\eta_{\rm eq}$, and ending at $\eta_{\rm d}$.

To evolve the density perturbations with a cosmic string network, we use the Fourier transformed version of equations (3.54)–(3.56), with initial conditions

$$\widetilde{s} = \dot{\widetilde{s}} = \widetilde{\tau}_{00} = \widetilde{\delta}_{\rm c} = 0. \tag{9.12}$$

The condition $\widetilde{\delta}_{\rm c} = 0$ here is employed for the sake of numerical convenience. Although this is not physically motivated, we have analytically proved in section 4.5.3 and numerically verified here that this choice does not affect the final density perturbations. The cosmological parameters we choose are $\Omega_{\rm c0} = 0.95$, $\Omega_{\rm B0} = 0.05$, $\Omega_\Lambda = 0$ and $h = 0.7$, resulting in $\eta_{\rm d} \approx 6\eta_{\rm eq}$ and $\eta_0 \approx 255\eta_{\rm eq}$ in accordance with equations (3.10), (3.13) and (3.61). We start the simulation from $0.75\eta_{\rm eq}$, when the string network is already in the scaling regime, and when the horizon size is sixteen times smaller than the box size. This comoving box corresponds to an angular size of about 1.34 degrees today. The resolution of the box is 128^3. At the epoch $\eta_{\rm d}$, we use equations (3.38) and (3.39) to calculate $\widetilde{\delta}_\gamma(\mathbf{k})$ and $\widetilde{\mathbf{v}}_\gamma(\mathbf{k})$ respectively, and then inverse Fourier transform them back to obtain three-dimensional information. We then take a slice from the three-dimensional box and use equation (9.10) to compute its angular power spectrum.

Figure 9.1 shows preliminary results for intrinsic temperature contrast $\delta_\gamma(\eta_{\rm d})/4$ (the first term in equation (9.3)) in a string-induced CMB map, while Figure 9.2 shows the Doppler effect $-\hat{\mathbf{n}} \cdot \mathbf{v}_\gamma(\eta_{\rm d})$ (the second term in equation (9.3)) of the same realisation. The sum of the two is presented in Figure 9.3, which shows the total primary anisotropies $(\Delta T/T)|_{\rm prim}$. In each plot, the colour scale shows the temperature contrast $\Delta T/T$ in units of one standard deviation. With the choice of $G\mu_6 = 1 \times 10^6 = 1$, the standard deviations of these preliminary maps are $\sigma_{\rm intr} \approx 2.4 \times 10^{-5}$, $\sigma_{\rm Dopp} \approx 3.1 \times 10^{-5}$, and $\sigma_{\rm prim} \approx 3.8 \times 10^{-5}$ respectively. We note that the amplitude of the Doppler effect is stronger than that of the intrinsic temperature contrast, i.e. $\sigma_{\rm Dopp} > \sigma_{\rm intr}$. In addition, all maps show hot and cold regions which are beyond the three-sigma range of the temperature contrast, although they look very smooth due to photon diffusion damping. These hot and cold spots can depend strongly on the resolution of the string simulation [134], and this has to be studied in more detail before firm conclusions can be drawn.

Using equation (9.10), we can also compute the angular power spectra $l(l+1)C_l/2\pi$ of these maps. They are shown in Figure 9.4. As one can see, the angular power spectrum of the intrinsic temperature contrast peaks at $l \sim 1200$, while that of the Doppler effect reaches its maximum at $600 \lesssim l \lesssim 900$. In the comparison of these two curves, we see that they peak at different angular scales, and this is due to the different oscillation phases in the density contrast and the velocity of the photon-baryon fluid. We also note that the intrinsic temperature contrast has a lower amplitude at low l, reflecting the smaller standard deviation of the temperature contrast computed earlier. The sum of these two contributions result in the total primary perturbations, whose power

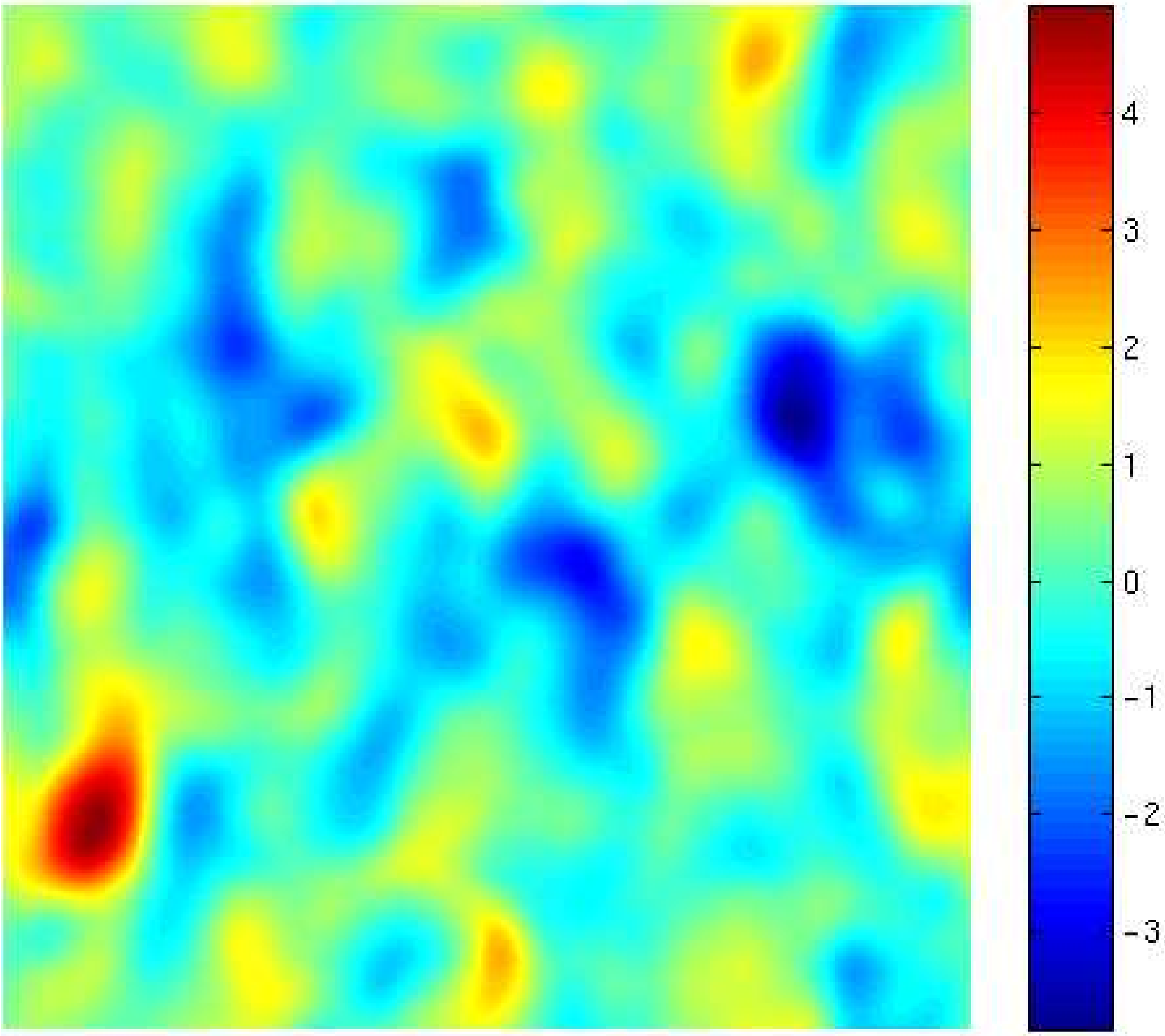

Figure 9.1: The intrinsic temperature contrast in the string-induced primary CMB anisotropies. The angular size is $(1.34°)^2$, and the resolution is 128^3. We have used $\Omega_{c0} = 0.95$, $\Omega_{B0} = 0.05$, $\Omega_\Lambda = 0$, $h = 0.7$ and $G\mu_6 = 1$. The dynamic range is from $0.75\eta_{eq}$ to $\eta_d \approx 6\eta_{eq}$. The colour scale shows the temperature contrast $\Delta T/T$ in units of one standard deviation. The standard deviation is $\sigma_{intr} \approx 2.4 \times 10^{-5}$.

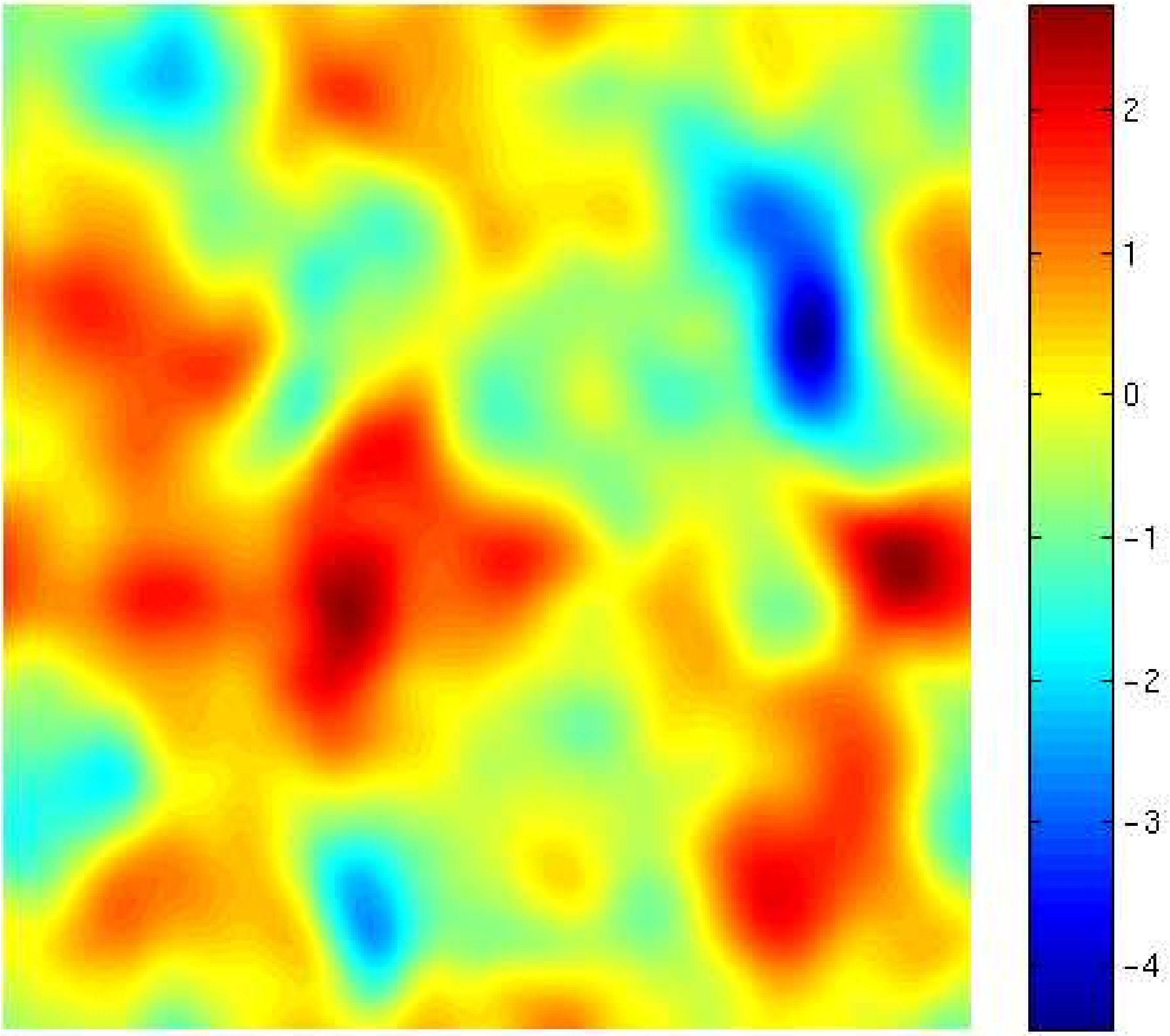

Figure 9.2: The Doppler effect in the string-induced primary CMB anisotropies. The scale and other parameters are the same as Figure 9.1, except that the standard deviation here is $\sigma_{\rm Dopp} \approx 3.1 \times 10^{-5}$.

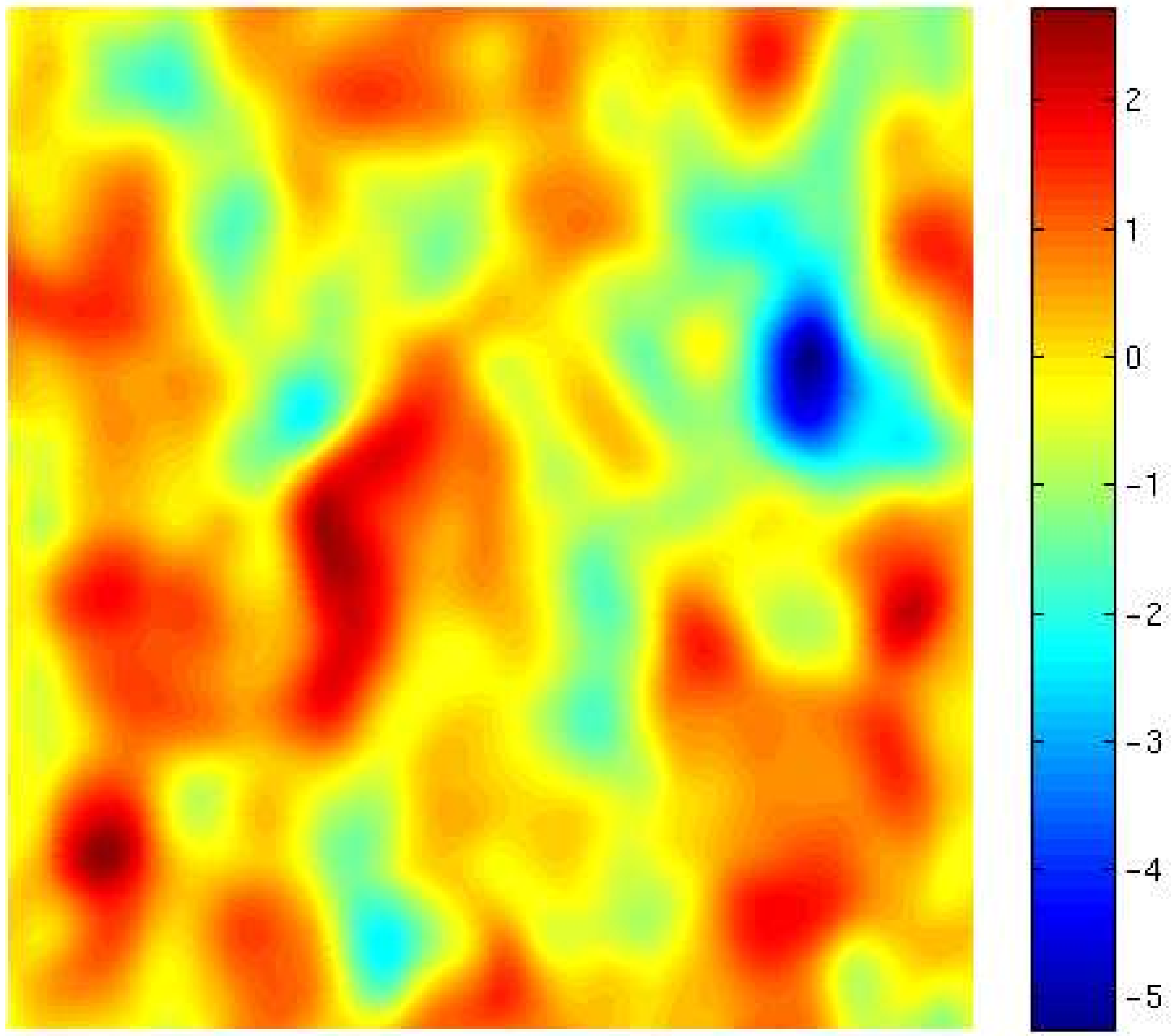

Figure 9.3: The total string-induced primary CMB anisotropies from the last scattering surface. This is the linear sum of Figures 9.1 and 9.2. The standard deviation here is $\sigma_{\rm prim} \approx 3.8 \times 10^{-5}$.

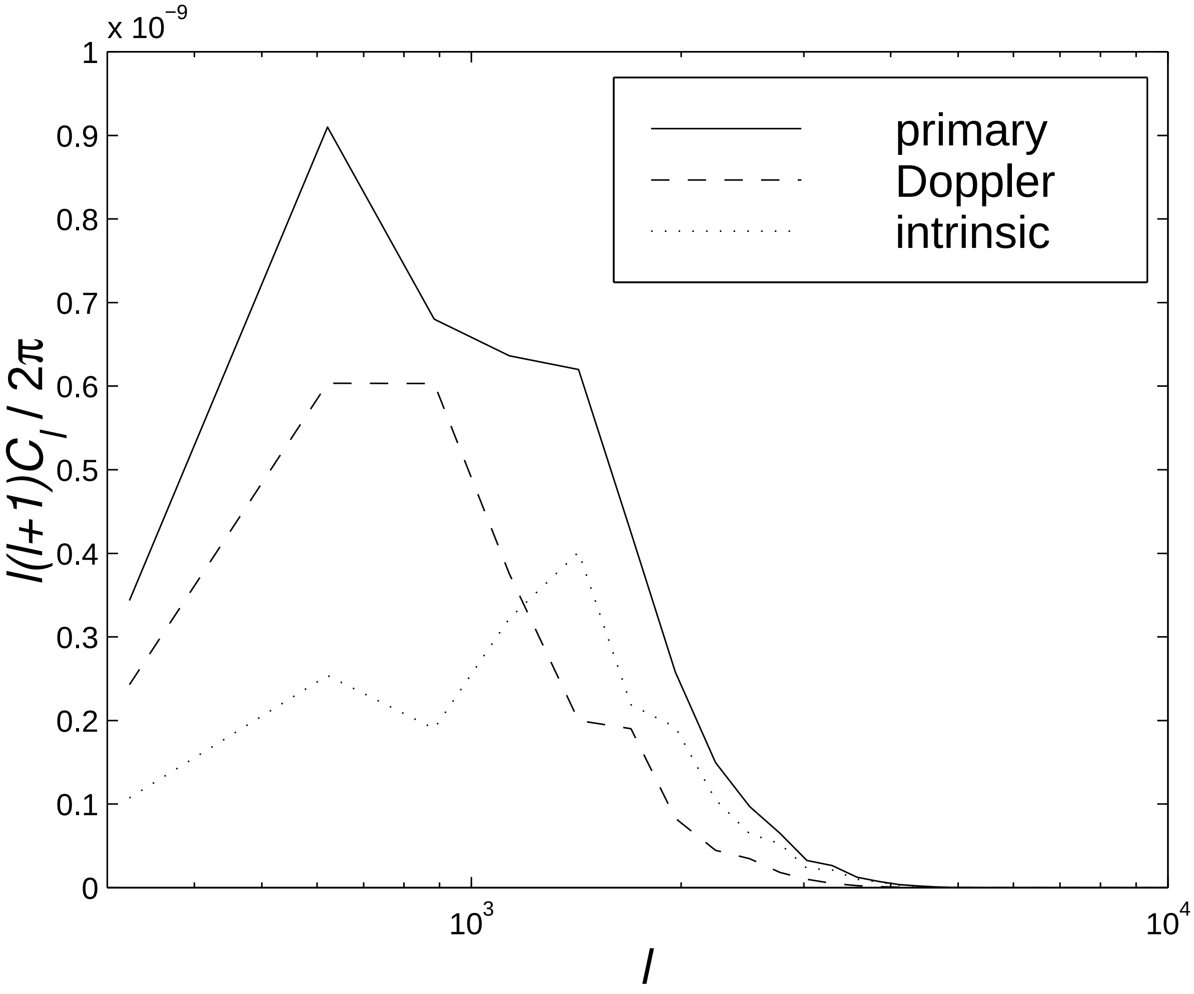

Figure 9.4: Preliminary results for the angular power spectra of the intrinsic temperature contrast (dotted), the Doppler effect (dashed), and their sum (the primary anisotropies; solid). We have used $\Omega_{c0} = 0.95$, $\Omega_{B0} = 0.05$, $\Omega_\Lambda = 0$, $h = 0.7$, and $G\mu_6 = 1$.

spectrum is also shown in Figure 9.4. This curve has a very broad peak, reaching its maximum at around $l \sim 600$, and starts to drop dramatically at $l \gtrsim 1400$.

We now comment on the uncertainties and potential problems in the current calculations. First, when the simulation started, the horizon size was only sixteen times smaller than the box size. This is not adequate to set $\widetilde{\tau}_{00} = 0$ on all scales of interest. Second, the byproducts of the long-string network, like small loops and gravitational waves, have not been included. These components could potentially not only affect the superhorizon compensation, but also change the subhorizon behaviour of the CMB power spectrum. Due to these reasons, indeed we saw that throughout the simulation, the power spectrum of $\widetilde{\tau}_{00}(k)$ decay as k^2 on superhorizon scales rather than k^4 as it should be, indicating that the energy was not well conserved. Thirdly, the initial time of the current simulation is only $0.75\eta_{\rm eq}$, which is not well before the radiation-matter transition. This will result in an inadequate power in the final CMB angular power spectrum due to the lost dynamic range in the radiation era. Therefore, in the future work, we shall need a simulation with a larger dynamic range to overcome the power lost here, with an initial horizon small enough in order to apply the initial condition $\widetilde{\tau}_{00} = 0$, and with all small loops and the gravitational waves that they produce properly included so as to conserve the energy more accurately on large scales. The uncertainties in this preliminary work remain too large for us to draw any robust conclusions.

As mentioned in Chapter 2, one of the main goals of computing the CMB angular power spectra for different models is to identify the so-called Doppler peaks. These are oscillatory peaks in the C_l curve, which are evidence of phase coherence in the evolution of perturbations [55, 174]. Such phase coherence would indicate that perturbations are largely in the growing mode, meaning that they were already in place while well before the last scattering. This is the characteristic of inflationary perturbations, but unlikely to be present in isocurvature models like topological defects [56]. One of our aims in the future work will be to verify whether or not these peaks can exist in cosmic string models. This goal can not be achieved unless we considerably improve the uncertainties which are carried in the current work presented above.

9.5 Integrated Sachs-Wolfe effect

9.5.1 The SFL flat-space formalism

It is a great challenge to compute the ISW effect, because it is a path-dependent integral and in principle one needs to determine, in all directions, the time derivative of the spatial perturbations $\dot{h}_{ij}$ on the whole past light cone between the present observer and the last scattering surface. Nevertheless, since we are working in the SFL at present, the ISW effect (9.4) can be greatly simplified as shown in the literature [175, 176, 177, 178, 179]. Under the SFL in flat space, Ref. [175] showed that the ISW effect can be re-expressed as an integral over the stress-energy distribution which generates the gravitational fields. Ref. [176] re-examined this formalism with corrections, in a further simplified manner which also applies to gravitational lensing. Extensions to large angles for flat space and for a matter-dominated FRW model are given in Ref. [177] and Ref. [178] respectively. Here we use the SFL in flat space [176, 177]. We note that this formalism has also been employed to study texture-induced CMB anisotropies in Ref. [179].

In the SFL, the direction along the line of sight $\hat{\mathbf{n}}$ can be treated as a constant, so the coordinates of any position on the flat patch can be decomposed as $\mathbf{x}_\gamma(\eta) = \mathbf{x}_\perp + \hat{\mathbf{n}}(\eta_0 - \eta)$, where the subscript '$\perp$' denotes the component perpendicular to $\hat{\mathbf{n}}$ (i.e. the projection on the patch), and the time dependence accounts for the motion of the patch on the past light cone. We shall call the plane where the vector $\mathbf{x}_\perp$ is the '$\hat{\mathbf{n}}_\perp$ plane'. With this convention, the CMB anisotropies can be calculated as [176, 177]

$$\nabla_\perp^2 \frac{\Delta T}{T}\bigg|_{\rm ISW} (\mathbf{x}_\perp) = -8\pi G \nabla_\perp \cdot \mathbf{U}(\mathbf{x}_\perp)\,, \tag{9.13}$$

where $\mathbf{U}$ is a vector on the $\hat{\mathbf{n}}_\perp$ plane:

$$U_i(\mathbf{x}_\perp) = -(\delta_i^j - \hat{n}_i\hat{n}^j)\int_{-\infty}^{\infty} \left\{\Theta_{0j}[\eta, \mathbf{x}_\gamma(\eta)] - \hat{n}^k\Theta_{jk}[\eta, \mathbf{x}_\gamma(\eta)]\right\} d\eta\,, \quad i = 1,2,3\,, \tag{9.14}$$

and $\Theta_{\mu\nu}$ is the stress-energy tensor of the source. Note that equation (9.13) determines the temperature pattern up to an arbitrary function satisfying $\nabla_\perp^2\psi$. Nevertheless, because we shall use the periodic boundary condition, the fact that $\langle \Delta T/T\rangle = 0$ automatically determines ψ. If we choose $\hat{\mathbf{n}} = (0,0,1)$ and consider cosmic strings as the source, then equations (9.14) become

$$U_i(\mathbf{x}_\perp) = -(1 - \hat{n}_i\hat{n}^i)\int_{-\infty}^{\infty} \left\{\Theta_{0i}[\eta, \mathbf{x}_\gamma(\eta)] - \Theta_{i3}[\eta, \mathbf{x}_\gamma(\eta)]\right\} d\eta\,, \quad i = 1,2,3\,, \tag{9.15}$$

leading to

$$U_1(\mathbf{x}_\perp) = -\int_{-\infty}^{\infty} d\eta\mu \int d\sigma\, \delta^{(3)}[\mathbf{x}_\gamma(\eta) - \mathbf{x}_\mathrm{s}(\sigma,\eta)][\epsilon v_1(1-v_3) + \epsilon^{-1} {x'_\mathrm{s}}^1 {x'_\mathrm{s}}^3], \tag{9.16}$$

$$U_2(\mathbf{x}_\perp) = -\int_{-\infty}^{\infty} d\eta\mu \int d\sigma\, \delta^{(3)}[\mathbf{x}_\gamma(\eta) - \mathbf{x}_\mathrm{s}(\sigma,\eta)][\epsilon v_2(1-v_3) + \epsilon^{-1} {x'_\mathrm{s}}^2 {x'_\mathrm{s}}^3], \tag{9.17}$$

$$U_3(\mathbf{x}_\perp) = 0. \tag{9.18}$$

according to equation (2.93). Considering a two-dimensional Fourier transform on the $\hat{\mathbf{n}}_\perp$ plane, equation (9.13) further yields

$$\left.\widetilde{\frac{\Delta T}{T}}\right|_{\mathrm{ISW}}(\mathbf{k}) = \frac{8\pi G i}{k^2}\left[k_1 \widetilde{U}_1(\mathbf{k}) + k_2 \widetilde{U}_2(\mathbf{k})\right], \tag{9.19}$$

which implies

$$\left[\left.\widetilde{\frac{\Delta T}{T}}\right|_{\mathrm{ISW}}(\mathbf{k})\right]_{\mathcal{R}} = -\frac{8\pi G}{k^2}[k_1 \widetilde{U}_{1(\mathcal{I})}(\mathbf{k}) + k_2 \widetilde{U}_{2(\mathcal{I})}(\mathbf{k})], \tag{9.20}$$

$$\left[\left.\widetilde{\frac{\Delta T}{T}}\right|_{\mathrm{ISW}}(\mathbf{k})\right]_{\mathcal{I}} = \frac{8\pi G}{k^2}[k_1 \widetilde{U}_{1(\mathcal{R})}(\mathbf{k}) + k_2 \widetilde{U}_{2(\mathcal{R})}(\mathbf{k})], \tag{9.21}$$

where $\mathbf{k} = (k_1, k_2)$, $k = |\mathbf{k}|$, and the subscripts '$\mathcal{R}$' and '$\mathcal{I}$' denote the real and imaginary parts respectively. The results for $\hat{\mathbf{n}} = (1,0,0)$ or $(0,1,0)$ are similar.

9.5.2 Kaiser-Stebbins effect

As illustrated in section 2.3.4, in the rest frame of a cosmic string there exists a deficit angle when lights pass through it. This implies that if a string is perpendicular to the line of sight in the CMB sky and traversing through it, the light coming from the last scattering surface but just behind the string will experience a blueshift, while the light in front of the string will not. This mechanism will result in a discontinuity in the CMB temperature, and is called the Kaiser-Stebbins effect [165].

As an example of using the formalism described in the previous section, we compute the Kaiser-Stebbins effect. Figure 9.5 shows the CMB anisotropies induced by a wiggly string on the $\hat{\mathbf{n}}_\perp$ plane moving to the right-hand side. We can see clearly that the temperature behind the string is blueshifted, so that when the fluctuations are normalised to their mean, we gain a redshift in front of the string and a blueshift behind it. As we shall see, the Kaiser-Stebbins effect is the dominant feature of the ISW effect induced by cosmic strings. For other defect models, non-Gaussian mechanisms can also operate. For example, cosmic textures can lead to both hot and cold spots in the CMB temperature pattern [167].

9.5.3 String-induced ISW effect

We now use the formalism described in section 9.5.1 to compute the string-induced ISW effect. The numerical scheme we shall use is the same as that presented in Ref. [180]. First, we evolve the cosmic string network in the scaling regime from the decoupling epoch η_d through to deep into the matter-dominated era. The string box has a fixed comoving size with periodic boundary conditions. The two-dimensional vector $\mathbf{U}(\mathbf{x}_\perp)$ is then computed by integrating the source history along the past light cone, using equations (9.16)–(9.18) for $\hat{\mathbf{n}} = (0,0,1)$, and similarly for $\hat{\mathbf{n}} = (1,0,0)$ and $(0,1,0)$. This vector field $\mathbf{U}(\mathbf{x}_\perp)$ is then Fourier transformed on the $\hat{\mathbf{n}}$ plane, and plugged into equation (9.19). A further inverse Fourier transform will yield the ISW effect that we are aiming to compute.

The cosmological parameters we choose are $\Omega_{\mathrm{c}0} = 0.95$, $\Omega_{\mathrm{B}0} = 0.05$, $\Omega_\Lambda = 0$ and $h = 0.7$, resulting in $\eta_\mathrm{d} \approx 6\eta_\mathrm{eq}$ and $\eta_0 \approx 255\eta_\mathrm{eq}$ The dynamic range in conformal time is 5, starting from $\eta_\mathrm{d} \approx 6\eta_\mathrm{eq}$. This is equivalent to a redshift range of $z \approx 1300$–70. The box size of the simulation is $(24\eta_\mathrm{eq})^2$, with a resolution of 256^3. If we locate the centre of the box at $\eta = 18\eta_\mathrm{eq}$ on the past light cone, with one of the major axes aligning with the line of sight, then the farther side of the box will coincide with the last scattering surface. It follows that a plane in the box moving on the past light cone will impact the centre of the box at $\eta = 18\eta_\mathrm{eq}$ and subtend an angular size of about $(5.8^\circ)^2$ at that moment. We shall take this as the angular size of the resulting CMB maps. Figure 9.6 shows one of the maps. Figure 9.7 shows the string paths on the past light cone of the same realisation. As one can see, the Kaiser-Stebbins effect in Figure 9.6 is apparent. When compared with Figures 9.1, 9.2 and 9.3, we find that the ISW effect has much stronger visually distinct non-Gaussian features. For comparison, we Gaussianise this map and plot it in Figure 9.8.

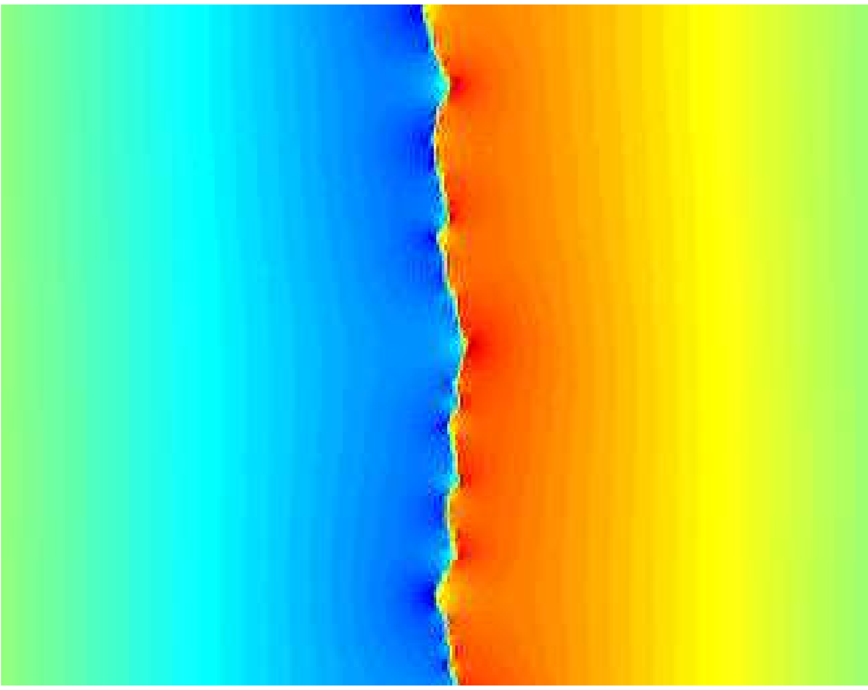

Figure 9.5: The Kaiser-Stebbins effect in the CMB anisotropies. There is one vertical wiggly string moving to the right hand side. The red (blue) colour means hotter (cooler) regions.

Using the SFL approximation (9.10), we also compute the angular power spectrum of Figure 9.6, which is shown in Figure 9.9. We see that the spectrum seems to be flat at smaller l, and decays as a power law l^{-1} towards larger l. This shows the intrinsic behaviour of the power spectrum of the string-induced ISW effect. However, because the dynamic range here is rather small with a low resolution, and we have used the SFL flat-space approximation which is marginal for the current angular size and dynamic range, we do not intend to draw information for the normalisation of the power spectrum presented here. Nevertheless, this preliminary result still provides a rough idea of the power-law decay of the spectrum for the string-induced ISW effect, and the maps simulated here can potentially provide good samples for identifying the non-Gaussian features of cosmic strings.

9.6 Conclusion

In this chapter, we computed both the primary CMB anisotropies and the ISW effect in the context of cosmic string models. Although the current work still carry considerable uncertainties in the normalization of the angular power spectra, these maps could potentially provide good samples for identifying the non-Gaussian features in the string-induced CMB anisotropies. The future quantitative improvement of the work will be to carry out a simulation with a larger dynamic range, an earlier initial time, a smaller initial horizon size, and the proper inclusion of small loops and gravitational waves. With sufficient computational resources, eventually one should use the Boltzmann code (e.g. the CMBFAST [85]) to simulate the string-induced CMB maps. This is also one of the main programmes currently under development.

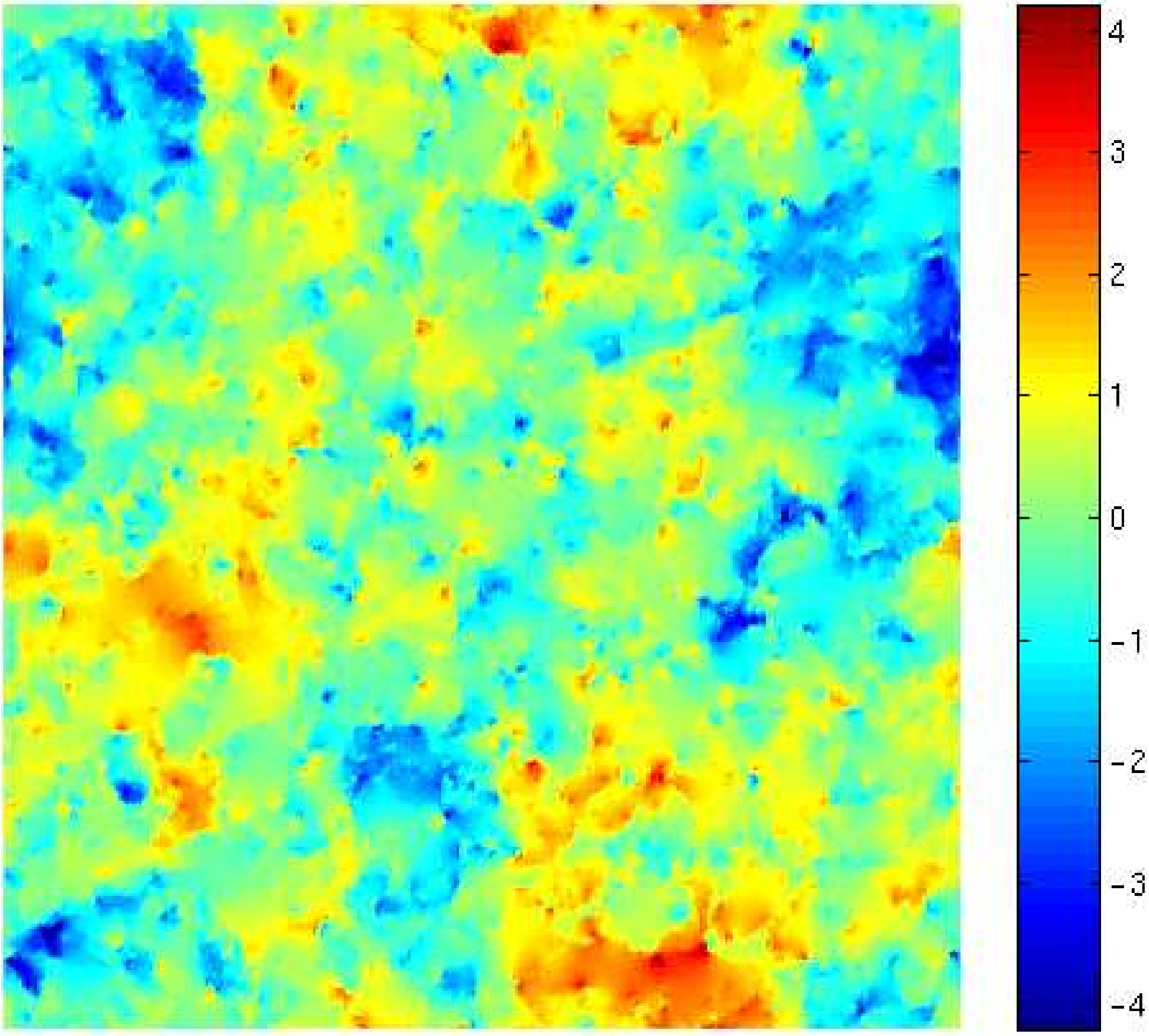

Figure 9.6: A map of the string-induced ISW effect. The dynamic range is from $\eta_{\rm d} \approx 6\eta_{\rm eq}$ to $30\eta_{\rm eq}$. The angular size is $(5.8°)^2$, and the resolution is 256^2. We have used $\Omega_{\rm m0} = 1$, $\Omega_\Lambda = 0$ and $h = 0.7$. The dynamic range is from $\eta_{\rm d} \approx 6\eta_{\rm eq}$ to $30\eta_{\rm eq}$. The colour scale shows the temperature contrast $\Delta T/T$ in units of one standard deviation. The standard deviation is $\sigma_{\rm ISW} \approx 4.7 \times 10^{-6}$ for $G\mu_6 = 1$.

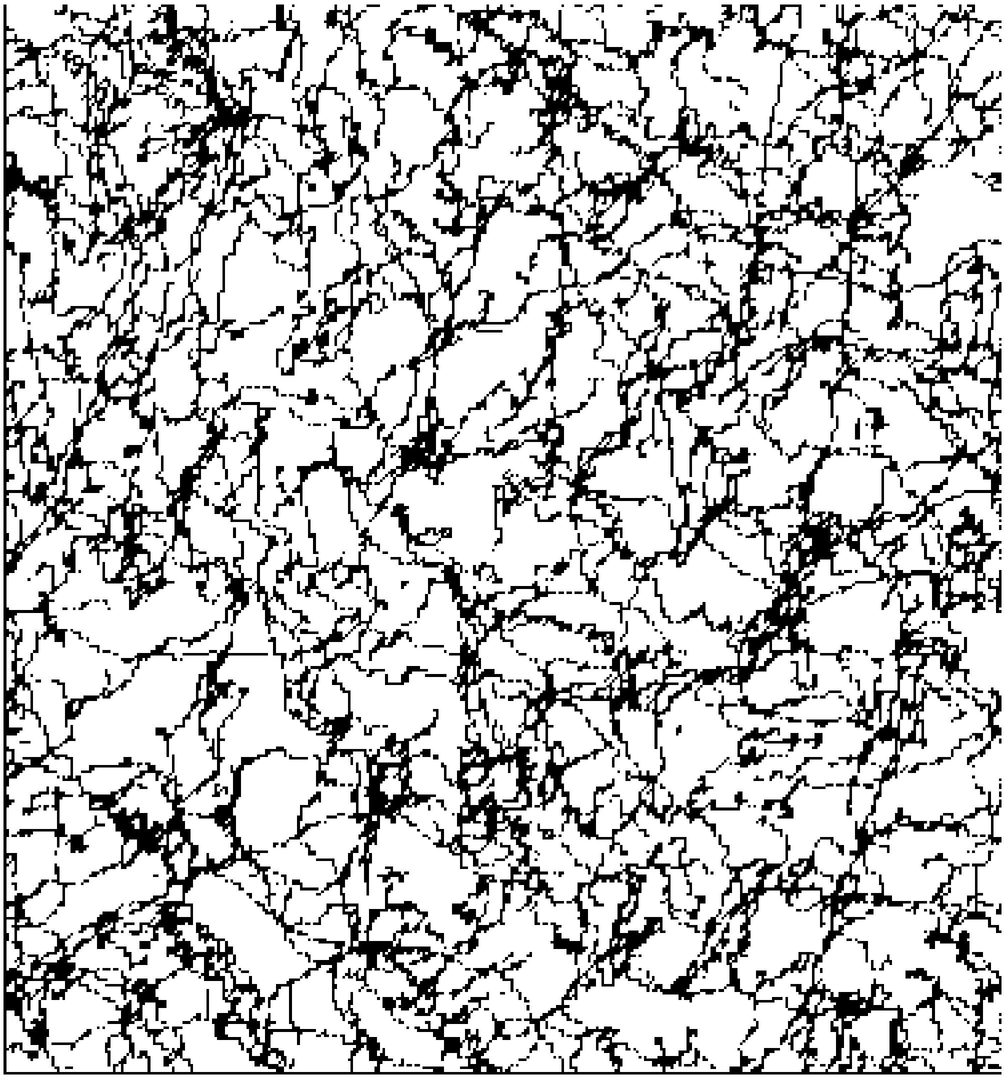

Figure 9.7: The string paths on the past light cone between the last scattering surface $\eta_d \approx 6\eta_{eq}$ and the time $30\eta_{eq}$. These paths correspond to the same realisation as Figure 9.6.

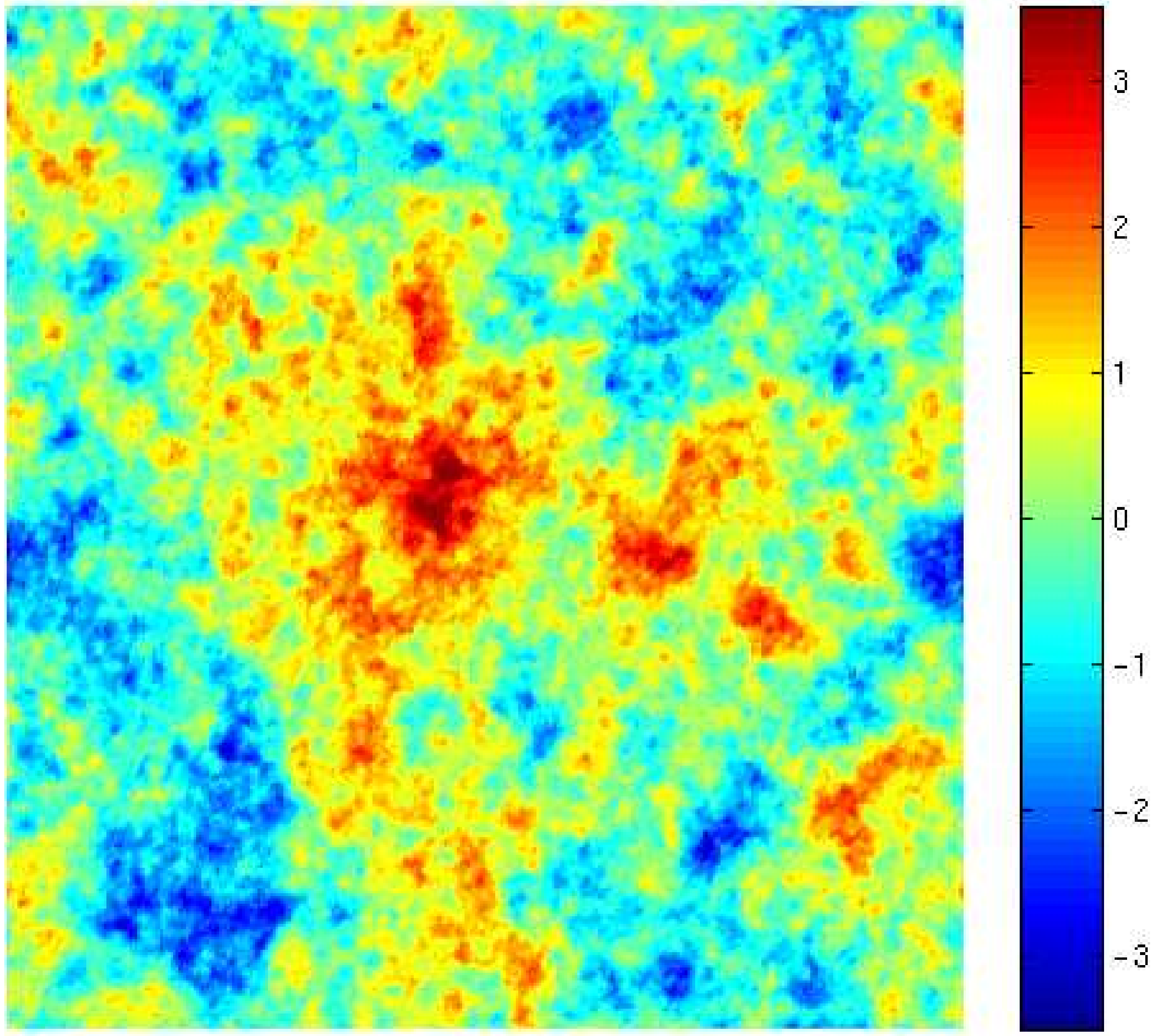

Figure 9.8: A CMB map Gaussianised from Figure 9.6.

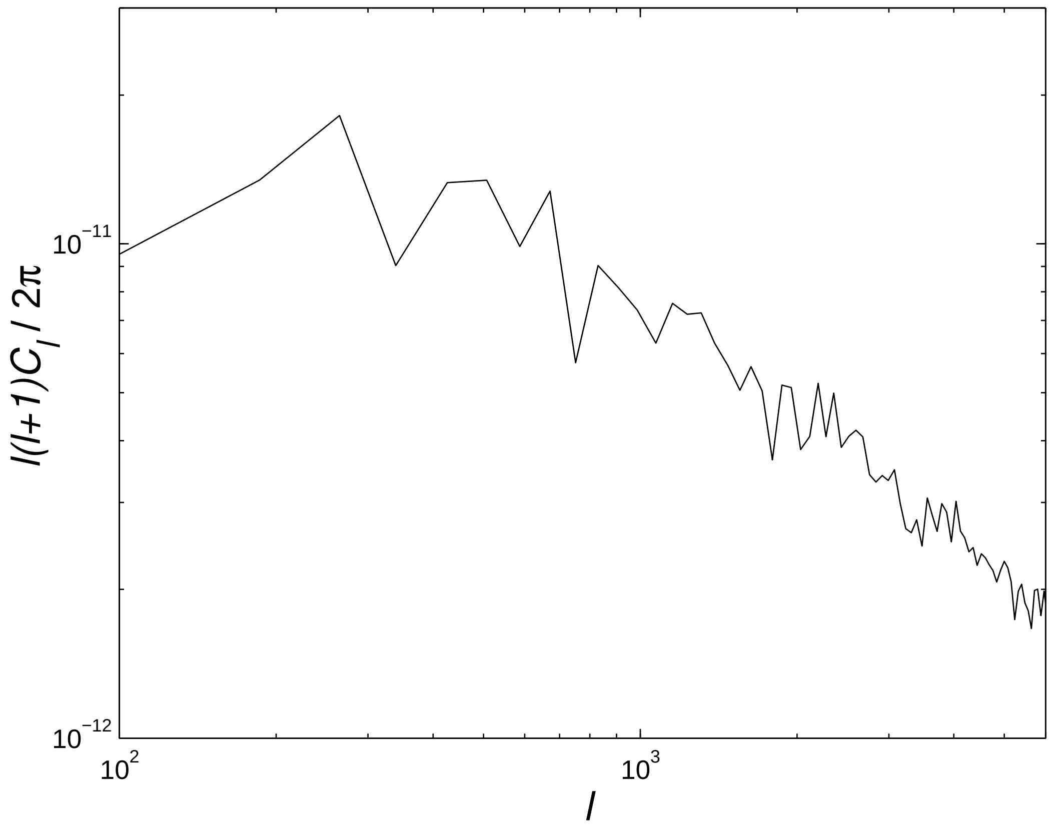

Figure 9.9: The angular power spectrum of Figure 9.6, the string-induced ISW effect. The dynamic range is five in conformal time, starting from $\eta_{\mathrm{d}} \approx 6\eta_{\mathrm{eq}}$ and finishing at $30\eta_{\mathrm{eq}}$. The normalisation is arbitrary.

Chapter 10

CMB Non-Gaussian Tests I—Extracting the NG Signal

10.1 Introduction

As mentioned at the end of Chapter 1, apart from the calculation of LSS and CMB spectra, the search for non-Gaussianity also plays an important role in testing current theories of structure formation. While standard inflationary models predict Gaussianity, there exist theories giving non-Gaussianity. For example, isocurvature inflation [181, 182] and topological defects [38] both provide mechanisms for generating non-Gaussian perturbations. Thus identifying and pinning down the non-Gaussian features of these theories is a task needing a comprehensive formalism with which to confront the flood of high-accuracy data in the near future (for a review see Ref. [164]). Moreover, should non-Gaussianity be found, further cosmological challenges would be faced. First, the standard inflationary cosmology that helps solve some important cosmology conundrums would be insufficient to explain the structures of the observed universe since it predicts only Gaussianity. Second, the power spectra of LSS and CMB, namely $\mathcal{P}(k)$ and the C_l curve respectively, would be insufficient to characterise the statistical properties of our universe and therefore higher-order statistics would be needed. Finally, because most frameworks for data analysis like those for the CMB assume Gaussianity, they would obviously be inadequate for non-Gaussian data. This would affect for example the estimate of cosmological parameters (e.g. $\Omega_{\mathrm{m}0}$, $\Omega_{\mathrm{B}0}$, h, Λ, etc.), because a non-Gaussian likelihood function would be required (e.g. [183]).

The best way to search for non-Gaussianity is via the CMB data, due to its cleanness in preserving the original statistical properties of perturbations. This search includes two steps. First we need to invoke statistical tools to characterise the non-Gaussian features of a underlying theory. Then we apply the same statistics to observations to test for consistency. However, in reality when dealing with the observational data, there are normally other components obscuring the primordial anisotropies which we would like to identify. The first component is foreground contamination. It can result from the radio emission of our own galaxy and point sources like individual galaxies. The point sources can directly block the CMB signal, as well as causing gravitational lensing effects in their surrounding regions. It has been shown that some attempts at removing point sources result in insufficient power in the CMB [184, 185, 186]; the theoretically expected lensing effect will operate at a small angular scale of around ten arc minutes [187]. The second component is due to instrumental effects. This could be in the form of white noise or a signal with a particular spectrum depending on the hardware design. This component sets a lower limit for the signal, below which it becomes unreliable. The last component is the cosmic and sample variance. It is especially important on large angular scales or in situations where only a very limited amount of data can be obtained for statistical analysis.

In addition to the observational obstruction mentioned above, there are also theoretical difficulties with regard to the design and use of statistics. The first aspect to face is the central limit theorem (CLT). Even if the underlying mechanism induced very non-Gaussian anisotropies at any single time, the time-integrated result could be almost Gaussian provided that both the orientations and positions were sufficiently random and numerous within the area of observation. This mechanism provides an extra Gaussian component in the underlying CMB map, which further obscures the non-Gaussian signal (if present), and this is why it is anticipated that the primary perturbations from before last scattering (see section 9.2) should be more or less Gaussian as a whole. The second challenge is the incomplete sky coverage due to the Galactic cut. This geometric effect is problematic for large-angle or global statistics like genus, Minkowski Functionals, etc. With both these observational deficiencies and theoretical difficulties it is possible to create spurious non-Gaussian signals. A good statistic therefore would be one that is easy to apply to CMB observations, which is less sensitive to the CLT effect, and capable of extracting primordial non-Gaussianity in the face of obscuring components such as foregrounds and instrument noise.

To this end, we shall propose a powerful yet simple method later in this chapter to extract the non-Gaussian signal in the CMB. The main idea is to renormalise the power in Fourier space to white noise, and thus make the non-Gaussian signal in real space more pronounced. we shall show that as long as the NG signal dominates within a certain range of the accessible scales, no matter how strong the superposed Gaussian background is (e.g. a million times the RMS amplitude of the non-Gaussian signal), our treatment can always successfully extract the non-Gaussian signal such that it becomes distinctly visible in real space and hence detectable by most traditional non-Gaussian statistics. With this method, the non-Gaussian signal can be obscured only by instrument noise of at least the same strength on scales where the NG signal dominates, and this is the theoretical limit for all statistics in any case. We verify our method against several different types of artificial non-Gaussian maps, as well as the theoretically simulated CMB maps induced by cosmic strings using a toy model. In particular, this method is shown to be capable of locating string paths on the past light cone since last scattering.

In the next section, we shall first review various statistics proposed for the CMB NG test, and their success in confronting the observations. In section 10.3, the new method of extracting non-Gaussianty from a CMB map will be introduced, with a main focus on its theoretical basis. We shall then apply this new method to CMB maps in section 10.4. This includes justifications via both simple artificial CMB maps and more realistic simulated situations. Finally in section 10.5, we discuss the future of this method.

10.2 A review of NG tests in the CMB

10.2.1 Statistical spaces

There are broadly three different statistical spaces, in which various statistics are usually defined. They are the real, multipole (or Fourier in the SFL, see section 9.3), and wavelet spaces. We shall briefly introduce their definitions, as well as the advantages and disadvantages of each choice.

1. Real space: $\frac{\Delta T}{T}(\hat{\mathbf{n}}) \equiv \frac{\Delta T}{T}(\theta,\phi)$, which is the CMB anisotropies observed along a direction $\hat{\mathbf{n}} \equiv (\theta,\phi)$. This is the conventional space where statistics are normally defined. The advantage of designing statistics in real space is that it preserves the local information of the underlying field. However, in real space all components in the CMB map mentioned earlier are mixed up in position, so that the separation or filtering of these components becomes difficult.

2. Multipole/Fourier space: As shown in section 1.5.2, the Multipole transform is

$$a_{lm} = \int d\Omega \frac{\Delta T}{T}(\theta,\phi) Y^*_{lm}(\theta,\phi), \tag{10.1}$$

where $d\Omega \equiv \sin\theta d\theta d\phi$. In the SFL, i.e. when $l \gg 1$ (see section 9.3), a small patch in the CMB sky can be treated as flat so that the multipole transform decays to a Fourier transform. This means

$$a_{\mathbf{k}} = \int dx^2 \frac{\Delta T}{T}(\mathbf{x}) e^{-i\mathbf{k}\cdot\mathbf{x}}, \tag{10.2}$$

so that

$$\begin{aligned} C_l &\approx C_k = \frac{\langle |a_{\mathbf{k}}|^2\rangle}{A}, && (10.3)\\ l(l+1)C_l &\approx k^2 C_k = \frac{k^2}{A}\langle |a_{\mathbf{k}}|^2\rangle, && (10.4)\\ l &\approx k = |\mathbf{k}|. && (10.5) \end{aligned}$$

Here A is the area of the small patch, $\mathbf{x}$ the two-dimensional coordinates on the patch, and $\mathbf{k}$ the wave number. The advantage of statistics designed in this space is that they are scale-dependent. They are also convenient for interferometric experiments to use, because interferometers measure signals directly in Fourier space, i.e. on the so-called (u,v)-plane. On the other hand, one disadvantage of this space is that non-Gaussian signals are easily obscured by the central limit theorem (CLT), and this can be seen implicitly from equation (10.2).

space	advantage	disadvantage
real	local information preserved	NG merged with G b.g. in position
multipole/ Fourier	scale-dependent; handy for interferometers	NG easily obscured by CLT
wavelet	local information preserved; scale-dependent	subtlety in the choice of $\phi(\mathbf{x})$

Table 10.1: Summary of advantages and disadvantages of various statistical spaces.

3. Wavelet space: This transform is obtained by convolving a field with both translations and dilations of a so-called parent function. In the SFL, a general form of the wavelet transform is

$$\left.\frac{\widehat{\Delta T}}{T}\right|_{l,j} = \int d^2x \frac{\Delta T}{T}(\mathbf{x})\phi_{l,j}(\mathbf{x}), \tag{10.6}$$

where $\phi_{l,j}(\mathbf{x}) = \phi(\frac{\mathbf{x}-\mathbf{x}_l}{k_j})$, and $\phi(\mathbf{x})$ is the parent function which satisfies $\int \phi(x)dx = 0$. This space not only separates scales but also preserves local information. For this reason, it has attracted considerable attention in the literature recently (e.g. [188, 189, 190]). However, the choice of the parent function $\phi(\mathbf{x})$ remains a subtle task, depending on the feature of the non-Gaussian signal under consideration.

In Table 10.1, we summarise both the advantages and disadvantages of these three commonly used statistical spaces.

10.2.2 Various statistics and literature review

We now review various statistics proposed for CMB NG tests, and their success in confronting observations. We have divided these statistics to several categories, each starting with introductory definitions:

1. Moments of order $n = k_1 + \ldots + k_m$: For m variables, the moments of order n are defined as

$$\mu_{k_1\ldots k_m} \equiv \left\langle x_1^{k_1}\ldots x_m^{k_m}\right\rangle = i^n \left[\prod_{j=1}^{m}\left(\frac{\partial}{\partial p_j}\right)^{k_j}\right] f(\mathbf{p}), \tag{10.7}$$

where $f(\mathbf{p}) \equiv \langle \exp(-i\mathbf{p}\cdot\mathbf{x})\rangle$ is the moment generating function. The moments of variables normalised to their mean are called the 'central moments', which are normally denoted as $\bar{\mu}_{k_1\ldots k_m}$.

(a) For $m = 1$, one defines skewness as $\bar{\mu}_3/\bar{\mu}_2^{3/2}$, kurtosis as $\bar{\mu}_4/\bar{\mu}_2^2 - 3$, etc. This class of statistics have been investigated in Ref. [185, 191], and is normally applied to describe the property of the one-point PDF. In Ref. [82], the authors used the one-point PDF to study the non-Gaussianity of cosmological perturbations induced by global defects.

(b) For $n = 3$ with $k_1 = k_2 = k_3 = 1$, the resulting moment μ_{111} is called the 3-point correlation function. The COBE data at the frequency of 53 GHz was tested using this statistic, and found to be Gaussian [185, 186]. A new statistic S_3 defined in a similar manner was proposed by Ref. [184], which found the CMB data UCSB SP91 to be non-Gaussian.

(c) For $n > 3$, it is normally too computationally expensive to apply to observational data.

- Bispectrum: This is an application of the moment, which is defined as

$$B_{l_1l_2l_3,m_1m_2m_3} = \langle a_{l_1m_1}a_{l_2m_2}a_{l_3m_3}\rangle\,. \tag{10.8}$$

For $l_1 = l_2 = l_3$, the COBE DMR data was found to be Gaussian for $l = 2,4,6,8,10$ in Ref. [192]; the same data set was found non-Gaussian for $l \approx 16$ at 98% confidence level in Ref. [193]. It was argued in Ref. [194] that this NG component resulted from contamination by foregrounds and noise. Ref. [195] also argued that this NG signal originated from only a small patch of the whole CMB sky.

- Hotelling's T^2 statistic: This is constructed from the bispectral estimates [185]. For small angular scales ($\theta \lesssim 1^\circ$) in the SFL, the data RING from OVRO was tested and found to be non-Gaussian [185, 186].

2. Cumulants of order $n = k_1 + \ldots + k_m$ are defined as

$$\chi_{k_1 \ldots k_m} \equiv i^n \left[\prod_{j=1}^{m} \left(\frac{\partial}{\partial p_j} \right)^{k_j} \right] \ln f(\mathbf{p}). \tag{10.9}$$

They are obtained by replacing the generating function of moments with its logarithm, so that the resulting statistics—the cumulants—become additive. Another advantage is that they are computationally cheap, i.e. the computation time scales as $O(N)$ [190], where N is the number of the samples. One example of how to design them as scale-dependent statistics is to combine them with the Fourier/wavelet transform, which decomposes the signal into domains of different scales.

   (a) For $m = 1$, a new technique invoking cumulants was proposed in Ref. [190].

   (b) For $m = 2$, it was theoretically estimated that the NG mechanism in the CMB resulting from the weak lensing should operate at the angular size of about $10'$ [187], but this scale is too small for current observations to probe.

   (c) For $m = 3$, as we shall see in Chapter 11, one can use them together with the Fourier transform to design a new method as a NG test of the CMB.

3. Minkowski Functionals V_i [196, 197, 198]: This is based on integral geometry. According to Hadwiger's Theorem, to measure the morphology of a d-dimensional pattern, the $d + 1$ Minkowski functionals are sufficient:

$$V_i = \frac{\Gamma(i/2)}{2\pi^{i/2} C_j^d} \oint_{\partial Q} dA\ \mathcal{Q}_i(\kappa_1, \ldots, \kappa_{d-1}), \tag{10.10}$$

where Γ is the normal Gamma function, C_j^d is the binomial coefficient, ∂Q is a $(d-1)$-dimensional boundary, κ_i is the geodesic curvatures, and $\mathcal{Q}_i$ is the ith elementary symmetric function defined by the polynomial expansion

$$\prod_{j=1}^{d-1} (x + \kappa_j) = \sum_{i=1}^{d} x^{d-i} \mathcal{Q}_i(\kappa_1, \ldots, \kappa_{d-1}). \tag{10.11}$$

Because $d = 2$ for the CMB data, one needs three Minkowski Functionals (V_0, V_1, V_2) to describe its property. Ref. [198] found that the COBE data was Gaussian at $\sim 4°$.

- Euler-Poincaré (EP) characteristics (Γ_u), or genus (g) [185, 186, 199, 200]: This is a variant of Minkowski Functionals. For the CMB sky, it can be deduced from the Minkowski Functionals as

$$g = V_2 + \frac{V_0}{2\pi R^2}, \tag{10.12}$$

where R is the radius of the last-scattering sphere. In the SFL, R approaches infinity, so that $g = V_2$. Conceptually, genus can be defined as in equation (7.3), i.e. $g \equiv \#\ \text{of holes} - \#\ \text{of parts} + 1$. It has been shown that genus is a poor discriminator between G and NG fields when the noise is comparable to the signal [185]; The COBE DMR data was tested and found to be Gaussian [186, 201, 202].

- Peak's statistics [203, 204]: It also describes the geometry of a field. It is defined as the number density of peaks (maxima or minima) with a specified threshold. Ref. [205] used it to address the NG feature of global defects.

- Another statistical method relevant to describing the geometry of a field is the correlation of excursion sets above a given threshold. It was discussed in Ref. [206].

4. Rings in Fourier space (in the SFL): This class of statistics are defined on a ring of a certain radius in the Fourier space. It has been used to define the 'non-Gaussian spectra', which abstract the phase information from an underlying field [207]. Another new statistic was proposed by implementing a second Fourier transform along the ring [208]. This method was shown to be very sensitive to the CLT. In the next chapter, we shall further discuss and invoke the concept of using these rings.

5. Techniques with wavelets [188, 190, 209]: Using wavelets transform, Ref. [209] found correlation between 11° and 22° near North Galactic Pole. Ref. [188] also used this technique to propose a scheme in the SFL for testing the NG in the CMB.

As we can see, various tests of the CMB data using different techniques have drawn no robust conclusion about the Gaussianity of our CMB sky. The main difficulty has been due to the low resolution and low accuracy of the current data. This situation is expected to be considerably improved in the next decade (for a review see Ref. [164]). However, the design of the NG statistic itself is also as vital as the quality of the observational data. Since the degrees of freedom in a data set are equal to the number of sampled points for the temperature distribution, in theory one needs the same number of statistical measures in order to probe the whole domain of the observation. Because this is impossible to implement in a normal situation, there is a need to design statistical schemes which can efficiently identify the NG signal in a CMB field. As seen above, although different statistic schemes have different capabilities for identifying specific features of a field, a complete and reliable treatment has yet to be found. Here we shall propose a new method with regard to improving this situation.

10.3 A new method—removing the Gaussian component

Let us first consider what is meant by Gaussianity [8]. In terms of n-point correlation functions, we know by definition that the only non-zero components in Gaussian models are the one- and two-point correlation functions. Since for the CMB we expect the mean to be zero by definition, the only information that any Gaussian field carries is the two point correlation function, or equivalently, the power spectrum in Fourier space. Therefore we can refer to the power spectrum as the 'Gaussian component' of a field, and any extra information apart from the power spectrum provides the non-Gaussianity. Thus, the main concept used in the new method is rather simple. That is to remove this Gaussian component, i.e. to 'remove' the power on all scales in Fourier space, so that the underlying field is 'whitened' with a white noise power spectrum while keeping all the phase information unchanged. More explicitly, we outline the procedure in the small-field limit to achieve this purpose:

1. Fourier transform the underlying field $\Delta(\mathbf{x}) \equiv \Delta T(\mathbf{x})/T$ to obtain $a_{\mathbf{k}}$ as in equation (10.2). This step is unnecessary for interferometer experiments since they measure Fourier signals directly on the (u, v)-plane.

2. Use equation (10.3) to estimate the power spectrum $C_l \approx C_k$.

3. Divide each $a_{\mathbf{k}}$ by $C_k^{1/2}$, so that the field is 'whitened':

$$\widetilde{\Delta}^{\mathrm{W}}(\mathbf{k}) \equiv a_{\mathbf{k}}^{\mathrm{W}} = \frac{a_{\mathbf{k}}}{C_k^{1/2}}, \tag{10.13}$$

   where we have used the superscript 'W' to denote a whitened field.

4. Fourier transform $a_{\mathbf{k}}^{\mathrm{W}}$ back to real space:

$$\Delta^{\mathrm{W}}(\mathbf{x}) = \int dk^2 a_{\mathbf{k}}^{\mathrm{W}} e^{i\mathbf{k}\cdot\mathbf{x}}. \tag{10.14}$$

   We notice that the normalisation here is irrelevant since the field has now been whitened, carrying no information regarding the perturbation amplitudes on all scales.

With such a procedure, one can now write the underlying field in Fourier space as

$$\widetilde{\Delta}(\mathbf{k}) = \widetilde{\Delta}^{\mathrm{W}} C_k^{1/2}, \tag{10.15}$$

or equivalently in real space as

$$\Delta(\mathbf{x}) = \Delta^{\mathrm{W}}(\mathbf{x}) \otimes D(\mathbf{x}), \tag{10.16}$$

where the symbol $\otimes$ denotes a convolution, and

$$D(\mathbf{x}) \equiv D_x = \int dk^2 C_k^{1/2} e^{i\mathbf{k}\cdot\mathbf{x}}, \tag{10.17}$$

which is the inverse Fourier transform of a spherically symmetric field $\widetilde{D}(\mathbf{k}) \equiv C_k^{1/2}$. Thus we see that we have decomposed the underlying field $\Delta(\mathbf{x})$ into two components: the 'Gaussian component' $D(\mathbf{x})$ which carries the information in the power spectrum, and the 'non-Gaussian component' $\Delta^{\mathrm{W}}(\mathbf{x})$ which possesses all the remaining information. Therefore, if $\Delta(\mathbf{x})$ is a Gaussian field, one should see $\Delta^{\mathrm{W}}(\mathbf{x})$ as purely white noise. If this is not the case, then $\Delta(\mathbf{x})$ is non-Gaussian. That is to say, we expect $\Delta^{\mathrm{W}}(\mathbf{x})$ to show us all the non-Gaussian features in $\Delta(\mathbf{x})$ if there are any.

To see how this formalism can work in practice, we consider a linear sum of Gaussian and non-Gaussian fields, namely Δ_{G} and Δ_{NG} respectively:

$$\Delta = \Delta_{\mathrm{G}} + \Delta_{\mathrm{NG}}. \tag{10.18}$$

Assuming no correlation exists between Δ_{G} and Δ_{NG}, we have $C_k = C_{k(\mathrm{G})} + C_{k(\mathrm{NG})}$. This leads to

$$\widetilde{\Delta}^{\mathrm{W}} = \widetilde{\Delta}^{\mathrm{W}}_{\mathrm{G}} \frac{C^{1/2}_{k(\mathrm{G})}}{\left[C_{k(\mathrm{G})} + C_{k(\mathrm{NG})}\right]^{1/2}} + \widetilde{\Delta}^{\mathrm{W}}_{\mathrm{NG}} \frac{C^{1/2}_{k(\mathrm{NG})}}{\left[C_{k(\mathrm{G})} + C_{k(\mathrm{NG})}\right]^{1/2}}. \tag{10.19}$$

Thus we see that if $C_{k(\mathrm{G})} < C_{k(\mathrm{NG})}$ at a certain k, the non-Gaussian component $\widetilde{\Delta}^{\mathrm{W}}_{\mathrm{NG}}$ will dominate. As we shall show in the next section, as long as $\widetilde{\Delta}^{\mathrm{W}}_{\mathrm{NG}}$ dominates within a certain range of k, even if this range is only a small part of the full range of k that the field can cover, the resulting Δ^{W} will show us this non-Gaussian feature. This sensitivity can greatly benefit non-Gaussian tests in the CMB, as we shall see later.

With the extracted non-Gaussian signal Δ^{W}, one can in principle further apply conventional methods to identify its non-Gaussian features. For example, one can use some non-Gaussian statistics like moments, cumulants, 3-point functions, Minkowski functionals, peak statistics, etc. However, as we shall see later, in most cases the non-Gaussian feature in Δ^{W} is already largely visually recognisable. Therefore, instead of trying further statistics to test our Δ^{W}, we shall concentrate more on the realistic problems facing observations. When necessary, we shall square the non-Gaussian signal Δ^{W} to amplify its non-Gaussian features.

10.4 Applications to CMB maps

10.4.1 Simulating the Gaussian background

To simulate the Gaussian background, we use the following power spectrum:

$$l(l+1)C_l = 3 \times 10^{-9} \times \exp\left[-\left(\frac{l}{1200}\right)^2\right]. \tag{10.20}$$

This form fits well on large scales ($l \gg 1$) with that of a standard CDM model with $\Omega_{\mathrm{c}0} = 0.95$, $\Omega_{\mathrm{B}0} = 0.05$ and $h = 0.5$, which we compute from CMBFAST [85]. It possesses an exponentially decaying tail, which we use to mimic photon diffusion damping [94], a generic effect on the last scattering surface (see section 3.3.6). The Gaussian field is then generated in multipole space by assigning a_{lm} for each mode (l, m), with both real and imaginary parts drawn from a Gaussian distribution of RMS amplitude $\sqrt{C_l/2}$. In the SFL using the conventions of (10.2)–(10.5), we draw the real and imaginary parts of $a_{\mathbf{k}}$ both from a Gaussian distribution of RMS amplitude $\sqrt{AC_k/2}$. With such an assignment in multipole or Fourier space, one can then obtain a Gaussian field by transforming a_{lm} or $a_{\mathbf{k}}$ back to obtain $\Delta(\hat{\mathbf{n}})$ or $\Delta(\mathbf{x})$ respectively. In Appendix B, we describe an algorithm to implement the full-sky multipole transform invoking the FFT, and show a simulated full-sky CMB map of about 52 million pixels, with an angular power spectrum generated from Ref. [85] (see Figure B.2). This map is the theoretically predicted CMB anisotropies from the standard CDM model.

10.4.2 A simple test

As a starting point, we test our method on a simple case, in which there are only two components in the field: a Gaussian background Δ_{G}, and a non-Gaussian signal Δ_{NG} consisting of 10 filled rectangles. The field size we shall use is $2° \times 2°$ with a resolution of 256^2, corresponding to $l \approx 2 \times 10^2$–2×10^4. The size of rectangles is 32×48 point^2, corresponding to $l \approx 1440$ and 960 respectively. Figure 10.1 shows this test. In (a), we plot the non-Gaussian component Δ_{NG}, the 10 filled rectangles. This non-Gaussian component is then superimposed onto a Gaussian background Δ_{G}, which has the power spectrum (10.20). In the resulting field $\Delta = \Delta_{\mathrm{G}} + \Delta_{\mathrm{NG}}$, the RMS amplitude of Δ_{G} is one million times larger than that of Δ_{NG}. Δ is plotted in (b), where we see that the rectangles are now indistinguishable. We then apply our procedure described in section 10.3 to obtain the whitened field Δ^{W}, and it is plotted in (c). As one can see, the shapes of the 10 rectangles are clearly recovered in the extracted non-Gaussian component Δ^{W}. Referring to the power spectra plotted at the bottom of Figure 10.1, we see that although the RMS amplitude of Δ_{G} is one million times larger than that of Δ_{NG}, resulting in $C_{l(\mathrm{G})}$ being 10^{12} times larger than $C_{l(\mathrm{NG})}$ on small l (~ 200), $C_{l(\mathrm{NG})}$ still dominates on larger l ($\gtrsim 7000$) due to the exponential decay of $C_{l(\mathrm{G})}$ towards smaller scales. This verifies that although C_l is

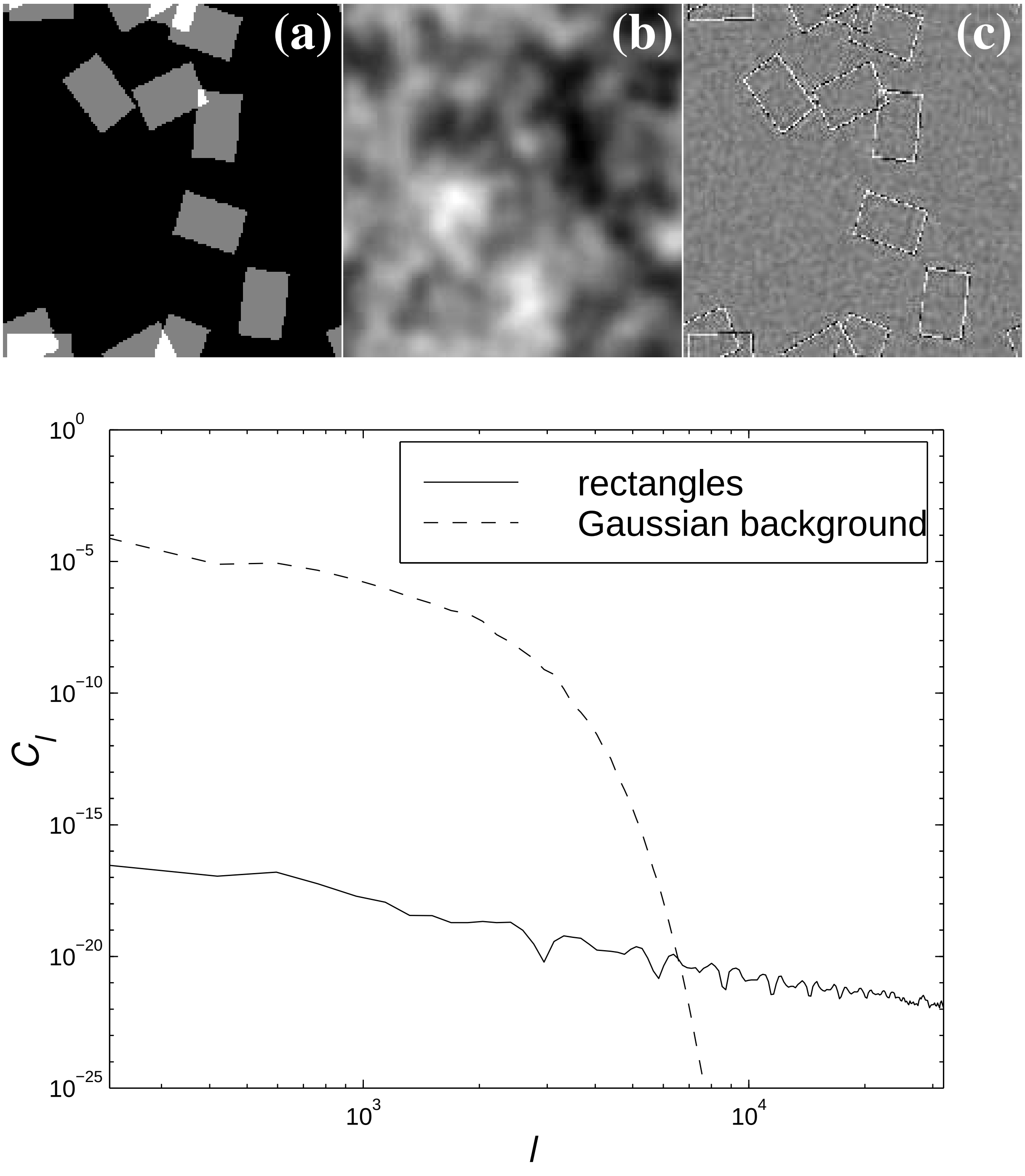

Figure 10.1: A simple test of our method for extracting the non-Gaussian signal. The field size is $2° \times 2°$, with a resolution of 256^2. On the top are: (a) $\Delta_{\rm NG}$ containing 10 filled rectangles; (b) $\Delta = \Delta_{\rm G} + \Delta_{\rm NG}$, i.e. 10 filled rectangles superimposed with a Gaussian background whose RMS amplitude is one million times larger (so that the rectangles are indistinguishable here); (c) the extracted NG signal $\Delta^{\rm W}$ using our new method. In each plot, the colour scheme interpolates linearly between the minima and the maxima, going from black to white, The bottom is the power spectra of the two components in (b): $C_{l({\rm G})}$ and $C_{l({\rm NG})}$.

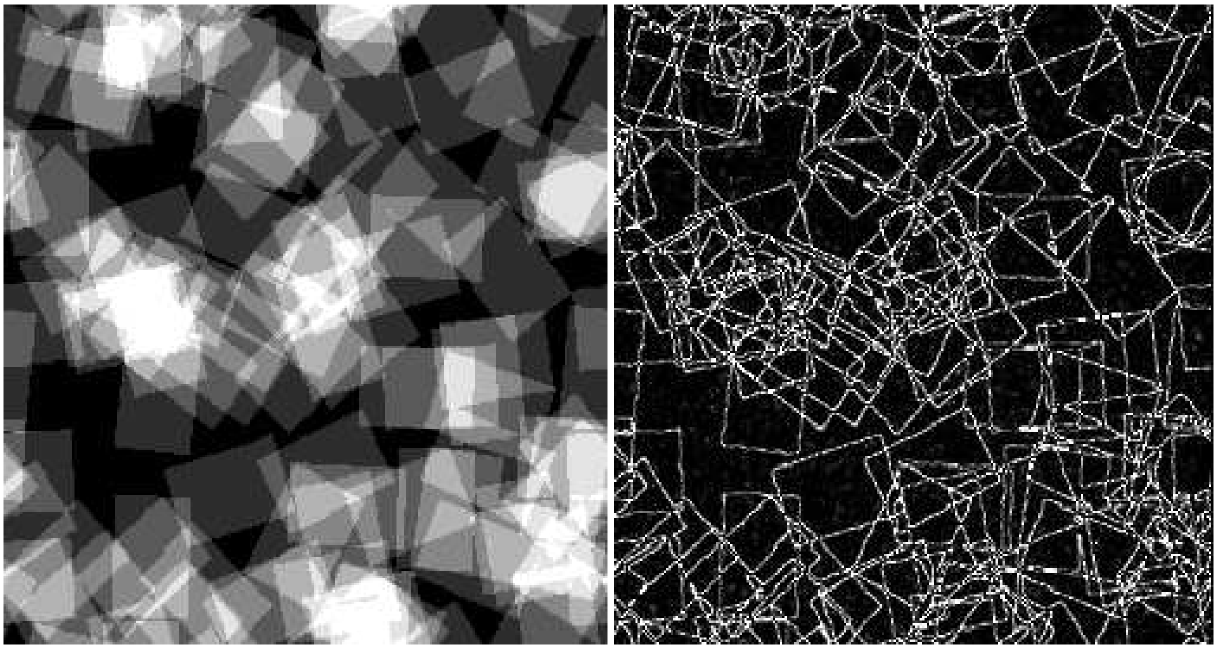

Figure 10.2: Insensitivity of the new method to the central limit theorem. The field size is $2^\circ \times 2^\circ$, with a resolution of 256^2. On the left is $\Delta_{\rm NG}$ containing 100 filled rectangles. This map is then linearly superimposed with a Gaussian background $\Delta_{\rm G}$ which we use in Figure 10.1. Its RMS amplitude is 10^6 times that of $\Delta_{\rm NG}$. We then use our new method to remove the power of the mixed map $\Delta = \Delta_{\rm G} + \Delta_{\rm NG}$, and on the right is the square of the extracted non-Gaussian signal $(\Delta^{\rm W})^2$. As seen, the shapes of 100 rectangles are clearly recovered without being obscured by the CLT.

dominated by $C_{l({\rm G})}$ for most l's, the non-Gaussian feature can still show up in $\Delta^{\rm W}$ as long as $C_{l({\rm NG})}$ dominates $C_{l({\rm G})}$ within a certain range of l.

In general cases, we do not know on which scale the NG signal dominates, and therefore can not apply a priori window function to remove the power on scales where the Gaussian component dominates. In addition, the overall amplitude of the field is normally dominated by the Gaussian background, as in the case presented above. Nevertheless, the method demonstrated here effectively makes power on all scales equally important, so that the NG signal shows up more easily in the whitened field $\Delta^{\rm W}$. One may argue that in the presence of point sources, all signals (the Gaussian and non-Gaussian ones) could be severely obscured by the non-Gaussianity of the points. However, as we shall show in section 10.4.4, our method is capable of not only removing the Gaussian signal, but also keeping the intrinsic NG features of each NG component, even some of which dominate others. In the simple example demonstrated above, we also note that although the NG signal dominates above a unrealistic high l, the purpose here is just to highlight the fact that our method is not affected by a strong Gaussian background, provided that the NG signal comes to dominate even just within a small range of the accessible l.

10.4.3 Insensitivity to the central limit theorem

In the literature, most statistics suffer from the CLT, especially for those designed in Fourier space. For example, NG events of a number as large as one hundred have shown to completely screw up a new NG statistic proposed in Ref. [208]. Therefore, to test the sensitivity of our formalism to the CLT, we consider a mixed field as in the previous section, but with $\Delta_{\rm NG}$ consisting of 100 filled rectangles this time instead of 10. Figure 10.2 shows the test. On the left is the $\Delta_{\rm NG}$ containing 100 filled rectangles. It is then linearly superimposed with a Gaussian background $\Delta_{\rm G}$ which we use in Figure 10.1. Its RMS amplitude is 10^6 times that of $\Delta_{\rm NG}$. We notice that the power spectra of these two components are the same as those plotted in Figure 10.1. We then use our new method to remove the power of $\Delta = \Delta_{\rm G} + \Delta_{\rm NG}$, resulting in $\Delta^{\rm W}$. On the right of Figure 10.2, we plot $(\Delta^{\rm W})^2$, the square of the extracted non-Gaussian signal. As seen, the shapes of 100 rectangles are clearly recovered without being obscured by the central limit theorem. Thus we have seen a treatment which combines both real and Fourier spaces to take the advantages of both, without being affected by their disadvantages (see Table 10.1).

10.4.4 Against various components in the CMB

In the face of realistic CMB observations, the signal normally contains various components resulting from different physical processes. Here we consider six different components, and then see if our method can extract out the non-Gaussian ones among them. The field size we shall use is $1^\circ \times 1^\circ$, with a resolution of 256^2. This corresponds to a range of $l \approx 4 \times 10^2$–5×10^4. Although such a large range is unrealistic for the observational data to be obtained in the near future, we choose this range only for the convenience of illustrating the power of our method under various circumstances. As illustrated in Figure 10.3, these six components are:

(a) 10 filled circles: 5 with a diameter of 64 points and the other 5 with a diameter of 128 points, corresponding to $l \approx 1440$ and 720 respectively. This component represents a certain type of non-Gaussianity.

(b) 10 rectangles: with a size of 32×48 point^2, corresponding to $l \approx 2880$ and 1920 respectively. This component represents a second type of non-Gaussianity, which is similar to the ISW effect produced by cosmic strings.

(c) 5 sharp circular points: each with a power spectrum

$$C_{l(\mathrm{c})} \approx C_{k(\mathrm{c})} = \exp\left[-(rRk)^2\right], \tag{10.21}$$

where R is the angular size of the field. Thus r indicates the size of the point as a fraction of the size of the field, and here we use $r = 0.005$. These five points represent point sources like galaxies, which block the CMB anisotropies and also induce strong point-like signals on the top of the CMB.

(d) 5 very diffuse circular points: each with a power spectrum

$$C_{l(\mathrm{d})} \approx C_{k(\mathrm{d})} = \frac{1}{1 + (sRk)^t}, \tag{10.22}$$

where R again is the angular size of the field. Thus s indicates the size of the point as a fraction of the size of the field and t shows the degree of diffusion. Here we use $s = 0.5$ and $t = 6$. This sort of signal can result from certain types of topological defects like textures [167]. As one can see from Figure 10.3 that they are so diffuse that their centres are difficult to locate visually, especially when mixed with the Gaussian background (see the left plot in Figure 10.5).

(e) a Gaussian background: with the power spectrum (10.20). This component represents the intrinsic CMB anisotropies, which are expected to be close to Gaussian due to the CLT. This is the dominant part in the observed CMB fluctuations.

(f) Gaussian white noise: with a constant power spectrum. This is expected from the instrument effects, and its amplitude normally sets a boundary above which useful information can be extracted from the underlying observation.

Thus we see that components (a)–(d) are non-Gaussian, while (e) and (f) are Gaussian.

To test our method in distinguishing the non-Gaussian components, we first generate these six components individually as shown in Figure 10.3, and then take the linear sum of them $\Delta = \Delta_{(\mathrm{a})} + \Delta_{(\mathrm{b})} + \Delta_{(\mathrm{c})} + \Delta_{(\mathrm{d})} + \Delta_{(\mathrm{e})} + \Delta_{(\mathrm{f})}$. The RMS amplitudes of each components in Δ are $\sigma_{(\mathrm{a})} : \sigma_{(\mathrm{b})} : \sigma_{(\mathrm{c})} : \sigma_{(\mathrm{d})} : \sigma_{(\mathrm{e})} : \sigma_{(\mathrm{f})} = 1 : 1 : 10 : 500 : 1000 : 0.2$. Figure 10.4 shows the power spectra of each individual components. We then apply our method to remove the Gaussian information (the power spectrum) in the mixed map Δ, and this results in an extracted non-Gaussian signal Δ^{W}. Figure 10.5 shows both Δ and $(\Delta^{\mathrm{W}})^2$. As we can see, all the non-Gaussian components (a)–(d) are clearly recovered in $(\Delta^{\mathrm{W}})^2$. Combined with Figure 10.4, we notice four things. First, the centres of the diffuse points in (d) are now recovered in $(\Delta^{\mathrm{W}})^2$ (see Figure 10.5), although hardly recognisable in $\Delta_{(\mathrm{d})}$ (see Figure 10.3). Second, although the non-Gaussian components ((a)–(d)) dominate the Gaussian ones ((e) and (f)) only within a certain range of l ($4000 \lesssim l \lesssim 40000$), their non-Gaussian features are almost fully recovered in Δ^{W}. Third, although the point sources (c) and (d) are much stronger than (a) and (b) on all scales, the non-Gaussian features of (a) and (b) can still be recovered in Δ^{W} without being obscured by (c) and (d). This is because these non-Gaussian components have different intrinsic features that are not correlated. Finally, we notice that although the ratios among these components seem to be unrealistic for the near future experiments, in particular the signal to noise being larger than five thousands, the choice of these ratios is just for the convenience of illustrating various capabilities of our method. Therefore, we see that our new method is indeed capable of extracting various non-Gaussian features in a CMB map, as long as they dominate the Gaussian power at least within a certain range of scales.

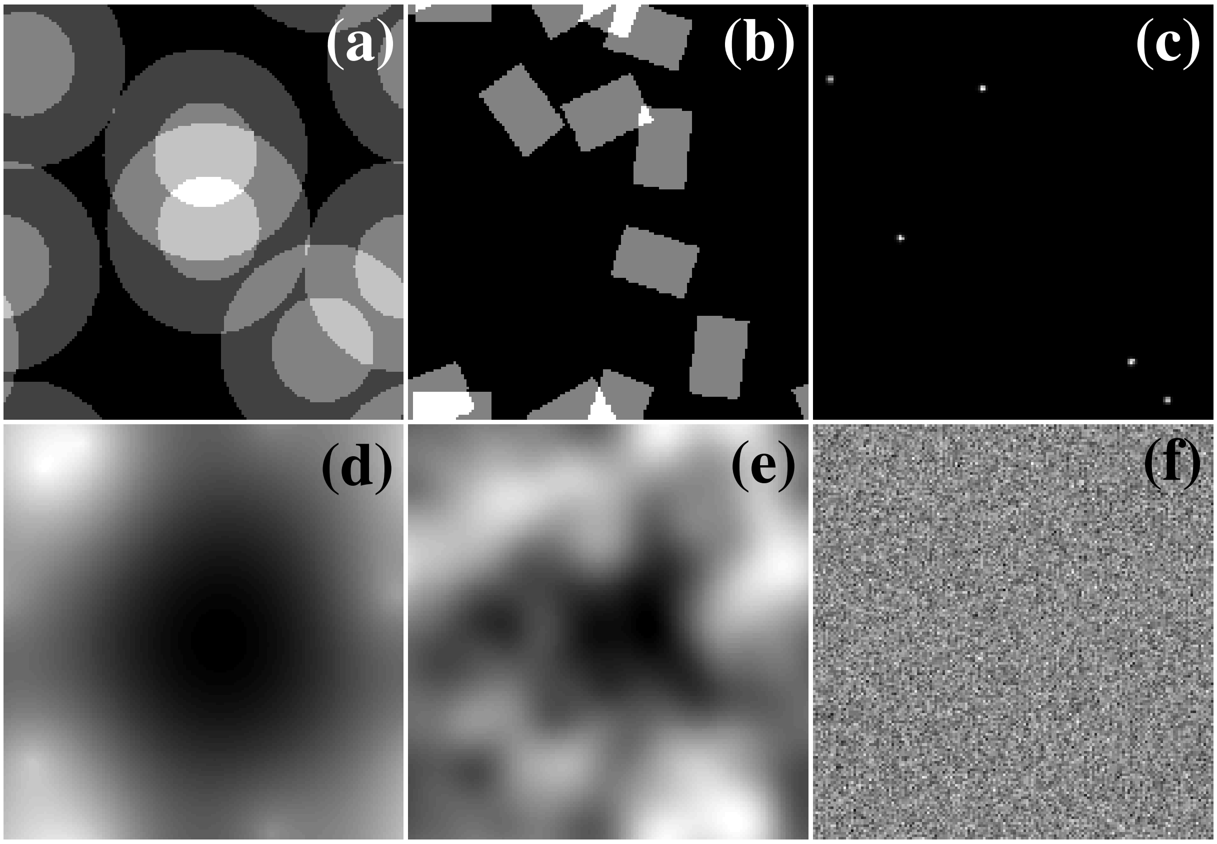

Figure 10.3: Six different components in an artificial CMB map of $1° \times 1°$: (a) 10 filled circles; (b) 10 filled rectangles; (c) 5 sharp point sources using equation (10.21) with $r = 0.005$; (d) 5 diffuse points using equation (10.22) with $s = 0.5$ and $t = 6$; (e) Gaussian background using equation (10.20); (f) white noise.

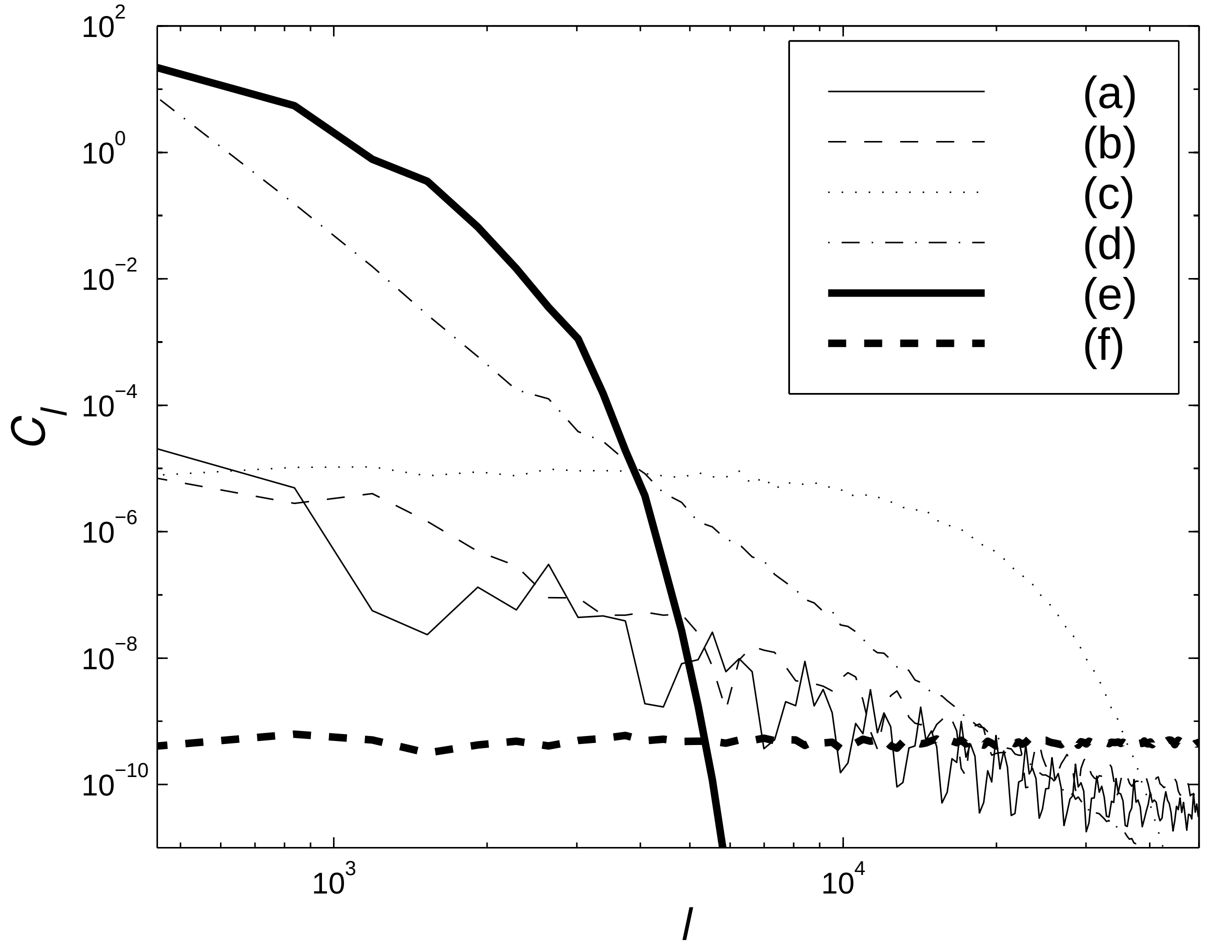

Figure 10.4: Power spectra of the six components (see Figure 10.3) in the artificial CMB map of $5^\circ \times 5^\circ$ (see the left plot of Figure 10.5). The labelling here is the same as in Figure 10.3. The RMS amplitudes of these components are $\sigma_{(\mathrm{a})} : \sigma_{(\mathrm{b})} : \sigma_{(\mathrm{c})} : \sigma_{(\mathrm{d})} : \sigma_{(\mathrm{e})} : \sigma_{(\mathrm{f})} = 1 : 1 : 10 : 500 : 1000 : 0.2$.

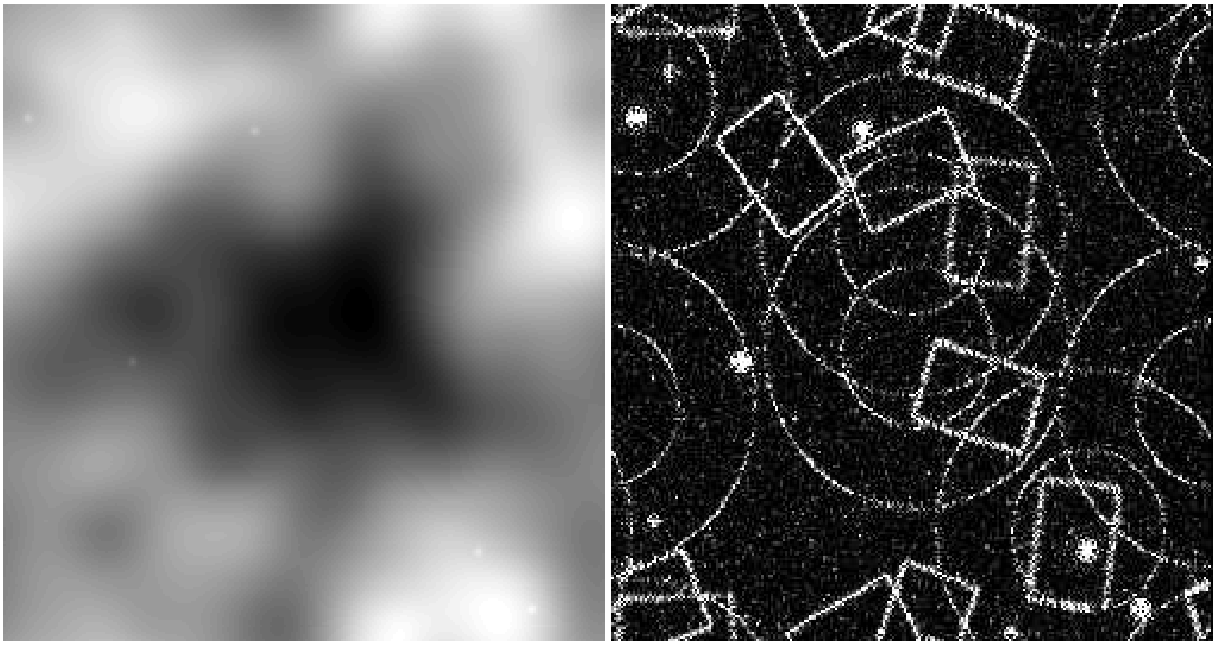

Figure 10.5: On the left is an artificial CMB map Δ, with the six components presented in Figure 10.3 all mixed together. Their individual RMS amplitudes are $\sigma_{(\mathrm{a})} : \sigma_{(\mathrm{b})} : \sigma_{(\mathrm{c})} : \sigma_{(\mathrm{d})} : \sigma_{(\mathrm{e})} : \sigma_{(\mathrm{f})} = 1 : 1 : 10 : 500 : 1000 : 0.2$, where the labelling here is the same as in Figure 10.3. On the right is $(\Delta^{\mathrm{W}})^2$, the square of the extracted non-Gaussian signal from Δ. As we can see that all the non-Gaussian components (a)–(d) are clearly recovered. The colour scheme in the left plot is linear between the minima and the maxima, going from black to white. That in the right plot is linear between 0 and 1, from black to white.

10.4.5 Window effects and instrument noise from experiments

In the face of small-field observational data, the $a_{\mathbf{k}}$ modes are normally seen with a given observing beam. Let us describe this primary beam by a window $W(\mathbf{x})$ and the observing beam by $B(\mathbf{x})$. For an interferometer, the sampled sky temperature is then

$$\Delta_{\mathrm{s}}(\mathbf{x}) = [\Delta(\mathbf{x})W(\mathbf{x})] \otimes B(\mathbf{x}). \tag{10.23}$$

For a single-dish experiment, it is

$$\Delta_{\mathrm{s}}(\mathbf{x}) = [\Delta(\mathbf{x}) \otimes B(\mathbf{x})]\, W(\mathbf{x}). \tag{10.24}$$

In reality, we can only observe the sampled distribution $\Delta_{\mathrm{s}}(\mathbf{x})$, while the underlying field to be investigated is $\Delta(\mathbf{x})$, which may consist of both Gaussian and non-Gaussian signals. This window effect will obscure the intrinsic field, and therefore its non-Gaussian features (if any). Thus in any theoretical framework, one needs to include this effect in order to convincingly justify the underlying method. We shall use a Gaussian form for both primary and observing beams. For the primary beam, we have

$$W(\mathbf{x}) = \exp\left(-\frac{x^2}{2\sigma_W^2}\right), \tag{10.25}$$

where σ_W is the RMS radius of the circularly symmetric Gaussian window $W(\mathbf{x})$. Since the observing beam is associated with a convolution in both equations (10.23) and (10.24), it would be more convenient if we define it in Fourier space:

$$\widetilde{B}(\mathbf{k}) = \exp\left(-\frac{k^2}{2k_B^2}\right), \tag{10.26}$$

where

$$k_B = \frac{1}{\sigma_B}, \tag{10.27}$$

and σ_B is the RMS radius of the circularly symmetric Gaussian window $B(\mathbf{x})$. Apart from σ_W and σ_B, one commonly used measure of the size of a window is the Full Width of Half Maximum (FWHM), x_{FWHM}. It is defined as the width of the window at it half maximum. For a Gaussian window, this can be calculated as

$$x_{\mathrm{FWHM}} = 2\,(2\ln 2)^{1/2}\,\sigma \approx 2.35\sigma, \tag{10.28}$$

where σ is either σ_W or σ_B in our case.

Another important instrumental effect is the noise originating from the observing mechanism of the instrument itself. This noise normally obscures the signal and has been a great challenge for all current experiments. It sets the lower limit in the amplitude of the signal, below which one can not yield useful information. In the following analysis, where we model the signal as measured from an interferometer, we shall simulate this effect by adding Gaussian noise into all points in Fourier space, with a strength 5% of the RMS amplitude of the Fourier signal.

10.4.6 CMB from cosmic strings

As seen in section 9.2, the string-induced CMB signal can be considered as a linear sum of two components: the primary perturbations $\Delta_{\rm prim}$ that originate from the last scattering surface, and the ISW effect $\Delta_{\rm ISW}$ that is generated on the past light cone between us and the last scattering surface. In the worst case for the non-Gaussian test, $\Delta_{\rm prim}$ is almost Gaussian due to the central limit theorem. Nevertheless, its power spectrum possesses a exponentially decayed tail towards larger l due to photon diffusion damping (e.g. [93]), so that we expect $\Delta_{\rm ISW}$ to dominate beyond certain l since it decays as power law as we have seen in section 9.5.3. Although Ref. [95] argued that the string-induced $\Delta_{\rm prim}$ will decay first as exponential and then as power law when $l \gtrsim 3000$ due to the survival of the late-time string-induced perturbations just before the decoupling, we expect the argument about the domination of non-Gaussianity over Gaussianity towards large l to be unchanged, because those late-time perturbations providing the power-law tail in $\Delta_{\rm prim}$ are apparently non-Gaussian.

To simulate the worst situation, we model $\Delta_{\rm prim}$ as purely Gaussian, with an angular power spectrum given by equation (10.20). As for the $\Delta_{\rm ISW}$, we first use a toy model which simulates important properties of cosmic strings, like self-avoiding and scaling, to generate a history of string evolution on the past light cone. We then use equations (9.13) and (9.16)–(9.17) to compute the string-induced ISW effect $\Delta_{\rm ISW}$. The dynamic range used to simulate this effect starts from $\eta_{\rm d} \approx 6\eta_{\rm eq}$ and finishes at around $240\eta_{\rm eq}$, which is nearly today. We also consider point sources which are inevitable in any observation. This component is simulated as the component (c) described in section 10.4.4, with the power spectrum (10.21) and $r = 0.01$. We then mix these three components together, i.e. $\Delta = \Delta_{\rm prim} + \Delta_{\rm ISW} + \Delta_{\rm pnt}$, with standard deviations $\sigma_{\rm prim} : \sigma_{\rm ISW} : \sigma_{\rm pnt} = 5 : 1 : 2$. The size of field is $(2°)^2$, with a resolution of 256^2. The window effects described in the previous section are then added in this Δ using equation (10.23), to simulate the temperature distribution sampled by an interferometer $\Delta_{\rm s}$. The size of the primary beam is $x_{{\rm FWHM}(W)} = 0.7$ (see equations (10.25) and (10.28)) in units of the field size. In the same units, the size of the observing window is $x_{{\rm FWHM}(B)} = 0.02$ (see equations (10.26) and (10.28)), which is approximately 5 grid spacings. The primary beam will constrain the visible area in the real space to be approximately 0.7 times of the field size, while the observing window will smooth away perturbations on scales smaller than five grid spacings. Finally, we add in the effect of instrument noise. This is achieved by Fourier transforming the $\Delta_{\rm s}$ obtained above, and then adding a Gaussian noise onto each grid in Fourier space. The RMS amplitude of the noise is 5% of the RMS amplitude of all Fourier modes. A further inverse Fourier transform yields the resulting temperature distribution, which is as observed from an interferometer.

With such a map, we can test our new method of extracting the NG signals. Figure 10.6 illustrates the procedure of the map making, as well as the extracted NG signals. Plot (a) is the point sources $\Delta_{\rm pnt}$, and plot (b) is the string-induced ISW effect $\Delta_{\rm ISW}$, with a dynamic range of 40 in conformal time starting from the last scattering. These two components are then mixed with a Gaussian background to yield plot (c), where the window effects and instrument noise are also included to simulate the CMB field sampled by an interferometer. The colour scheme here is within the three-sigma range. As one can see in plot (c), the signal strength is constrained in the central area of the field due to the primary beam. There is no fine structure either due to the convolution of the observing window. The point sources are clearly recognisible, but the string-induced ISW effect can hardly be distinguished in this plot. We then use our method (the procedure described as the steps 1–4 in section 10.3) to whiten the field (i.e. to remove the Gaussian component), and the extracted NG signals $\Delta_{\rm s}^{\rm W}$ are presented in plot (d) in Figure 10.6. The colour scheme is within three-sigma range. Apparently, although the point sources appear as strong as eighty sigmas in strength, the string-induced ISW effect has still been recovered as line structures. By comparing plot (b) and (d), we find that these line structures correspond exactly to the string paths on the past light cone.

We also calculate the power spectra of various components in plot (c) of Figure 10.6. They are plotted in Figure 10.7. As one can see, although the string-induced ISW effect $\Delta_{\rm ISW}$ is dominated by either $\Delta_{\rm prim}$ or $\Delta_{\rm pnt}$ almost on all scales, its NG feature is not obscured (see plot (d) of Figure 10.6). This is because, as explained previously, the power spectrum of $\Delta_{\rm ISW}$ is larger than that of $\Delta_{\rm prim}$ (which is the only Gaussian component apart from the noise) when $l \gtrsim 2000$, and the NG property of $\Delta_{\rm ISW}$ is intrinsically different from that of $\Delta_{\rm pnt}$ so that they do not mix up. Therefore, we now have certain confidence that our new method is capable of extracting the NG feature of string-induced CMB anisotropies from data sampled by an interferometer. If the

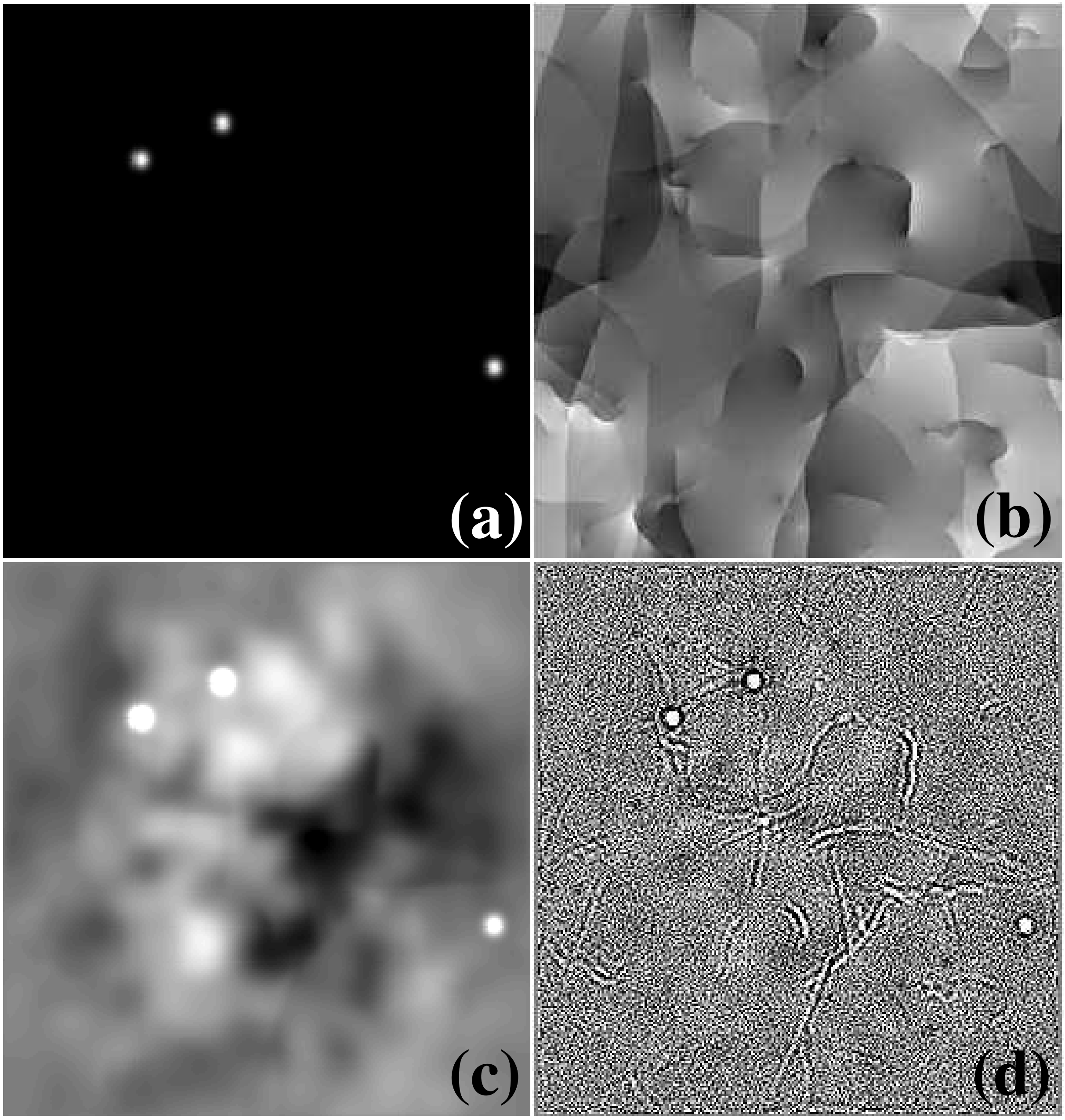

Figure 10.6: Simulated CMB maps of a size $(2^{\circ})^2$, with a resolution of 256^2. Plot (a) shows three point sources $\Delta_{\rm pnt}$. Plot (b) is the string-induced ISW effect $\Delta_{\rm ISW}$ using a toy model. Plots (a) and (b) are then mixed with a Gaussian primary perturbations $\Delta_{\rm prim}$, to form a field $\Delta = \Delta_{\rm prim} + \Delta_{\rm ISW} + \Delta_{\rm pnt}$, with standard deviations $\sigma_{\rm prim} : \sigma_{\rm ISW} : \sigma_{\rm pnt} = 5 : 1 : 2$. Window effects and a 5% instrument noise are then added into this field to form plot (c). Window sizes are $x_{_{\rm FWHM}(W)} = 0.7$ and $x_{_{\rm FWHM}(B)} = 0.02$ in units of the field size. Plot (d) is the NG signals $\Delta_{\rm s}^{\rm W}$ extracted from plot (c).

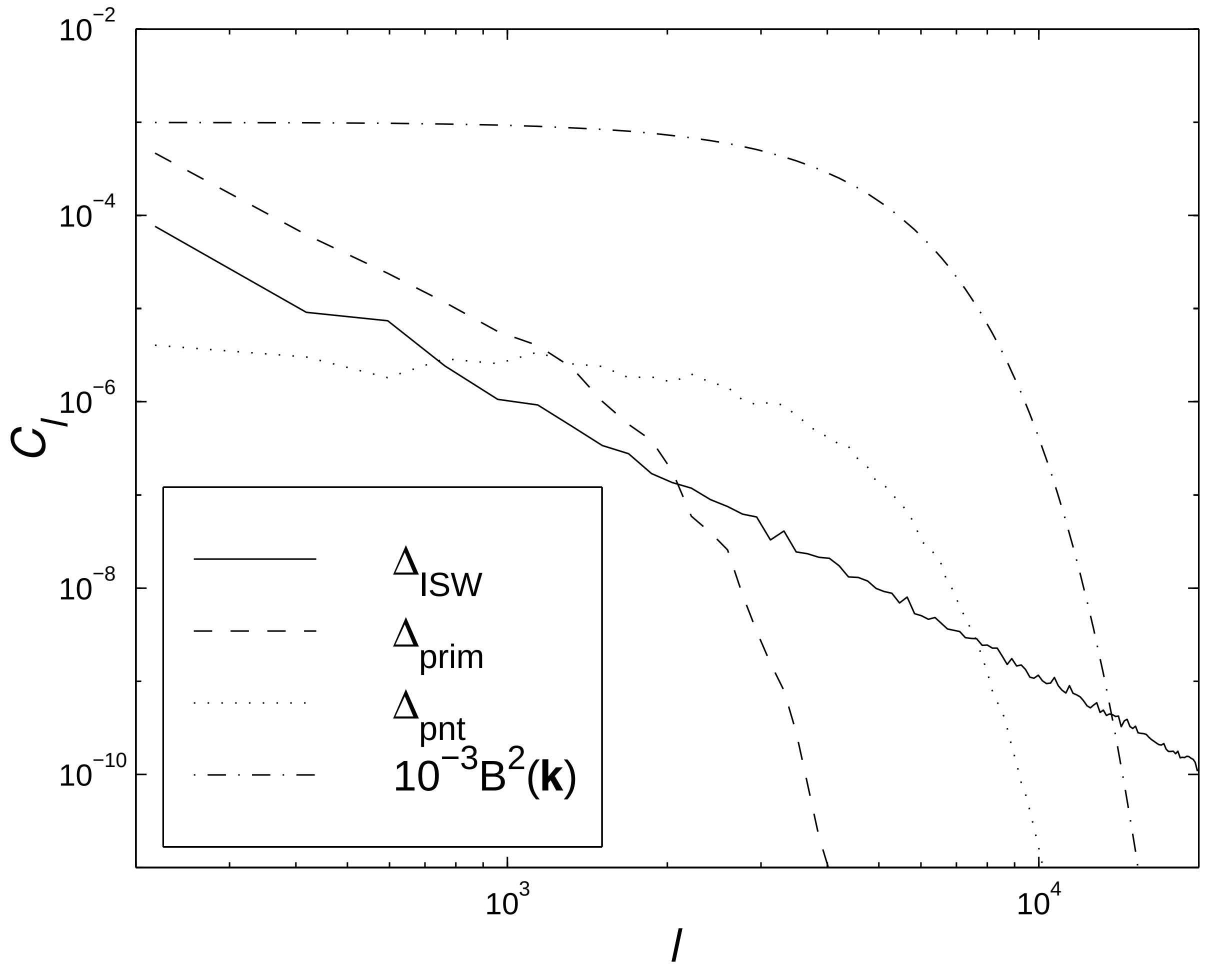

Figure 10.7: Power spectra of various components in plot (c), Figure 10.6, before adding in the noise and window effects. Their standard deviations are $\sigma_{\rm prim} : \sigma_{\rm ISW} : \sigma_{\rm pnt} = 5 : 1 : 2$. Also shown is the observing window (the dot-dasked line; see equation (10.26)), which removes the small-scale power in the resulting map.

string-induced primary perturbations are not exactly Gaussian, the picture which we just revealed can be even stronger.

10.5 Conclusion

We proposed a new method which is capable of extracting all NG signals from an underlying CMB map and which is insensitive to the CLT. This method basically whitens the power spectrum of the field and therefore removes the Gaussian component. We showed that the whitened field can keep the intrinsic features of all NG signals, even if some of which may dominate the others in the amplitude. This method has been justified by testing it on some toy models, as well as a simulated string-induced CMB map as observed by an interferometer, with both the window effects and instrument noise included. With the whitened field, in principle one can further apply other statistics to, for example, locate the non-Gaussian scale or further quantify the non-Gaussianity. We also note that the generalisation of this method to a whole-sky treatment is straightforward. What one needs to do in this case is just to whiten the full-sky map with respect to its angular power spectrum in multipole space, as what we did for Fourier spcae here. Therefore, we expect this new method to be of assistance in investigating the non-Gaussianity of the CMB maps to be obtained in the near future.

Chapter 11

CMB Non-Gaussian Tests II—Inter-Scale Correlations

11.1 Introduction

As we have discussed, the CMBR is one of the most important cosmological measurements currently available. The future satellite missions such as MAP and PLANCK are expected to provide high-resolution, low-noise CMB data that will strongly constrain current theories of structure formation. One of the key differences between the competing theories is the Gaussian distribution of the density fluctuations predicted in most inflationary theories versus the generically non-Gaussian inhomogeneities in scenarios based on topological defects. For this reason, the detection of non-Gaussian features in the CMB is regarded as an important distinguishing test, especially in view of forthcoming full-sky observations. In the previous chapter, we saw a new method which can potentially extract the NG signals from an underlying CMB map. In this chapter, we continue to propose a new statistical scheme by which we can test the hypothesis that an underlying map is Gaussian.

Since all theories give only statistical predictions of the CMB, conclusions from one observation of the CMB sky are by necessity probabilistic. As reviewed in the previous chapter, many criteria for Gaussianity have been proposed in the literature and applied to the CMB data yielding results consistent with the Gaussian hypothesis. Although some claims of finding a non-Gaussian signal in COBE have been advanced recently [193], not all of them are equally persuasive, as discussed in [195]. Apart from possible data contamination by foregrounds and noise (which according to Ref. [194] was the source of the spurious non-Gaussian signal reported in Ref. [193]), the utility of any CMB statistic is fundamentally limited by cosmic variance, which is a consequence of having only one measurement of the CMB sky. From this point of view, some descriptors of a non-Gaussian signal have greater potential than others.

Cosmic variance manifests itself differently in real space and in Fourier space, where the CMB temperature fluctuations are decomposed into multipole modes a_{lm} corresponding to the spherical harmonics. The modes at smaller angular scales (higher l) are more precisely defined, essentially because information about smaller-scale features is gathered from more independent areas of the sky, and because all Fourier modes are independent if Gaussianity is assumed. In other words, cosmic variance in the Fourier space is shifted from small scales toward large scales. One would therefore expect statistics of Fourier modes at very small scales to be relatively free of cosmic variance. On the other hand, real space data is not uncorrelated and is equally affected by cosmic variance throughout all scales. Nevertheless, as we discussed in the previous chapter, the recently introduced wavelet-based statistics (e.g. [188, 189, 190]) can be regarded as occupying the middle ground between the real and the Fourier space descriptors.

This motivates us to consider statistics of Fourier modes for small scales, in the hope of obtaining a sensitive statistic. We intend to investigate inter-scale correlations in Fourier space as non-Gaussian indicators. In this framework, a non-Gaussian signal manifests itself as a non-zero correlation between Fourier modes a_{lm} of different scales l, and this extra correlation must be of third order or higher. The third-order correlator of the Fourier modes or the 'bispectrum' has already been considered in Ref. [192, 193] for modes at a fixed scale l. To obtain the cross-correlation between some chosen scales l and l', we may regard the set of mode values a_{lm}, $a_{l'm}$ at these scales as coming from a joint two-variable distribution and test that distribution for Gaussianity. A general way of testing a distribution for Gaussianity is by using cumulants (see e.g. [190]). For our analysis, we define and employ suitable multi-variate cumulants. Since the variance of a cumulant estimator is decreased when the number of points in the sample grows, we expect the statistic to be increasingly sensitive on smaller scales (larger l). An attractive property of Fourier space cumulants at small scales is their statistical independence: for a hypothetical Gaussian field, the cumulant estimators are approximately independent and themselves approximately Gaussian distributed, as we shall show generally for cumulants of any order. This significantly simplifies the likelihood analysis of one sample.

The outline of the chapter is as follows. In section 11.2, we define the particular cumulants that describe inter-scale correlations of the Fourier modes, and show how to estimate them from a single field realization. We derive the means and variances of these cumulant estimators assuming a Gaussian random field, as would be needed for a likelihood analysis. We also show that the error bars decrease for large l as expected. We then in section 11.3 propose a scheme for our NG test, based on the theory described in section 11.2. We present numerical results in section 11.4, including tests involving simple artificial NG signals, instrument noise, point sources, and the string-induced ISW effect. Brief conclusions about the viability of the method are given in section 11.5. To keep a comprehensive line of development, we have moved some detailed derivations relevant to the chapter to Appendices D and E.

11.2 Cumulants in Fourier space

11.2.1 Translation and rotation invariance

For simplicity, we consider random fields in flat space, although the results can be straightforwardly generalised to random fields on a sphere, such as the CMB temperature perturbation $\Delta T/T$. From the Fourier space viewpoint, a random field $f(x)$ is a collection of random variables (modes) $a_{\mathbf{k}}$ (c.f. equation (10.2)),

$$a_{\mathbf{k}} = \int x^2 f(\mathbf{x})\, e^{-i\mathbf{k}\cdot\mathbf{x}}. \tag{11.1}$$

Homogeneity of the random field (translation invariance in real space) constrains the joint distribution of $\{a_{\mathbf{k}}\}$, namely the moments must satisfy

$$\langle a_{\mathbf{k}_1} a_{\mathbf{k}_2} ... a_{\mathbf{k}_n}\rangle = 0, \quad \text{if } \mathbf{k}_1 + ... + \mathbf{k}_n \neq 0, \tag{11.2}$$

and isotropy means that for any rotation R in Fourier space,

$$\langle a_{\mathbf{k}_1} a_{\mathbf{k}_2} ... a_{\mathbf{k}_n}\rangle = \langle a_{R\mathbf{k}_1} a_{R\mathbf{k}_2} ... a_{R\mathbf{k}_n}\rangle . \tag{11.3}$$

We note that in the spherical case, both conditions are embodied in one isotropy condition. The task of testing the Gaussianity of a random field is now translated into checking whether all the $a_{\mathbf{k}}$ are jointly Gaussian distributed.

In the general case one can describe the joint distribution of all Fourier modes by a suitable generating functional of moments,

$$\begin{aligned} Z[j(\mathbf{k})] &\equiv \sum_{n=0}^{\infty} \int d\mathbf{k}_1 ... d\mathbf{k}_n \frac{j(\mathbf{k}_1)...j(\mathbf{k}_n)}{i^n n!} \langle a_{\mathbf{k}_1} a_{\mathbf{k}_2} ... a_{\mathbf{k}_n}\rangle \\ &= \left\langle \exp\left[-i\int j(\mathbf{k})\, a_{\mathbf{k}} d\mathbf{k}\right]\right\rangle , \end{aligned} \tag{11.4}$$

from which one recovers all the moments by functional differentiation,

$$\langle a_{\mathbf{k}_1} a_{\mathbf{k}_2} ... a_{\mathbf{k}_n}\rangle = i^n \frac{\delta^n}{\delta j(\mathbf{k}_1)...\delta j(\mathbf{k}_n)} Z[j(\mathbf{k})]_{j=0} . \tag{11.5}$$

Here the functional argument $j(\mathbf{k})$ is a complex-valued function of $\mathbf{k}$ satisfying $j(-\mathbf{k}) = j^*(\mathbf{k})$. It is convenient to define the general cumulants of the Fourier modes from the logarithm of the functional (11.4),

$$\tilde{C}^{(n)}(\mathbf{k}_1, ..., \mathbf{k}_n) \equiv i^n \frac{\delta^n}{\delta j(\mathbf{k}_1)...\delta j(\mathbf{k}_n)} \ln Z[j(\mathbf{k})]_{j=0} . \tag{11.6}$$

For example, the distribution of the modes of a Gaussian random field with power spectrum $P(k)$ is characterised by the generating functional

$$Z_{\mathrm{G}}[j(\mathbf{k})] = \exp\left[-\frac{1}{2}\int P(k)\, |j(\mathbf{k})|^2 d\mathbf{k}\right] , \tag{11.7}$$

and we see that cumulants $\tilde{C}^{(n)}$ of order $n \geq 3$ vanish, as usual. It is easy to see that the general cumulants $\tilde{C}^{(n)}$ satisfy the same homogeneity and isotropy conditions of equations (11.2)–(11.3) as do the moments. This is because equations (11.2)–(11.3) imply that the generating functional $Z[j(\mathbf{k})]$ is invariant under rotations $j(\mathbf{k}) \to j(R\mathbf{k})$ and translations $j(\mathbf{k}) \to e^{i\mathbf{k}\cdot\mathbf{a}} j(\mathbf{k})$ of its functional argument, and therefore $\ln Z[j(\mathbf{k})]$, the generating functional for cumulants, must also be rotation- and translation-invariant. This constrains some cumulants to identically vanish. The joint distribution of modes is Gaussian if all remaining cumulants of order $n \geq 3$ vanish as well.

11.2.2 Rings of fixed moduli

Now we turn to the statistics of $a_{\mathbf{k}}$ that can be relevant for data analysis. Since a measurement of one realization of the random field gives only one set of values $\{a_{\mathbf{k}}\}$, there is a limit to the kind of information about the joint distribution of $\{a_{\mathbf{k}}\}$ which can be extracted without a priori knowledge. For instance, we cannot efficiently test equations (11.2)–(11.3) for any particular choice of $\{\mathbf{k}_i\}$ because in each case we would have only one data point $a_{\mathbf{k}_1} a_{\mathbf{k}_2} ... a_{\mathbf{k}_n}$ to compare; an inhomogeneous, anisotropic random field that has the "wrong" distribution of just a few selected modes $a_{\mathbf{k}}$ cannot be distinguished from a homogeneous and isotropic field on the basis of one sample. Fortunately, we do not expect such an artificial random field to have physical relevance.

Similarly, we cannot test the hypothesis that, say, two particular modes $a_{\mathbf{k}_1}$ and $a_{\mathbf{k}_2}$ come from a joint Gaussian distribution since we measure only one realization of these modes. Clearly one can test a distribution for Gaussianity only if one has a large number of samples of that distribution. Therefore, to obtain any result at all concerning the Gaussianity of $a_{\mathbf{k}}$, we must assume that several of the modes $a_{\mathbf{k}}$ come from the same distribution. From the natural assumption of isotropy it follows that the modes $a_{\mathbf{k}}$ within a ring of fixed modulus $|\mathbf{k}| = k$ are identically distributed, and thus the ring $|\mathbf{k}| = k$ provides a sample that allows us to test their distribution for Gaussianity. This is the assumption behind the bispectrum test [192, 193]. Strictly speaking, a negative result of such a test would indicate either a non-Gaussian or an anisotropic distribution. However, most theories predict the isotropy of the CMB, and so, after any foregrounds are adequately dealt with, we are much less interested in detecting anisotropy or inhomogeneity than we are in finding non-Gaussian signals.

We are therefore motivated to assume homogeneity and isotropy and to regard the modes $a_{\mathbf{k}}$ at a fixed scale $|\mathbf{k}|$ as independent samples of the same joint distribution of modes. This assumption also allows us to investigate correlations between different scales, which is the main interest of the present work. Consider two rings of modes $a_{\mathbf{k}}$ at two fixed scales k_1 and k_2. The values $a_{\mathbf{k}}$ along the two rings can be regarded as independent realizations of the joint two-variable distribution, and we would like to test whether that distribution is Gaussian in two variables.

11.2.3 Cumulants of two rings

A standard way to test a sample for Gaussianity is to compute cumulants of various orders [190]: for a Gaussian distribution, the cumulants of order ≥ 3 vanish. As introduced in section 10.2.2, cumulants of a distribution of two variables (x, y) can be labelled by two indices, e.g. χ_{mn} has "dimension" $x^m y^n$, and an order $(m+n)$; the general definition of χ_{mn} is given by equations (E.3)–(E.4). We are interested in cumulants that describe cross-correlation between the two variables, for instance the nontrivial third-order cross-cumulant is

$$\chi_{12} = \left\langle xy^2 \right\rangle - \langle x \rangle \left\langle y^2 \right\rangle - 2 \langle xy \rangle \langle y \rangle + 2 \langle x \rangle \langle y \rangle^2 . \tag{11.8}$$

Similarly, one can define cumulants of a three-variable distribution (x, y, z), for example

$$\chi_{111} = \langle xyz \rangle - \langle xy \rangle \langle z \rangle - \langle xz \rangle \langle y \rangle - \langle yz \rangle \langle x \rangle + 2 \langle x \rangle \langle y \rangle \langle z \rangle . \tag{11.9}$$

Given a sample consisting of N points (x_i, y_i) of a two-variable distribution, one could estimate the cumulant χ_{12} directly by evaluating the sample averages required by equation (11.8). As shown in Appendix D, such estimators are generally not unbiased, and their expectation value for a Gaussian field is of order N^{-1} instead of zero, as given by equations (D.1)–(D.2) for one-variable and by equation (E.10) for two-variable cumulants. This presents no problem for likelihood analysis since this expectation value for a Gaussian field is known and can be simply subtracted. Variance of cumulant estimators is derived first for one-variable cumulants in Appendix D, and then for Fourier space cumulants in Appendix E. It is shown in equations (D.3) and (E.11) that the variance is always of order N^{-1}, and this is consistent with the result obtained numerically in Ref. [190] for one-variable cumulants. We see that since the number N of points in the sample is equal to the number of modes $a_{\mathbf{k}}$ at the chosen scale $|\mathbf{k}| = k$ in our case, the variance is indeed decreased for larger k and thus for smaller scales.

We could have directly applied the cumulant technique to the Fourier modes, taking $x \equiv a_{\mathbf{k}}$ and $y \equiv a_{\mathbf{k}'}$, but the homogeneity constraint equation (11.2) suggests that of all possible cross-cumulants, only the cumulants of modes $a_{\mathbf{k}_i}$ for which $\sum_i \mathbf{k}_i = 0$ need to be considered; all others can be regarded as "noise" resulting from accidental inhomogeneity of the given sample. For instance, the third-order cumulant $\chi_{111}(k_1, k_1', k_2)$ relating two chosen scales $k_1 = |\mathbf{k}_1| = |\mathbf{k}'_1|$ and $k_2 = |\mathbf{k}_2|$ should be estimated from the set of triples $\{a_{\mathbf{k}_1}, a_{\mathbf{k}_1'}, a_{\mathbf{k}_2}\}$ where $\mathbf{k}_1 + \mathbf{k}_1' + \mathbf{k}_2 = 0$. Obviously, this cross-cumulant $\chi_{111}(k_1, k_1', k_2)$ of the scales k_1 and k_2 can be nonzero only if $k_2 \leq 2k_1$. In the notation of equation (11.6), we have $\chi_{111} = \tilde{C}^{(3)}(\mathbf{k}_1, \mathbf{k}_1', \mathbf{k}_2)$. In Appendix E we derive general expressions for the variances of such cross-cumulants. The sample size N is determined by the number of modes in the smallest of the rings k_1 and k_2, and the variance of $\chi_{111}(k_1, k_1', k_2)$ is

$$\mathrm{var}\left[\chi_{111}(k_1, k_1', k_2)\right] \equiv \mathrm{var}\left[\chi_{111}(k_1, k_2)\right] = N^{-1} P^2(k_1) P(k_2). \tag{11.10}$$

The distribution is Gaussian if its cumulants estimated from the sample of N data points fall within their sample variance. We also show in Appendix E that cumulant estimators for a Gaussian field are always uncorrelated Gaussian random variates, up to terms of order N^{-2}. Therefore the likelihood analysis in our case consists of computing the chosen cumulant estimator (χ_{111}) for various pairs (k_1, k_2) of scales and normalising them to their theoretical Gaussian variances (11.10), where the power spectrum is estimated beforehand from the same map. The resulting set of normalized cumulant estimators (containing as many numbers as pairs of different scales (k_1, k_2) of interest) should be a set of normally distributed, independent random values, if the underlying map is Gaussian. We propose to test this by computing the first few cumulants (mean, variance, skewness and kurtosis) of that set, which is equivalent to joint estimation of Gaussianity of the inter-scale cumulants for all pairs of scales. A more comprehensive description of the scheme will be given in section 11.3.

11.2.4 Sensitivity to NG signals

With a generating functional $Z_{\mathrm{NG}}[j(\mathbf{k})]$ for the Fourier components of a non-Gaussian random field, we can mix it with a Gaussian background of some power spectrum $P_{\mathrm{G}}(k)$ described by a generating functional of equation (11.7). The new generating functional will be the product of the two and the cumulants will be the sums of the two sets of cumulants. Since the Gaussian field has vanishing higher-order cumulants, the only change will be in the power spectrum which becomes

$$P(k) = P_{\mathrm{NG}}(k) + P_{\mathrm{G}}(k). \tag{11.11}$$

Consider one cumulant estimator $\tilde{C}^{(n)}(k_1, ..., k_n)$ at certain selected set of scales k_i: since its variance is proportional to the appropriate powers of $P(k_i)$ as given by equation (E.13), while its expectation value is unchanged after adding the Gaussian background, the sensitivity of the cumulant estimator will be diminished by the factor

$$\alpha = \left\{ \prod_{i=1}^{n} \left[1 + \frac{P_{\mathrm{G}}(k_i)}{P_{\mathrm{NG}}(k_i)} \right] \right\}^{1/2}. \tag{11.12}$$

When $P(k_i) \approx P_{\mathrm{NG}}(k_i)$ for all k_i, we see $\alpha \approx 1$; α will depart significantly from unity if $P(k_i) \approx P_{\mathrm{G}}(k_i)$ for some k_i. Clearly the sensitivity of the criterion is unaffected on scales where the Gaussian background is negligible compared to the non-Gaussian signal, and will be proportionately decreased otherwise. Therefore, we expect the NG feature to show up in $\tilde{C}^{(n)}(k_1, ..., k_n)$ within a range of k where the NG signal dominates. This criterion is the same as that for the method we proposed in the previous chapter.

11.3 The scheme

Consider a CMB map which we would like to test for the hypothesis that it is Gaussian. Based on the theory described in the previous section, the procedure is as follows:

1. Fourier transform this map to obtain the modes $\{a_{\mathbf{k}}\}$ (this step is unnecessary for data obtained from interferometers).

2. Divide the whole Fourier domain $\mathbf{k}$ into rings of fixed moduli $k = |\mathbf{k}|$. One ring of radius k_1 is then picked up in turn with k_1 going from $0.1k_{\mathrm{max}}$ to k_{max}, where k_{max} is the maximum accessible wavenumber. With a chosen ring at each time, a pair of modes $a_{\mathbf{k}_1}$ and $a_{\mathbf{k}_1'}$ are then taken with

$$\mathbf{k}_1 \cdot \mathbf{k}_1' = {k_1}^2 \cos\theta, \tag{11.13}$$

where $|\mathbf{k}_1| = |\mathbf{k}_1'| = k_1$, and θ goes from 0 to π. For a chosen pair of (k_1, θ), we pick up a second ring of radius k_2 such that

$$k_2 = 2k_1 \cos(\theta/2). \tag{11.14}$$

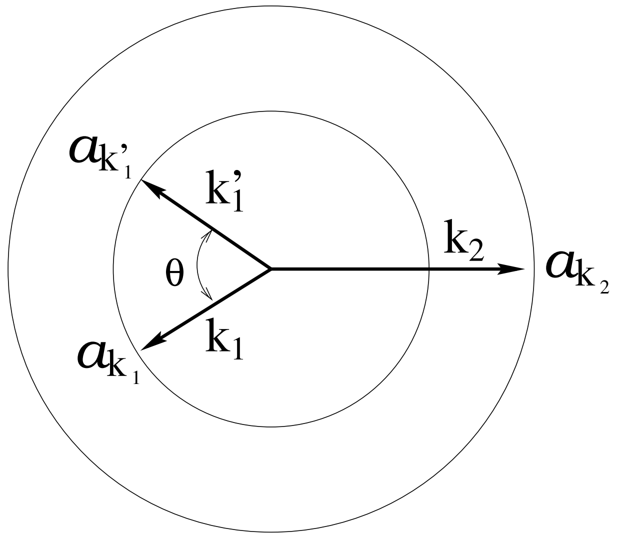

Figure 11.1: A configuration of $\{a_{\mathbf{k}_1}, a_{\mathbf{k}'_1}, a_{\mathbf{k}_2}\}$ with $\mathbf{k}_1 + \mathbf{k}'_1 + \mathbf{k}_2 = 0$, where $|\mathbf{k}_1| = |\mathbf{k}'_1| = k_1$ and $|\mathbf{k}_2| = k_2$ are the radii of two rings in Fourier space.

Then two modes on the first ring ($a_{\mathbf{k}_1}$ and $a_{\mathbf{k}'_1}$) and one mode on the second ring ($a_{\mathbf{k}_2}$) are then paired up to form a joint set of three variables $\{a_{\mathbf{k}_1}, a_{\mathbf{k}'_1}, a_{\mathbf{k}_2}\}$, with a configuration specified by $(k_1, \theta) \equiv (k_1, k_2)$ satisfying $\mathbf{k}_1 + \mathbf{k}'_1 + \mathbf{k}_2 = 0$. Figure 11.1 shows one example of such configurations.

3. Equation (11.9) is employed to calculate the estimator $\hat{\chi}_{111}(\mathbf{k}_1, \mathbf{k}'_1, \mathbf{k}_2) \equiv \hat{\chi}_{111}(k_1, k_2)$. We then normalise it to its standard deviation that assumes Gaussianity (i.e. the square root of equation (11.10)), resulting in the estimator

$$\hat{\bar{\chi}}_{111}(k_1, k_2) = \frac{\hat{\chi}_{111}(k_1, k_2)}{\mathrm{var}\left[\hat{\chi}_{111}\left(k_1, k_2\right)\right]^{1/2}} = \frac{N^{1/2}\hat{\chi}_{111}(k_1, k_2)}{P(k_1)P^{1/2}(k_2)}. \tag{11.15}$$

4. If the underlying joint distribution $\{a_{\mathbf{k}_1}, a_{\mathbf{k}'_1}, a_{\mathbf{k}_2}\}$ is Gaussian, then in the domain (k_1, k_2), the $\bar{\chi}_{111}(k_1, k_2)$ should be a Gaussian variable with mean zero and variance one. Therefore, we further calculate the first few normalised cumulants of $\hat{\bar{\chi}}_{111}$ in the domain (k_1, k_2):

$$\begin{aligned}
\mathrm{mean}\left[\hat{\bar{\chi}}_{111}\right] = \quad \hat{\varepsilon}_1 \quad &= \left\langle \hat{\bar{\chi}}_{111}(k_1, k_2) \right\rangle, && (11.16)\\
\mathrm{var}\left[\hat{\bar{\chi}}_{111}\right] = \quad \hat{\varepsilon}_2 \quad &= \left\langle \left[\hat{\bar{\chi}}_{111}(k_1, k_2) - \hat{\varepsilon}_1\right]^2 \right\rangle, && (11.17)\\
\mathrm{ske}\left[\hat{\bar{\chi}}_{111}\right] = \quad \hat{\varepsilon}_3 \quad &= \left\langle \left[\hat{\bar{\chi}}_{111}(k_1, k_2) - \hat{\varepsilon}_1\right]^3 \right\rangle \hat{\varepsilon}_2^{-3/2}, && (11.18)\\
\mathrm{kur}\left[\hat{\bar{\chi}}_{111}\right] = \quad \hat{\varepsilon}_4 \quad &= \left\langle \left[\hat{\bar{\chi}}_{111}(k_1, k_2) - \hat{\varepsilon}_1\right]^4 \right\rangle \hat{\varepsilon}_2^{-2} - 3, && (11.19)
\end{aligned}$$

where the ensemble averages are taken over the domain (k_1, k_2). We shall denote the number of samples in the domain (k_1, k_2) as N_χ.

5. If the joint distribution of $\{a_{\mathbf{k}_1}, a_{\mathbf{k}'_1}, a_{\mathbf{k}_2}\}$ is Gaussian, then ε_i should have mean zero, except ε_2 having mean one. They also have variances

$$\mathrm{var}\left[\varepsilon_i\right] = \frac{i!}{N_\chi}. \tag{11.20}$$

Therefore, if the estimators $\hat{\varepsilon}_i$ of the map depart from their means much further than $\mathrm{var}\left[\varepsilon_i\right]$, we can reject the hypothesis that the map is Gaussian. Thus it would be convenient to normalise these $\hat{\varepsilon}_i$ to their variances:

$$\hat{\omega}_1 = \sqrt{N_\chi}\,\hat{\varepsilon}_1, \ \ \hat{\omega}_2 = \sqrt{\frac{N_\chi}{2}}\left(\hat{\varepsilon}_2 - 1\right), \ \ \hat{\omega}_3 = \sqrt{\frac{N_\chi}{6}}\,\hat{\varepsilon}_3, \ \ \hat{\omega}_4 = \sqrt{\frac{N_\chi}{24}}\,\hat{\varepsilon}_4. \tag{11.21}$$

With this procedure, we can evaluate the estimators $\hat{\omega}_i$ from the CMB map under investigation. If these $\hat{\omega}_i$ depart significantly from unity, we can reject the hypothesis that the map is Gaussian.

11.4 Numerical tests

11.4.1 The Gaussian background and instrument noise

As mentioned in the previous chapter, even if the underlying mechanism for producing the CMB anisotropies is non-Gaussian, there are still two main possible Gaussian components in the observed CMB—first, a Gaussian background due to the CLT and, second, instrument noise. Although the latter is not necessarily Gaussian on general grounds, we shall assume it to be Gaussian with a white-noise power spectrum for simplicity. As a first step for testing the method described in the previous section, we shall compute the estimators $\hat{\omega}_i$ for these two Gaussian components and their combination, to see if they can pass the hypothesis that they are Gaussian, i.e. to see if the normalised cumulant estimators $\hat{\omega}_i$ lie within the range $(-1,1)$ for these cases.

First, the Gaussian background is generated using the method described in section 10.4.1, and the field size is $(5°)^2$ with a resolution of 256^2. The instrument noise is simulated in a similar manner, but with a white-noise power spectrum. We show these two maps on the top of Figure 11.2. These two maps are then mixed, with a noise fraction

$$f_{\rm noi} = \frac{\sigma_{\rm noi}}{(\sigma_{\rm G} + \sigma_{\rm noi})}, \tag{11.22}$$

where $\sigma_{\rm G}$ and $\sigma_{\rm noi}$ denote the RMS amplitudes of the Gaussian background and the white noise respectively. We then compute the estimators $\hat{\omega}_i$ as defined in equation (11.21), for the value of $f_{\rm noi}$ going from zero to unity. The results are plotted at the bottom of Figure 11.2. We first note that for $f_{\rm noi=0}$ and $f_{\rm noi} = 1$, the results correspond to those for a pure Gaussian background and pure white noise respectively. In both cases, all $\hat{\omega}_i$ lie within the range between -1 and 1, meaning that the hypothesis about Gaussianity holds in both cases. For the range $0 < f_{\rm noi} < 1$, although some $\hat{\omega}_i$ go outside the range $(-1,1)$, they are still well inside the two-sigma region $(-2,2)$ (the 95% confidence region) so that the hypothesis still holds. This verifies the reliability of our numerical procedure, as well as the theory about the use of ω_i. For reference, we also plot the power spectra of the Gaussian background and the instrument noise in Figure 11.3, i.e. the solid and dotted lines respectively.

In addition to this test, we also generate 10000 realisations of the Gaussian background, and then compute $\hat{\omega}_i$ for each realisation. This allows us to estimate the variances of ω_i, and we find that they are indeed unity for $i = 1,2,3$ and 4 as they should be, since we have previously proved that ω_i are Gaussian variables of mean zero and variance one. We have also numerically verified that this result is independent of the underlying power spectrum of the Gaussian background, as it should be on the theoretical grounds. This is one of the advantages of employing the ω_i for testing Gaussianity, because in this case one does not need to compute the so-called 'equivalent Gaussian realisations' (e.g. [188]) from the map under test in order to obtain information for likelihood analysis. Therefore, in the likelihood analyses hereafter, we shall simply shade the one-sigma area $(-1,1)$ of ω_i in all the relevant plots.

11.4.2 Point sources

As explained in the previous chapter, one important challenge for all CMB data analysis is dealing with the presence of point sources. This component is an 'unwanted' non-Gaussian signal, which normally causes most CMB NG tests to fail. For this reason it is important to see how this component will contribute to our estimators $\hat{\omega}_i$.

We first generate five point sources using equation (10.21), with $r = 0.01$. It is shown in the upper part of Figure 11.4. This map is then superimposed onto the Gaussian background used in the previous subsection, with a fraction of the point sources

$$f_{\rm pnt} = \frac{\sigma_{\rm pnt}}{(\sigma_{\rm G} + \sigma_{\rm pnt})}, \tag{11.23}$$

where $\sigma_{\rm pnt}$ is the RMS amplitude of the point signal. We then compute the estimators $\hat{\omega}_i(f_{\rm pnt})$ as in the previous case. The results are plotted in the lower part of Figure 11.4.

As one can see, when the fraction of the point signal is weak ($f_{\rm pnt} \lesssim 0.1$), their non-Gaussianity does not show up. Once this signal becomes stronger ($f_{\rm pnt} \gtrsim 0.1$), $\hat{\omega}_1$ starts departing from the one-sigma region while the other $\hat{\omega}_i$ also follow, though at larger $f_{\rm pnt}$. With $f_{\rm pnt} \lesssim 0.2$, one can use $\hat{\omega}_1$ to reject the hypothesis of Gaussianity at a confidence level of more than 99%, because the dotted line goes beyond the three-sigma region $(-3,3)$.

We also plot the power spectrum of the point sources as the dot-dashed line in Figure 11.3. By comparing this line to the solid line (the power spectrum of the Gaussian background), we confirm that for $f_{\rm pnt} \lesssim 0.1$, the point signal never comes to dominate on all scales and so unsurprisingly passes the hypothesis of Gaussianity using ω_i. Once $f_{\rm pnt} \gtrsim 0.2$, the power spectrum of the point signal starts to dominate at $l \sim 2000$, at which stage it fails the hypothesis. Thus, we conclude that our estimators $\hat{\omega}_i$ are sensitive to point sources. For this method to succeed in real CMB data, therefore, it will be necessary to first clean the point sources (e.g. [210]),

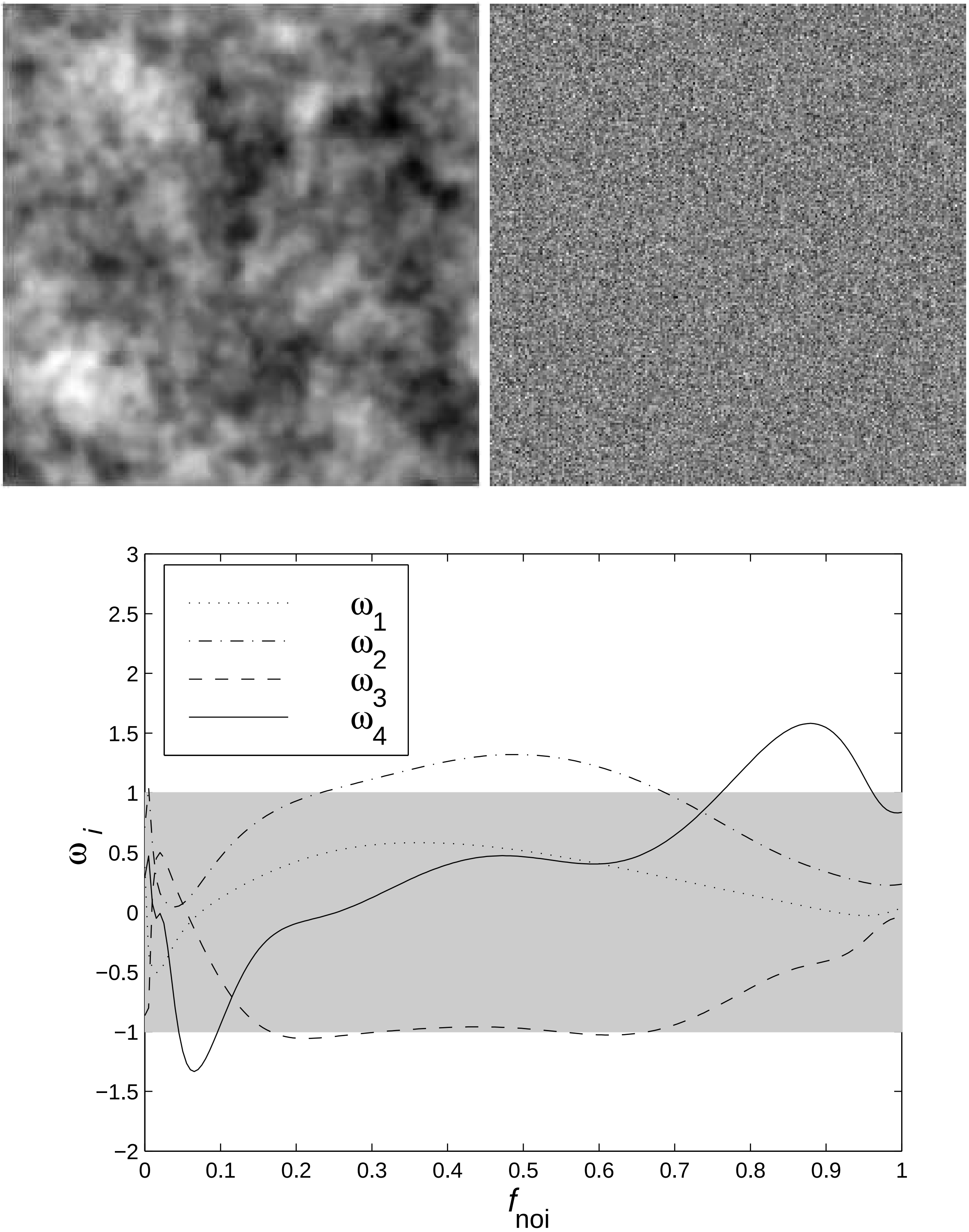

Figure 11.2: On the top are maps of a Gaussian background (left) and white noise (right). The field size is $(5°)^2$. These two components are then mixed, with a noise fraction $f_{\rm noi} = \sigma_{\rm noi}/(\sigma_{\rm G} + \sigma_{\rm noi})$. At the bottom is the resulting estimators $\hat{\omega}_i$ of the mixed map (see equation (11.21)), as a function of $f_{\rm noi}$. The shaded area is $\omega_i = (-1, 1)$.

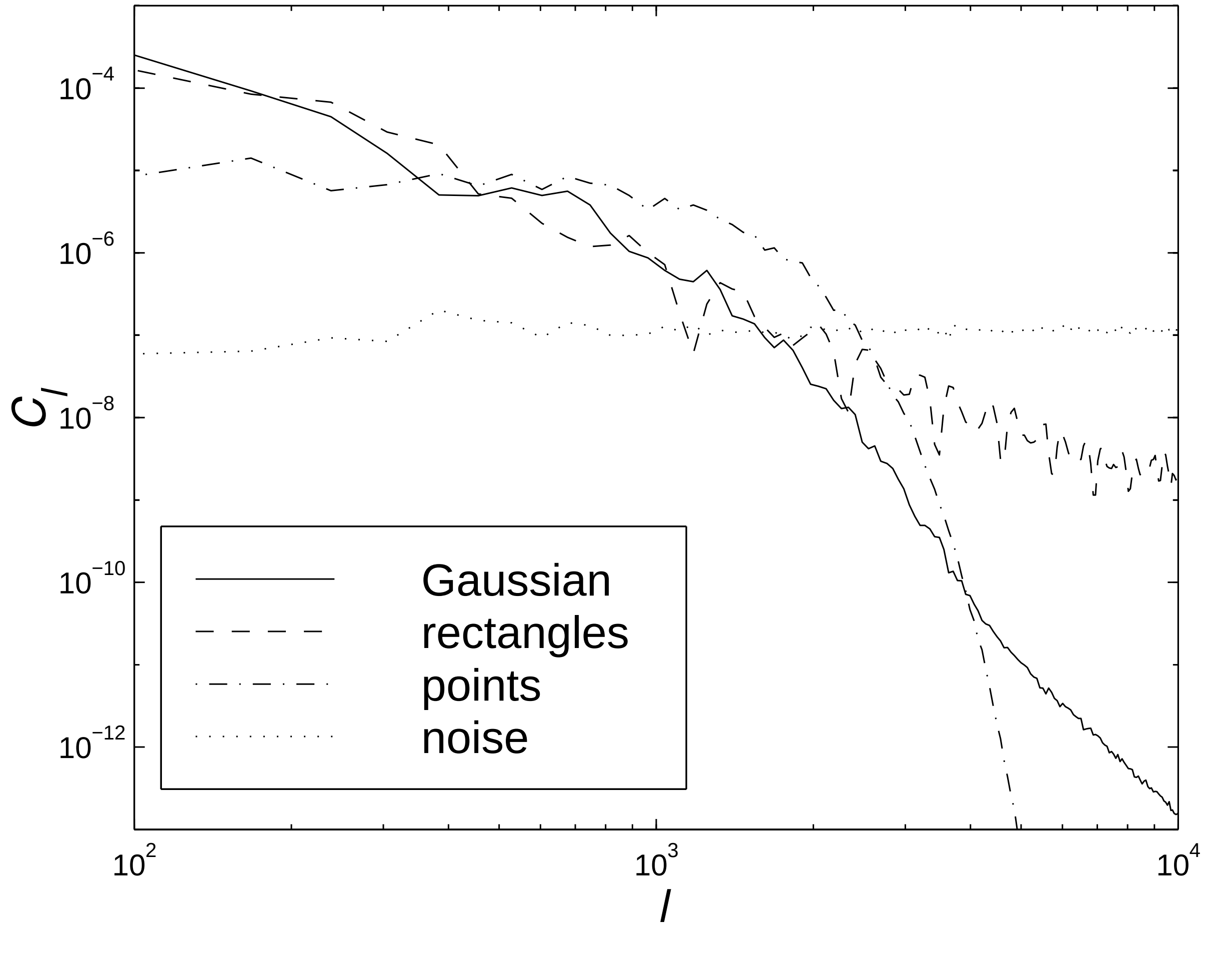

Figure 11.3: Power spectra of various components presented in Figures 11.2, 11.4, 11.5, 11.6. Their standard deviations are all unity.

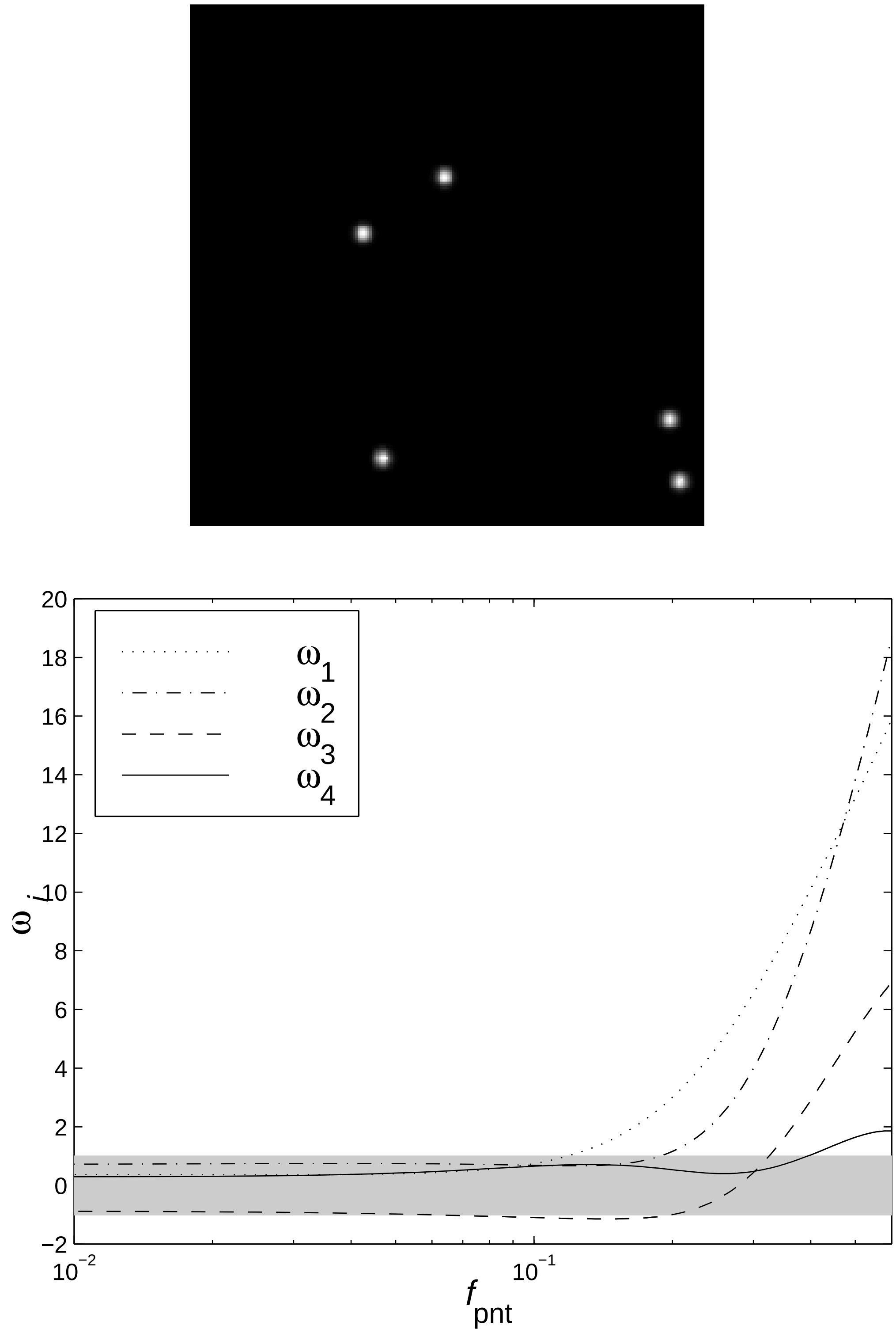

Figure 11.4: On the top is a map of five point sources. It is then mixed with the Gaussian background shown in Figure 11.2, with a fraction of the point sources $f_{\text{pnt}} = \sigma_{\text{pnt}}/(\sigma_{\text{G}} + \sigma_{\text{pnt}})$. At the bottom is the resulting estimators $\hat{\omega}_i$ of the mixed map (see equation (11.21)), as a function of f_{pnt}. The shaded area is $\omega_i = (-1, 1)$.

and then compute the estimators $\hat{\omega}_i$. In the following analyses, we shall ignore the point sources by assuming that they have been removed beforehand.

11.4.3 Test on filled rectangles

To test the sensitivity of our method to certain types of non-Gaussianity, we first try a simple model using filled rectangles. In a $(5°)^2$ field with a resolution of 256^2, We generate five filled rectangles, each of which has a size of 32×48 grid spacings. These rectangles are similar to the component (b) in section 10.4.4, and are shown in the upper part of Figure 11.5. This map is then mixed with the Gaussian background used in Figure 11.2, with a fraction of the rectangle map $f_{\rm rect}$ defined as usual. We then compute the estimators $\hat{\omega}_i(f_{\rm rect})$ as before, and the results are plotted at the bottom of Figure 11.5.

As we can see, one can use $\hat{\omega}_4$ to reject the hypothesis of Gaussianity at more than a 95% confidence level when $f_{\rm rect} \gtrsim 0.03$. This means that NG features of the rectangles shows up once their RMS amplitude is larger than a few percent of the total amplitude. Referring to the power spectrum of the rectangles in Figure 11.3 (the dashed line), we see that once $f_{\rm rect} \lesssim 0.01$, the power of the rectangles never comes to dominate when compared with the power of the Gaussian background (the solid line). This explains why NG features of the rectangles are suspended in $\hat{\omega}_i$ when $f_{\rm rect} \lesssim 0.03$ as seen previously.

We also test the sensitivity of our method to the rectangles in the presence of noise. In this case we mix three components: the rectangles, the Gaussian background, and the white noise. The last two components are exactly the same as the two we used in section 11.4.1 (see Figure 11.2). The noise fraction is fixed as $f_{\rm noi} = 0.05$, so the rectangles have a fraction:

$$f_{\rm rect} = \frac{\sigma_{\rm rect}}{(\sigma_{\rm rect} + \sigma_{\rm G} + \sigma_{\rm noi})} = 0.95 - f_{\rm G}\,. \tag{11.24}$$

The $\hat{\omega}_i$ are then computed as a function of $f_{\rm rect}$, and the results are presented in Figure 11.6. This time we see that the NG feature of the rectangles does not show up until a higher $f_{\rm rect}$. For example, one uses $\hat{\omega}_1$ to reject the hypothesis of Gaussianity at more then 95% confidence level only when $f_{\rm rect} \gtrsim 0.3$. Referring to the power spectra of the rectangles, the Gaussian background, and the instrument noise in Figure 11.3, one finds that when $f_{\rm rect} \lesssim 0.3$ with $f_{\rm noi} = 0.05$, the power of the rectangles is dominated by the power of the Gaussian components (the Gaussian background and the noise); when $f_{\rm rect}$ is larger than 0.3, its power starts to dominate at $l \sim 3000$. This verifies again that for $\hat{\omega}_i$ to succeed in detecting a NG signal, the power of the the NG component needs to dominate within at least a certain range of the accessible l. We notice that this criterion is the same as that for the method we proposed in the previous chapter. Thus we see that noise may reduce the sensitivity of our method for detecting NG signals, depending on whether the noise dominates the NG signal on scales where it originally dominated.

11.4.4 Detecting cosmic strings

Finally, we test our method against the string-induced CMB map. We first use the toy model mentioned in the previous chapter to simulate the string-induced ISW effect, which is on the top-left of Figure 11.7. The dynamic range is 40 in conformal time starting from last scattering, and the angular size is $(2°)^2$ with a resolution of 256^2. A Gaussian background is also simulated using our usual method (see section 10.4.1), with a power spectrum obtained from CMBFAST [85]. These two maps are then linearly summed, with a string fraction $f_{\rm str}$ defined as usual. The estimators $\hat{\omega}_i$ are then calculated as a function of $f_{\rm str}$, and the results are shown at the bottom in Figure 11.7.

From the dot dashed line in the figure ($\hat{\omega}_2$), one can reject the hypothesis of Gaussianity at more than a 95% confidence level when $f_{\rm str} \gtrsim 0.01$. The power spectrum of the Gaussian background is shown as the solid line in Figure 11.3, while that of the string-induced ISW effect is almost the same as the dashed line (the power spectrum of rectangles) in the same plot, though with a smoother shape. Again, if we compare these two lines carefully, we find that when $f_{\rm str} \lesssim 0.01$, the string-induced perturbations are dominated by the Gaussian background on all scales; when $f_{\rm str} \gtrsim 0.01$, the string-induced perturbations start to dominate at $l \sim 10^4$. This result is consistent with our previous argument about the sensitivity of the estimators $\hat{\omega}_i$.

In the presence of instrument noise, this sensitivity will be reduced if the noise dominates the string-induced perturbations on scales where they originally dominated. In other words, the presence of the noise will contribute extra power to the Gaussian component of the underlying map, so as to raise the threshold in power beyond which the NG signal can dominate. This argument is again verified in Figure 11.8, where we mix three components: the string-induced ISW effect, the Gaussian background, and the white noise whose strength is 5% of the total RMS amplitude. This gives

$$f_{\rm str} = \frac{\sigma_{\rm str}}{(+\sigma_{\rm str} + \sigma_{\rm G} + \sigma_{\rm noi})} = 0.95 - f_{\rm G}\,. \tag{11.25}$$

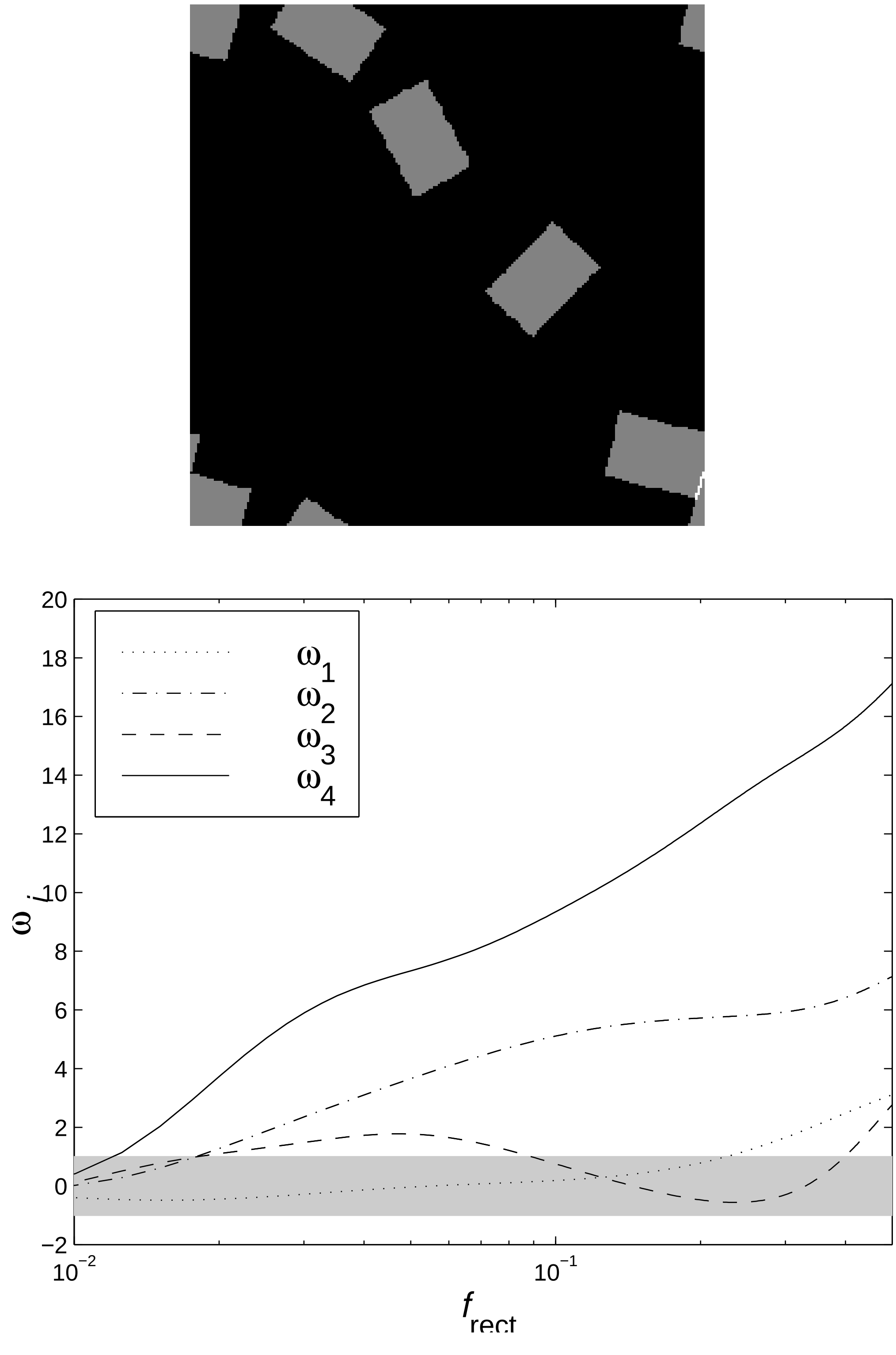

Figure 11.5: Above is a map of five filled rectangles. It is then mixed with the Gaussian background shown in Figure 11.2, with a fraction of the rectangles $f_{\rm rect} = \sigma_{\rm rect}/(\sigma_{\rm G}+\sigma_{\rm rect})$. At the bottom is the resulting estimators $\hat{\omega}_i$ of the mixed map (see equation (11.21)), as a function of $f_{\rm rect}$. The shaded area is $\omega_i = (-1, 1)$.

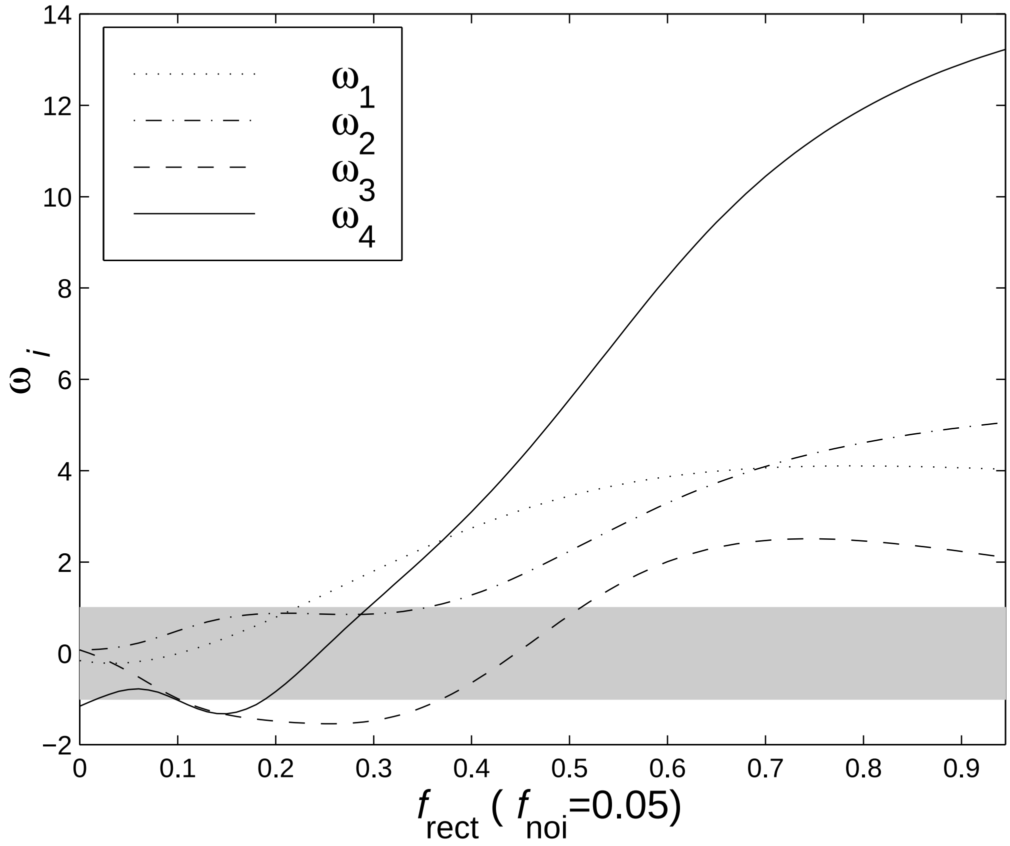

Figure 11.6: Same as Figure 11.5, but in the presence of 5% noise (shown in Figure 11.2), i.e. $f_{\rm noi} = \sigma_{\rm noi}/(\sigma_{\rm G} + \sigma_{\rm rect} + \sigma_{\rm noi}) = 0.05$.

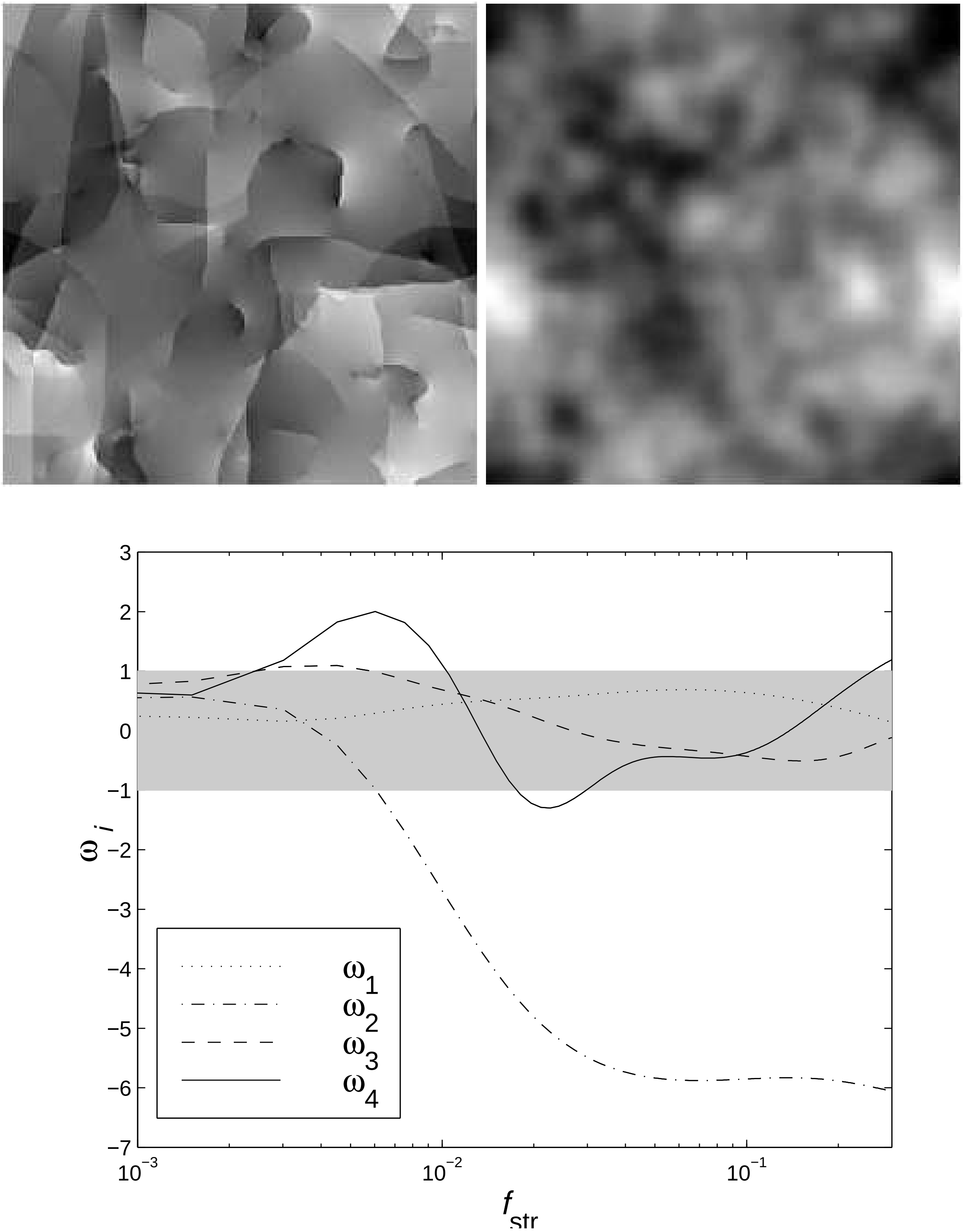

Figure 11.7: On the top are maps of a string-induced ISW effect (left), and a Gaussian background (right). The field size is $(2^\circ)^2$. These two components are then mixed, with a fraction of the string component $f_{\rm str} = \sigma_{\rm str}/(\sigma_{\rm G} + \sigma_{\rm str})$. At the bottom is the resulting estimators $\hat{\omega}_i$ of the mixed map (see equation (11.21)), as a function of $f_{\rm str}$. The shaded area is $\omega_i = (-1, 1)$.

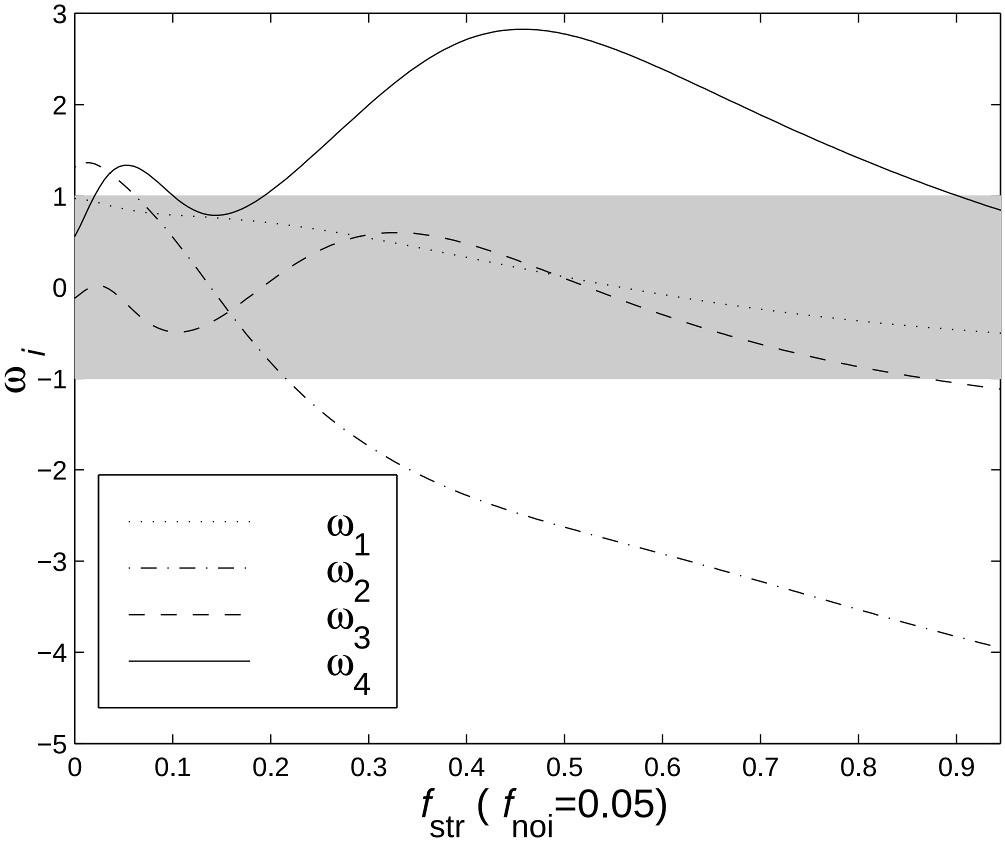

Figure 11.8: Same as Figure 11.7, but in the presence of 5% noise (shown in Figure 11.2), i.e. $f_{\rm noi} = \sigma_{\rm noi}/(\sigma_{\rm G} + \sigma_{\rm rect} + \sigma_{\rm noi}) = 0.05$.

As one can see, the hypothesis of Gaussianity is rejected at more than a 95% confidence level when $f_{\rm str} \gtrsim 0.3$, using either $\hat{\omega}_2$ or $\hat{\omega}_4$ (although $\hat{\omega}_4$ falls back into the 95% confidence region again when $f_{\rm str} \gtrsim 0.7$). This threshold $f_{\rm str} \approx 0.3$ can be again verified by comparing the dashed, solid, and dotted lines in Figure 11.3.

By comparing results for strings and rectangles (see section 11.4.3), we find that they have the same threshold for rejecting the hypothesis of Gaussianity, i.e. the thresholds are $f_{\rm rect} \sim f_{\rm str} \sim 0.01$ without noise, and $f_{\rm rect} \sim f_{\rm str} \sim 0.3$ with 5% noise. This is because strings and rectangles have almost identical power spectra, so that they come to dominate the Gaussian component above the same threshold and at the same scale.

By comparing Figures 11.7 and 11.8, with Figures 11.4, 11.5 and 11.6, we find that only string-induced perturbations result in a significantly negative estimator $\hat{\omega}_2$ which can be used to reject the hypothesis of Gaussianity, while all other NG components produce only positive $\hat{\omega}_i$ outside the 95% confidence region (the two-sigma region). This provides a potential discriminator in distinguishing string-induced NG features from other NG signals.

11.5 Conclusion

In this chapter, we designed a new class of statistics ω_i—the normalised inter-scale cumulants in Fourier space—to test an underlying CMB map for the hypothesis of Gaussianity. We provided carefully constructed theoretical foundations on which these new statistics are based, and numerically verified for $i = 1$–4 that they are sensitive to NG signals. These statistics are also sensitive to point sources. Therefore in order to apply them to detect certain types of NG signals like those resulting from defects, one needs to subtract potential point sources on scales where they dominate the NG signals which we would like to detect. The instrument noise was also shown to be capable of reducing the sensitivity of our new method, due to its extra contribution to the Gaussian component on small scales, where we expect the defect-induced perturbations to dominate. Nevertheless, as we theoretically proved and numerically verified, our new statistics are capable of detecting NG signals as long as the NG components dominate within a certain range of accessible scales. We also showed that the NG feature of the string-induced ISW effect appears different from others (e.g. points and rectangles) in the estimators $\hat{\omega}_i$: the string-seeded perturbations induce significantly negative $\hat{\omega}_2$. This is a potential discriminator in distinguishing string-induced perturbations from others. In the near future, we intend to apply this method to characterise the NG features resulting from different defect models in more realistic contexts, which should enhance our ability to detect cosmic defects.

To sum up, while an accurate estimate of defect-induced power spectra is difficult and computationally expensive to obtain, the methods for non-Gaussian test proposed in the previous and the current chapters could provide another observational window, through which to easily justify or disprove defect models. Even if defects are unimportant as the underlying mechanism for structure formation, we can still apply these methods to detect their existence, though at a lower energy level. With such techniques at hand, we should be ready to draw some robust conclusion about defects when high-accuracy and high-resolution observations are available in the near future.

OUTLOOK

It has been two decades since inflation and topological defects were proposed in the early eighties. While the development of inflationary models has been shifting from mathematical modelling and making cosmological predictions towards gaining an insight into its origin, phenomenological studies for defects have yet to be completed. It is the purpose of this dissertation to investigate the cosmological consequences of topological defects, in particular cosmic strings, which have proved to be more difficult to study due to their complicated non-linear behaviour.

In Part I, we began with the foundations of the standard cosmology, together with its successes and problems. We also described the basics of both inflationary and defect models, especially cosmic strings, and the observational methods in testing these models. These methods included the calculation of the power spectra for the LSS and CMB, and non-Gaussian tests in both. These are the main theme which we have been following in this dissertation.

In Part II, we presented a formalism in the synchronous gauge, which is suitable for studying cosmological perturbations with causal seeds, even in cosmologies with curvature or a cosmological constant. The accuracy of this formalism was justified in the context of the standard CDM model. Its analytic solutions for matter perturbations in the presence of causal sources were also given, and numerically verified to high accuracy. With these solutions and the use of the pseudo-stress-energy tensor, we further addressed the compensation mechanism in a mathematically and physically explicit way. Although in the literature, there have been detailed treatments of theories with causal seeds, the formalism and its analytic solutions presented here provide not only a physically transparent way for understanding the evolution of matter perturbations, but also a computationally economical method which is particularly pertinent when one needs to investigate the phase information of the resulting cosmological perturbations. Following the same line of development, we also have been working on obtaining the analytic solutions for radiation perturbations using the current formalism. This result will be useful in computing the full-sky CMB anisotropies seeded by topological defects.

In Part III, we investigated the LSS seeded by cosmic strings. This includes an opening chapter with detailed studies of string properties relevant for LSS. String-induced CDM power spectra were then calculated, with particular emphasis on the effects of HDM, curvature, a cosmological constant, radiation, loops, and baryons. Although the uncertainties in the present work remain to be further improved, mainly for the compensation mechanism and a more detailed treatment of loop effects, the picture revealed here has helped advance our understanding of cosmic string models. We then further investigated the phase information of the resulting string-seeded matter perturbations. We first used the PDF, skewness, kurtosis and genus to identify the NG features of string-induced matter perturbations. We found the non-Gaussian scale to be $\lesssim 1.5(\Omega_0 h^2)^{-1}\mathrm{Mpc}$, which is about the cluster scale. Above this scale, perturbations were shown to be very close to Gaussian, due largely to the CLT. A semi-analytic model was also proposed to verify the numerically obtained NG scale. In addition, we estimated the cluster abundance in the context of cosmic strings, so as to calculate the resulting σ_8. This result was shown to be consistent with the COBE normalisation, and is slightly smaller than that obtained in the standard CDM model. This is due to the slightly larger tail in the PDF of string-induced matter perturbations. We note that, although the underlying string-induced matter spectrum still carries considerable uncertainties, by investigating the phase information we still revealed several important qualitative features of matter perturbations seeded by cosmic strings. In the near future, we plan to improve the uncertainties in the current work, so as to obtain more robust quantitative results.

We proceeded to study the CMB in Part IV. The string-induced CMB maps were first computed, with considerations given to both the primary anisotropies and the ISW effect. Some preliminary analysis regarding their power spectra were also presented. Again, the uncertainties in the current calculations are still considerable, mainly resulting from the limitation of the simulation size, as well as the treatment for energy conservation. This programme of CMB simulations is currently underway, but unfortunately was not sufficiently complete to provide firm results. Nevertheless, the preliminary results presented here still provide a rough idea of the small-scale power-law decay of the spectrum for the string-induced ISW effect ($l(l+1)C_l \propto l^{-1}$), and the maps simulated here can potentially provide good samples for identifying the non-Gaussian features of cosmic strings. We also note that although the formalism used here is computationally economical for defect models, in order to make robust predictions for the CMB perturbations they induce, one should eventually use a Boltzmann code (e.g. CMBFAST [85]).

With some difficulties in obtaining quantitative predictions for string-induced cosmological perturbations, we then proposed two methods for NG tests in the CMB: one being capable of extracting NG signals from a strong Gaussian background, and the other being sensitive to scale-scale correlations which are generically present in defect models. These methods are shown to be reliable for the numerical tests carried out to date, not only for detecting cosmic strings but also for identifying other types of non-Gaussianity. Therefore, we expect these methods to be of great assistance in testing the Gaussianity of high-accuracy and high-resolution CMB data that will be soon obtained. This provides us with another independent way—apart from the calculation of the power spectra for LSS and CMB—to understand the current status of our universe and thus to test theories of structure formation.

With planned observations and further improvements in calculating the power spectra of string-induced cosmological perturbations, one should be able to either verify or rule out string models as the underlying mechanism for structure formation in the near future. However, even if cosmic strings, or more generally cosmological defects, are shown to be unimportant for forming the structures of the universe, the methods for CMB NG tests which we proposed here can still provide a way to constrain the existence of defects, though at significantly lower energies. We conclude, therefore, that although future observations with higher accuracy may favour one of the competing models for structure formation, this will not necessarily be catastrophic for either inflation or defects, which both have a much broader and generic role in cosmology and fundamental physics. Instead, combined with such observations, we hope the work presented in this dissertation will help clarify the respective roles of these theories, and thus advance our understanding of the origin of the universe.

APPENDICES

Appendix A

Abbreviations

AIT	Hammer-Aitoff
AS	Allen-Shellard (cosmic string simulation)
BB	Bennett-Bouchet (cosmic string simulation)
BBN	Big Bang Nucleosynthesis
CAR	Cartesian
CDM	Cold Dark Matter
CIC	Cloud in Cell
CLT	Central Limit Theorem
CMB	Cosmic Microwave Background
CMBR	Cosmic Microwave Background Radiation
COBE	Cosmic Background Explorer
DMR	Differential Microwave Radiometer
ECP	Equidistant Cylindrical Projection
FFT	Fast Fourier Transform
FRW	Friedmann-Robertson-Walker
FWHM	Full Width of Half Maximum

GLS	Sanson-Flamsteed, Sinusoidal
GUT	Grand Unified Theory
HDM	Hot Dark Matter
HSRA	Hubble-Slow-Roll Approximation
ISW	Integrated Sache-Wolf
IUSG	Initially Unperturbed Synchronous Gauge
ΛCDM	Cold Dark Matter with a Cosmological Constant
LSS	Large-Scale Structure
MT	Multipole Transform (spherical harmonic transform)
NG	Non-Gaussian
OCDM	Cold Dark Matter in Open Universe
PDF	Probability Distribution Function
PSRA	Potential-Slow-Roll Approximation
RMS	Root Mean Square
$\mathcal{S}$CDM	String with Cold Dark Matter
SFL	Small Field/Angle Limit
$\mathcal{S}$HDM	String with Hot Dark Matter
$\mathcal{S}\Lambda$CDM	String with Cold Dark Matter in Λ-Cosmologies
SN Ia	Supernova Type Ia
$\mathcal{S}$OCDM	String with Cold Dark Matter in Open Cosmologies
UETC	Unequal Time Correlator
ZEA	Zenithal Equal-Area

Appendix B

Spherical Harmonic Transform invoking FFT

We derive simple algorithms making use of the efficiency of the Fast Fourier Transform (FFT) to implement/invert multipole transforms, i.e. the spherical harmonic decomposition and its inverse. In the context of CMB anisotropy analysis, these algorithms are useful for azimuthal pixelisation schemes, where pixels lie on latitudinal circles and divide each evenly in longitude, e.g. the Equidistant Cylindrical Projection (ECP) (see [211] for a review), and the Igloo Pixelisation [212]. Algorithms for the spherical harmonic decomposition and its inverse will be derived separately in the following two sections. The former can be used to calculate the power spectrum of CMB anisotropies, the so-called C_l curve, while the latter can be employed to generate simulated full-sky CMB maps from given multipole coefficients, or from given C_l's in Gaussian theories.

B.1 Multipole transform and C_l calculation

The CMB anisotropies observed at a position $\mathbf{x}$ along a line of sight $(\theta,\ \phi)$ in spherical coordinates can be expanded in spherical harmonics (the multipole transform):

$$\frac{\Delta T}{T}(\mathbf{x};\theta,\phi) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l} a_{lm}(\mathbf{x}) Y_{lm}(\theta,\phi), \tag{B.1}$$

where

$$a_{lm}(\mathbf{x}) = \int d\Omega \frac{\Delta T}{T}(\mathbf{x};\theta,\phi) Y^*_{lm}(\theta,\phi), \tag{B.2}$$

with $d\Omega \equiv \sin\theta d\theta d\phi$. The power spectrum is then defined and estimated at a given $\mathbf{x}$, respectively, as

$$C_l = \left\langle |a_{lm}(\mathbf{x})|^2 \right\rangle_l \approx \frac{1}{2l+1}\sum_{m=-l}^{l} |a_{lm}(\mathbf{x})|^2. \tag{B.3}$$

To avoid confusion, we shall drop the $\mathbf{x}$ dependence in all relevant quantities hereafter. We know that the spherical harmonic can be decomposed into a θ-dependent part and a ϕ-dependent part as

$$Y_{lm}(\theta,\phi) = \lambda_l^m(\cos\theta)\exp(im\phi), \tag{B.4}$$

where $\lambda_l^m(\cos\theta)$ is real reflecting the oscillation amplitude of $Y_{lm}(\theta,\phi)$ at each θ, hence we can rewrite equation (B.2) as

$$a_{lm} = \int_0^{\pi} \sin\theta d\theta \lambda_l^m(\cos\theta) b_m(\theta), \tag{B.5}$$

where

$$b_m(\theta) = \int_0^{2\pi} d\phi \frac{\Delta T}{T}(\theta,\phi)\exp(-im\phi). \tag{B.6}$$

We thus see that if the $\frac{\Delta T}{T}(\theta,\phi)$'s are measured on latitudinal circles at certain θ's, each with N_ϕ pixels evenly distributed in ϕ, equation (B.6) decays into an FFT, for which efficient algorithms are well-known, i.e.

$$b_m(\theta) = \frac{2\pi}{N_\phi}\sum_{j=0}^{N_\phi-1} \frac{\Delta T}{T}(\theta,\phi_j)\exp(-\frac{2\pi i}{N_\phi}mj). \tag{B.7}$$

When combined with equation (B.3) and a discrete version of equation (B.5), this simplification enables us to speed up the calculation of the multipole transform and therefore that of the C_l curve. In addition, there exist recurrence relations among $\lambda_l^m(\cos\theta)$'s:

$$\begin{aligned}
\lambda_m^m(y) &= (-1)^m \sqrt{\frac{2m+1}{4\pi}} \frac{(2m-1)!!}{\sqrt{(2m)!}} (1-y^2)^{m/2}, &\text{(B.8)}\\
\lambda_{m+1}^m(y) &= y\sqrt{2m+3}\,\lambda_m^m(y)\,, &\text{(B.9)}\\
\lambda_l^m(y) &= \left[y\lambda_{l-1}^m(y) - \sqrt{\frac{(l+m-1)(l-m-1)}{(2l-3)(2l-1)}}\lambda_{l-2}^m(y) \right] \sqrt{\frac{4l^2-1}{l^2-m^2}}\,. &\text{(B.10)}
\end{aligned}$$

To use these recurrence relations in our treatment for the C_l calculation, we calculate the a_{lm}'s in the same order as the summation $\sum_{m=-l_{\max}}^{l_{\max}} \sum_{l=|m|}^{l_{\max}}$ on the $l-m$ plane, instead of using the conventional order $\sum_{l=0}^{l_{\max}} \sum_{m=-l}^{l}$. There are also other properties regarding symmetry that we can use to reduce the number of calculations needed to obtain the C_l's:

$$\begin{aligned}
a_{l,-m} &= (-1)^m a_{lm}^*\,, &\text{(B.11)}\\
b_{-m}(\theta) &= b_m^*(\theta)\,, &\text{(B.12)}\\
\lambda_l^m[\cos(\pi-\theta)] &= (-1)^{l+m}\lambda_l^m(\cos\theta)\,. &\text{(B.13)}
\end{aligned}$$

A numerical scheme

To apply the above methods, we take the ECP pixelisation scheme as a working example, where pixels are evenly distributed in both θ and ϕ. The following notation will be used:

$for(k=p,q)\{process\}$	repeating $process$ $(q-p+1)$ times while taking the value of k from p to q in turn each time with an increment of 1; here k, p and q are all integers and $p<q$.
$FFT_{mj}[s,N,f_j]$	implementing the FFT, $\sum_{j=0}^{N-1} f_j \exp(s\frac{2\pi i}{N}mj)$, where $s=\pm1$.
$a=b$	assigning the value of b to a.
$a\mathrel{+}=b$	accumulating the value of b into a.
$a\mathrel{/}=b$	dividing a by b.

Thus a numerical scheme for calculating the C_l's of an ECP sky with $2N^2$ pixels (N pixels along θ and $2N$ pixels along ϕ) can be given as (with $a_{lm}=C_l=0$ initially):

$$\begin{array}{l}
for(k=0,N/2-1)\{\\
\quad \theta_k = \frac{\pi}{2N}(2k+1)\\
\quad b_m(\theta_k) = \frac{\pi}{N}FFT_{mj}[-1,2N,\frac{\Delta T}{T}(\theta_k,\phi_j)]\\
\quad b_m(\pi-\theta_k) = \frac{\pi}{N}FFT_{mj}[-1,2N,\frac{\Delta T}{T}(\pi-\theta_k,\phi_j)]\\
\quad for(m=0,N)\{\\
\qquad for(l=m,N)\{\\
\qquad\quad \text{use equations (B.8)–(B.10) to get } \lambda_l^m(\cos\theta_k)\\
\qquad\quad a_{lm} \mathrel{+}= \frac{\pi}{N}\sin\theta\,\lambda_l^m(\cos\theta_k)\left[b_m(\theta_k)+(-1)^{l+m}b_m(\pi-\theta_k)\right]\\
\qquad\quad C_l \mathrel{+}= \begin{cases} a_{lm}a_{lm}^* & \text{for } m=0\\ 2a_{lm}a_{lm}^* & \text{otherwise}\end{cases}\\
\qquad \}\\
\quad \}\\
\}\\
for(l=0,N)\{C_l \mathrel{/}= (2l+1)\}\,.
\end{array} \qquad \text{(B.14)}$$

B.2 Inverse transform and CMB map making

As an analogue of the algorithm introduced in the previous section, we now derive a method to calculate $\frac{\Delta T}{T}(\theta,\phi)$ from given a_{lm}'s. Combining equations (B.1) and (B.4), we obtain

$$\frac{\Delta T}{T}(\theta,\phi) = \sum_{m=-l}^{l} b_m(\theta)\exp(im\phi), \qquad \text{(B.15)}$$

where

$$b_m(\theta) = \sum_{l=|m|}^{\infty} a_{lm}\lambda_l^m(\cos\theta) \; . \tag{B.16}$$

When the a_{lm} are given, equations (B.15) and (B.16) provide a way to obtain a full-sky map. Although this formalism is valid for any pixelisation scheme, it is optimised when the desired pixels are evenly distributed on azimuthal circles. In this case, equation (B.15) decays to an FFT:

$$\frac{\Delta T}{T}(\theta,\phi) = \sum_{m=-N_\phi/2}^{N_\phi/2-1} b_m(\theta)\exp(\frac{2\pi i}{N_\phi}mj). \tag{B.17}$$

We notice the consistency between this equation and equation (B.7). Again, equations (B.8)-(B.13) can be employed to reduce the calculations required here.

A numerical scheme

To demonstrate the formalism developed here, we shall use the ECP pixelisation scheme, with $N_\theta = N$ and $N_\phi = 2N$. Assuming the a_{lm}'s are given, we can simulate a CMB map in the following way:

$$\begin{array}{l}
for(k=0,N/2-1)\{\\
\quad \theta_k = \frac{\pi}{2N}(2k+1)\\
\quad for(m=0,N)\{\\
\quad\quad b_m(\theta_k)=0\\
\quad\quad b_m(\pi-\theta_k)=0\\
\quad\quad for(l=m,N)\{\\
\quad\quad\quad \text{use equations (B.8)–(B.10) to get } \lambda_l^m(\cos\theta_k)\\
\quad\quad\quad d = a_{lm}\ \lambda_l^m(\cos\theta_k)\\
\quad\quad\quad b_m(\theta_k) \mathrel{+}= d\\
\quad\quad\quad b_m(\pi-\theta_k) \mathrel{+}= (-1)^{l+m}d\\
\quad\quad\}\\
\quad\quad b_{-m}(\theta_k) = b_m^*(\theta_k)\\
\quad\quad b_{-m}(\pi-\theta_k) = b_m^*(\pi-\theta_k)\\
\quad\quad \text{enforce } b_0 \text{ and } b_N \equiv b_{-N} \text{ to be real}\\
\quad\}\\
\quad \frac{\Delta T}{T}(\theta_k,\phi_j) = FFT_{jm}[1,2N,b_m(\theta_k)]\\
\quad \frac{\Delta T}{T}(\pi-\theta_k,\phi_j) = FFT_{jm}[1,2N,b_m(\pi-\theta_k)]\\
\}\ .
\end{array} \tag{B.18}$$

In Gaussian theories, the power spectrum C_l is all what we need to generate the a_{lm}'s. The following routine generates such a_{lm}'s, assuming the C_l's are known:

$$\begin{array}{l}
for(l=0,N)\{\\
\quad A_l = \sqrt{C_l/2}\\
\quad for(m=0,l)\{\\
\quad\quad \text{generate Gaussian variables } R \text{ and } I\text{, both with mean 0 and rms 1}\\
\quad\quad a_{lm} = A_l(R+iI)\\
\quad\quad a_{l,-m} = (-1)^m a_{lm}^*\\
\quad\}\\
\}\ .
\end{array} \tag{B.19}$$

A well-known method of generating Gaussian variables (i.e. R and I in the above routine) is the Box-Muller transform (e.g. [172]), which transforms a uniform distribution into a Gaussian distribution.

B.3 Numerical performance

To investigate the numerical performance of the algorithms developed above, we first estimate the operation count of their numerical schemes. For both schemes (the multipole transform and its inverse), the operation count is the same and consists of two main parts inside the outer most loop of θ:

1. the inner two loops of m and l, with a step count of $N^2/2 + 3N/2 + 1$.

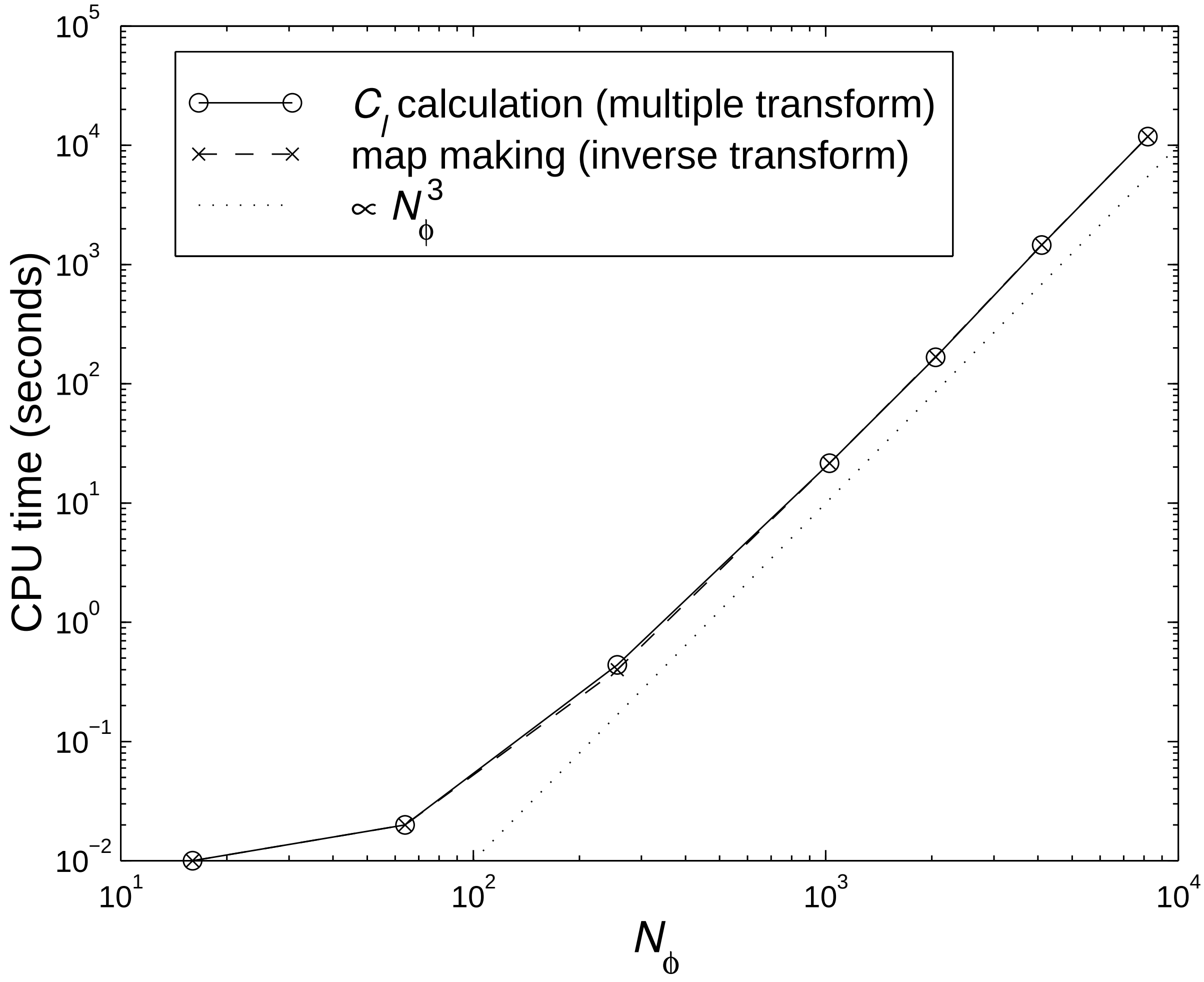

Figure B.1: CPU time spent for spherical harmonic transform and its inverse as functions of N_ϕ, in an ECP sky with a pixel number of $N_\theta N_\phi = N_\phi^2/2$. We see that for large N_ϕ, the CPU time scales as N_ϕ^3. The normalization of the dotted line is arbitrary.

2. the FFT, with a count scaling as $N \ln N$.

When N is large, the first dominates. Multiplying the sum of the above two parts by the count of the outer most loop $N/2$, we obtain the total step count as $N^3/4$ for large N, which scales as N^3. This is indeed what we observe from numerical tests. Figure B.1 plots the CPU time spent on the multipole transform and its inverse as functions of N_ϕ in an ECP sky. It shows that, the CPU time required by a spherical harmonic transform and its inverse are more or less the same, both scaling as N_ϕ^3 for large N_ϕ.

Note that the scaling is actually $N_{\rm pix}^{3/2}$, where $N_{\rm pix} \sim O(N_\phi^2)$ is the total number of pixels in the simulated or measured sky. We implement the test on a single SGI R10000 processor. For $N_\phi = 8192$ giving about 33 million pixels, the required CPU time for either transform is about 200 minutes. This enables us to deal with a large CMB data set, which is expected from the forthcoming high-accuracy experiments. Figure B.2 shows an example of a simulated full-sky CMB map with about 52 million pixels ($N_\phi = 10240$). The underlying power spectrum is that generated from Ref. [85].

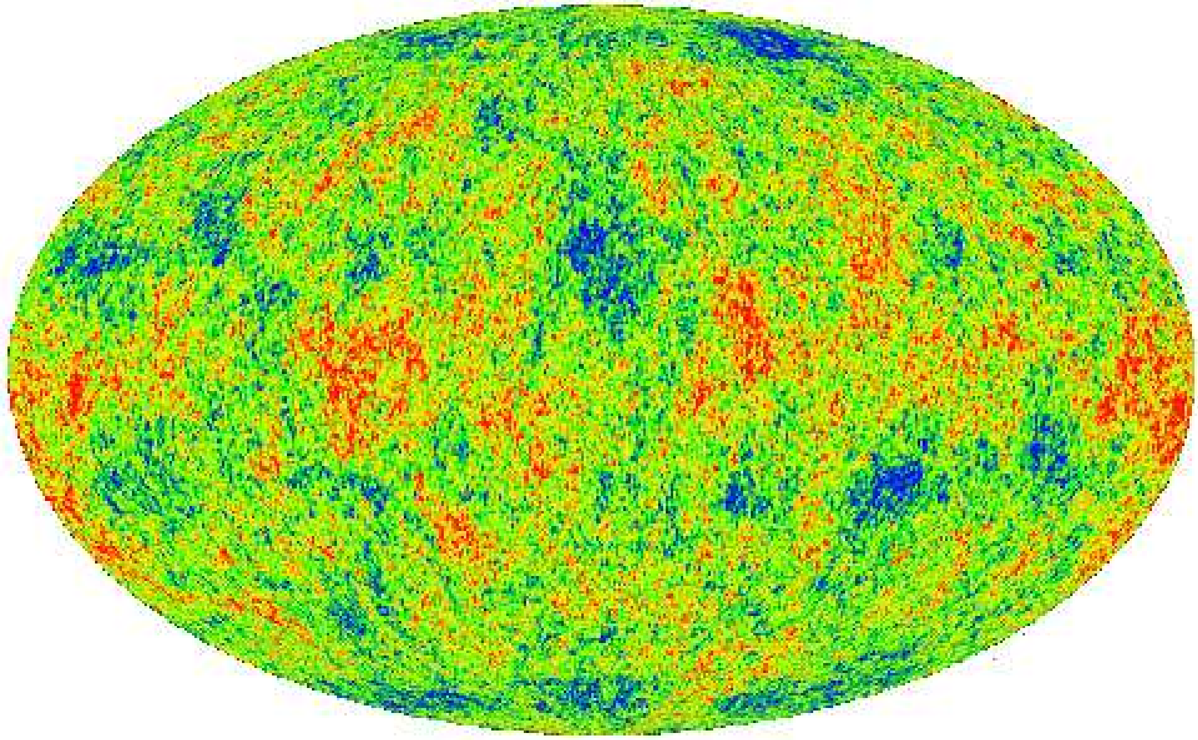

Figure B.2: A simulated full-sky CMB map in the standard CDM model. The pixelizatin scheme is the ECP with $N_\phi = 2N_\theta = 10240$, resulting in 52428800 pixels. Here we have used the Hammer-Aitoff (AIT) projection to expand the full-sky data (see Ref. [211] for a review).

Appendix C

A Package for CMB Analysis

We present a CMB graphics package which we have been developing to visualise full-sky CMB data and calculate its power spectrum, up to a resolution of around 50 million pixels. This is under the project of Cambridge Planck Analysis Centre (CPAC), part of the Planck mission.

This 3D package is written in a combination of C, C++, Fortran, and Open Inventor/GL, on a SGI Origin2000 parallel supercomputer, COSMOS. The user interface is through graphic windows and interactive mouse control, allowing users to read in CMB data of any pixelisation scheme, choose the projection method and the colour scheme, zoom in/out the three-dimensional image, rotate it, print/save it, calculate its power spectrum via the algorithm introduced in Appendix B, etc. The currently available projection methods are three-dimensional spherical projection, ZEA (Zenithal Equal-Area) projection, GLS (Sanson-Flamsteed, sinusoidal) projection, CAR (Cartesian) projection, and AIT (Hammer-Aitoff) projection (see [211] for a review). Figure C.1 shows examples of these functions.

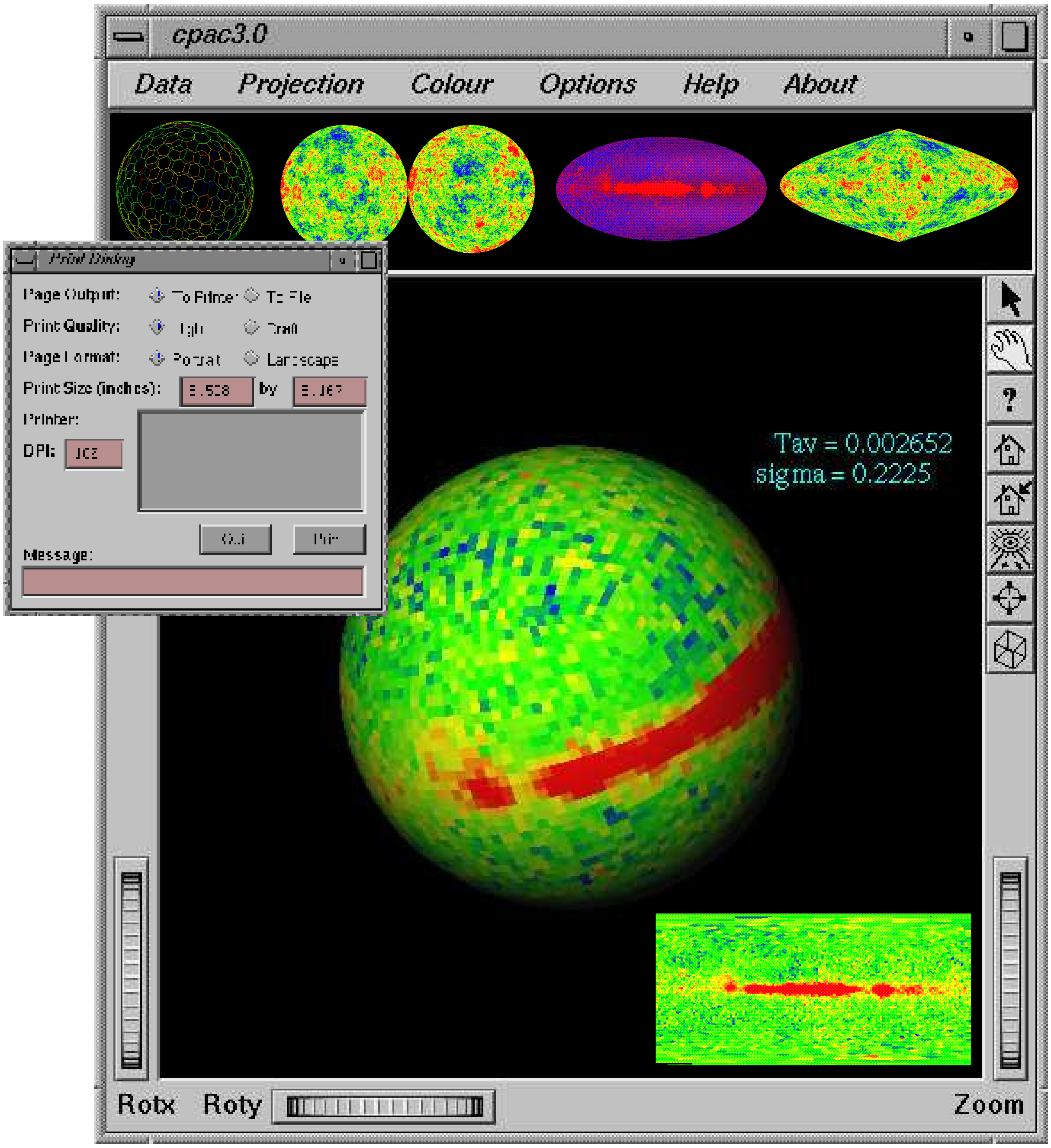

Figure C.1: A CMB graphics package. Shown in the main window is the COBE data.

Appendix D

Distribution of Cumulant Estimators for Independent Gaussian Samples

To establish or reject the hypothesis that a random variable x has a Gaussian distribution, one can estimate higher-order cumulants χ_n, $n > 2$ of that variable and check that they vanish. Given a number of realizations of x, one can compute the moment estimators $\hat{\mu}_n$ and cumulant estimators $\hat{\chi}_n$ for that purpose. However, obtaining the likelihood that the data x_i satisfies the hypothesis requires knowledge of the distribution of the cumulant estimators $\hat{\chi}_n$ for an underlying hypothetical Gaussian ensemble of realizations of x. If one assumes that the cumulant estimators $\hat{\chi}_n$ are Gaussianly distributed, one only needs to know their ensemble mean values $\langle\hat{\chi}_n\rangle$ and covariances $\langle\hat{\chi}_m\hat{\chi}_n\rangle$.

Here we give a self-contained derivation for the covariances of the cumulant estimators $\hat{\chi}_m$ computed for a set of N independent samples of one (Gaussian) random variable with standard deviation σ. Below, we show that to leading order in N^{-1}, the cumulant estimators are biased:

$$\begin{aligned} \langle\hat{\chi}_{2n}\rangle &= \chi_{2n} - N^{-1}\frac{(2n)!}{2n!}\sigma^{2n} + O\left(N^{-2}\right), && \text{(D.1)} \\ \langle\hat{\chi}_{2n+1}\rangle &= \chi_{2n+1} + O\left(N^{-2}\right), && \text{(D.2)} \end{aligned}$$

and are statistically independent:

$$\langle\hat{\chi}_m\hat{\chi}_n\rangle - \langle\hat{\chi}_m\rangle\langle\hat{\chi}_n\rangle = N^{-1}\sigma^{2n}n!\delta_{mn} + O\left(N^{-2}\right). \tag{D.3}$$

We shall generalise this formula for cumulants of two-point distributions. We shall also show that the higher-order correlations between χ_n such as $\langle\hat{\chi}_l\hat{\chi}_m\hat{\chi}_n\rangle$ are absent up to terms of order $O\left(N^2\right)$ and that these estimators can, therefore, be treated as approximately Gaussianly distributed, independent random variables with known means and variances. This is immediately relevant to the inter-scale cumulants of Fourier space amplitudes, and by using our formalism we show that these are also independent.

Consider a given set of N independent realizations x_i, $i = 1, ..., N$ of a Gaussian random variable x with mean μ and standard deviation σ,

$$\langle x_i\rangle = \mu, \quad \langle x_i x_j\rangle = \mu^2 + \sigma^2\delta_{ij}. \tag{D.4}$$

The problem at hand is to characterise the estimators for the cumulants χ_n of the distribution of x. First we take the unbiased estimators $\hat{\mu}_n$ of the moments of x, which are

$$\hat{\mu}_n \equiv N^{-1}\sum_i x_i^n. \tag{D.5}$$

The cumulant estimators $\hat{\chi}_n$ are defined through these moment estimators by the usual formulae relating cumulants and moments, e.g. the second cumulant

$$\hat{\chi}_2 = \hat{\mu}_2 - \hat{\mu}_1^2, \tag{D.6}$$

and so on (these standard relations follow from equations (D.11)–(D.13) below).

Defined in this manner, the cumulant estimators $\hat{\chi}_n$ are, however, not unbiased. For instance $\hat{\mu}_1^2$ in equation (D.6) is a biased estimator for the square of the first moment μ_1^2:

$$\left\langle\hat{\mu}_1^2\right\rangle = \langle N^{-2}\sum_{ij} x_i x_j\rangle = \mu_1^2 + N^{-1}\sigma^2. \tag{D.7}$$

We shall show now that the resulting bias in $\hat{\chi}_n$ is always of order N^{-1} or smaller. We need to consider this bias in more detail since the expectation values of χ_n are zero for $n > 2$ and also because terms of order $O\left(N^{-1}\right)$ dominate the variance of the estimators $\hat{\chi}_n$. Terms of order $O\left(N^{-1}\right)$ arise in expressions such as equation (D.6) from products of moment estimators. By analogy with equation (D.7) one obtains

$$\langle \hat{\mu}_a \hat{\mu}_b \rangle = \mu_a \mu_b + N^{-1} \left(\mu_{a+b} - \mu_a \mu_b \right). \tag{D.8}$$

Products of more than two moment estimators will also contain terms of higher order than N^{-1} but here we are only interested in the leading terms. For products of three moments, we obtain

$$\langle \hat{\mu}_a \hat{\mu}_b \hat{\mu}_c \rangle = \mu_a \mu_b \mu_c + N^{-1} \left(\mu_{a+b} \mu_c + \mu_{b+c} \mu_a + \mu_{c+a} \mu_b - 3 \mu_a \mu_b \mu_c \right) + O\left(N^{-2}\right), \tag{D.9}$$

and it is straightforward to generalise to products of k estimators,

$$\langle \hat{\mu}_{a_1} ... \hat{\mu}_{a_k} \rangle = \mu_{a_1} ... \mu_{a_k} + N^{-1} \sum_{1 \le i < j \le k} \left[\mu_{a_1} ... \mu_{a_i + a_j} ... \mu_{a_k} - \mu_{a_1} ... \mu_{a_k} \right] + O\left(N^{-2}\right). \tag{D.10}$$

The sum above is performed over all pairs of moments entering the original expression.

Now we need to find out where and how such terms appear in expressions for cumulant estimators. The relation of cumulants χ_n to moments μ_n is most easily understood by means of the generating function of moments (also called the characteristic function) $Z[p]$ which satisfies

$$\mu_n = \frac{1}{Z[0]} \left(i \frac{\partial}{\partial p} \right)^n Z[p]_{p=0}. \tag{D.11}$$

Because of the normalization $Z[0]^{-1}$ the generating function $Z[p]$ is defined up to an irrelevant constant factor. For instance, the Gaussian distribution from equation (D.4) has the generating function

$$Z[p] = \exp\left[-\frac{p^2}{2} \sigma^2 - i \mu p \right]. \tag{D.12}$$

The cumulants can be expressed (or, equivalently, defined) as

$$\chi_n = \left(i \frac{\partial}{\partial p} \right)^n \ln Z[p]_{p=0}, \tag{D.13}$$

where we note that the constant factor in the definition of $Z[p]$ drops out. The n-th derivative of $\ln Z[p]$ in equation (D.13) will contain terms such as

$$\frac{1}{Z[p]} \frac{\partial^l Z[p]}{\partial p^l} ... \frac{1}{Z[p]} \frac{\partial^m Z[p]}{\partial p^m}, \tag{D.14}$$

which give rise to products of moments $\mu_l ... \mu_m$. We would like to replace the moments μ_n by their estimators $\hat{\mu}_n$ and then to combine them pairwise to obtain the $O\left(N^{-1}\right)$ terms as shown by equation (D.10).

We notice that if we take an expression such as $\left(1 + N^{-1/2}\mu_1\right) ... \left(1 + N^{-1/2}\mu_k\right)$ and expand it in $N^{-1/2}$, then the term of order $O\left(N^{-1}\right)$ will be equal to the sum over all pairs of the product $\mu_i \mu_j$. This is almost what is required by equation (D.10), except that we also need to somehow convert $\mu_i \mu_j$ into μ_{i+j}. This would happen if μ_n behaved like a differential operator $(\partial/\partial t)^n$ acting on functions of a dummy variable t. The required technical trick is implemented as follows. We first replace the moments μ_n in equation (D.13) by the differential operator $\left(\mu_n + N^{-1/2} [i\partial/\partial t]^n\right)$. This is done by changing the generating function of moments $Z[p]$ to

$$\tilde{Z}[p;t] \equiv \sum_{n=0}^{\infty} \frac{(-ip)^n}{n!} \left(\mu_n + N^{-1/2} i^n \frac{\partial^n}{\partial t^n} \right) = Z[p] + N^{-1/2} \exp\left[p \partial_t\right]. \tag{D.15}$$

Then we introduce another copy of the old generating function of moments, $Z[t]$, on which this new generating function $\tilde{Z}[p;t]$ will act with ∂_t at the end of the calculation, so that the derivatives $[i\partial_t]^n$ will at the end be replaced again by the moments μ_n. We now claim that the expectation value of cumulant estimators $\hat{\chi}_n$ and of their products can be obtained up to and including terms of order $O\left(N^{-1}\right)$ from the modified generating function $\tilde{Z}[p;t]$ as

$$\langle \hat{\chi}_n \rangle = [i\partial_p]^n \ln \tilde{Z}[p;t] \, Z[t]_{p=t=0} + O\left(N^{-3/2}\right), \tag{D.16}$$

and

$$\langle\hat{\chi}_m\hat{\chi}_n\rangle = \left([i\partial_{p'}]^m \ln\tilde{Z}[p';t]\right)\left([i\partial_p]^n \ln\tilde{Z}[p;t]\right) Z[t]_{p=p'=t=0} + O\left(N^{-3/2}\right). \quad (D.17)$$

Here the parameters p, p' and t are set to 0 only after all derivatives have been computed. Expectation values of products of three or more cumulant estimators $\hat{\chi}_n$ can be obtained in the same manner,

$$\langle\hat{\chi}_{n_1}...\hat{\chi}_{n_k}\rangle = \left([i\partial_{p_1}]^{n_1} \ln\tilde{Z}[p_1;t]\right)...\left([i\partial_{p_k}]^{n_k} \ln\tilde{Z}[p_k;t]\right) Z[t]_{p_j=t=0} + O\left(N^{-3/2}\right). \quad (D.18)$$

We shall show first that equations (D.16)–(D.17) actually give the desired $O\left(N^{-1}\right)$ terms and then proceed to compute these terms.

Consider equation (D.16): after evaluating all derivatives in p we would obtain many terms of the form of equation (D.14) but with the modified generating function $\tilde{Z}[p;t]$ acting (with the derivatives in t) on $Z[t]$,

$$\left(\frac{1}{\tilde{Z}[p;t]}\frac{\partial^{l_1}\tilde{Z}[p;t]}{\partial p^{l_1}}...\frac{1}{\tilde{Z}[p;t]}\frac{\partial^{l_n}\tilde{Z}[p;t]}{\partial p^{l_n}}\right)_{p=0} Z[t]_{t=0}. \quad (D.19)$$

Take one such term, substitute $p=0$ and expand in $N^{-1/2}$:

$$\left(\frac{1}{\tilde{Z}[p;t]}\frac{\partial^{l_1}\tilde{Z}[p;t]}{\partial p^{l_1}}...\frac{1}{\tilde{Z}[p;t]}\frac{\partial^{l_n}\tilde{Z}[p;t]}{\partial p^{l_n}}\right)_{p=0} Z[t]_{t=0}$$

$$= \frac{\mu_{l_1} + N^{-1/2}(i\partial_t)^{l_1}}{1+N^{-1/2}}...\frac{\mu_{l_n} + N^{-1/2}(i\partial_t)^{l_n}}{1+N^{-1/2}} Z[t]_{t=0}$$

$$= \mu_{l_1}...\mu_{l_n} + N^{-1/2}\left(-n\mu_{l_1}...\mu_{l_n} + \sum_{k=1}^{n}\mu_{l_1}...(i\partial_t)^{l_k}...\mu_{l_n}\right) Z[t]_{t=0}$$

$$+N^{-1}\left(-\frac{n(n-1)}{2}\mu_{l_1}...\mu_{l_n} + \sum_{1\le i<j\le n}\mu_{l_1}...(i\partial_t)^{l_{i+j}}...\mu_{l_n}\right) Z[t]_{t=0} + O\left(N^{-3/2}\right). \quad (D.20)$$

Note that after taking the derivatives in t the $O(1)$ and $O\left(N^{-1}\right)$ terms in equation (D.20) are exactly the ones in equation (D.10), while the $O\left(N^{-1/2}\right)$ terms cancel. Therefore each term of the form of equation (D.19) yields the desired combination of moments for equation (D.16). Note that the modified generating function $\tilde{Z}[p;t]$ only gives the correct result up to $O\left(N^{-3/2}\right)$, and the further expansion terms are not useful.

The same argument can be easily shown to hold for equation (D.17) and generally for analogous expressions for averages of products of several $\hat{\chi}_n$. One only needs to take all derivatives in the parameters p, p', ... prior to taking the derivative in t.

Our strategy to obtain the expectation values from equations (D.16)–(D.17) is to first expand in $N^{-1/2}$ up to $O\left(N^{-1}\right)$, then substitute $p=0$ in the correct order, and then evaluate the derivatives in t. Since we know that the $O(1)$ terms give the unbiased result and that terms of order $O\left(N^{-1/2}\right)$ vanish (both of these statements are straightforwardly checked by a similar calculation), while the $O\left(N^{-3/2}\right)$ and higher-order terms are not useful for us, we only concentrate on the $O\left(N^{-1}\right)$ terms. Expanding the logarithm in equation (D.15),

$$\ln\tilde{Z}[p;t] = \ln Z[p] + N^{-1/2}\frac{\exp[p\partial_t]}{Z[p]} - \frac{N^{-1}}{2}\frac{\exp[2p\partial_t]}{Z[p]^2} + O\left(N^{-3/2}\right), \quad (D.21)$$

we obtain the expression

$$\ln\tilde{Z}[p;t]Z[t] = \chi_n Z[t] + N^{-1/2}\frac{Z[p+t]}{Z[p]} - \frac{N^{-1}}{2}\frac{Z[2p+t]}{Z[p]^2} + O\left(N^{-3/2}\right), \quad (D.22)$$

which will give the answer to equation (D.16) if we take the derivative in p and substitute $p=t=0$. The result is

$$\langle\hat{\chi}_n\rangle = \chi_n - \frac{N^{-1}}{2}[i\partial_p]^n_{p=0}\frac{Z[2p]}{Z[p]^2} + O\left(N^{-2}\right). \quad (D.23)$$

For a Gaussian variate with generating function given by equation (D.12) this becomes

$$\langle \hat{\chi}_n \rangle = \chi_n - \frac{N^{-1}}{2} \left[i\partial_p \right]_{p=0}^n e^{-\sigma^2 p^2} + O\left(N^{-2}\right), \tag{D.24}$$

which is the same as equations (D.1)–(D.2).

Equation (D.23) expresses the cumulant estimator bias for an arbitrary distribution through its generating function $Z[p]$. Although it can be seen that the bias for a non-Gaussian distribution will be different from that of equation (D.24), we do not need its general expression since we are only interested in testing the hypothesis of an underlying Gaussian distribution.

The calculation of the covariances of cumulants (equation (D.17)) is a little more involved since we need to expand both logarithmic terms prior to taking the derivatives in t. We begin with one logarithmic term (c.f. equation (D.22)), use the Gaussian generating function $Z[t]$ and evaluate the n-th derivative in p while keeping the t variable intact,

$$\begin{aligned} & \left[i\partial_p \right]^n \left(Z[t] \ln Z[p] + N^{-1/2} Z[t] \exp\left(-\sigma^2 pt\right) - \frac{N^{-1}}{2} Z[t] \exp\left(-\sigma^2 p^2 - 2\sigma^2 pt\right) \right) \\ = \; & Z[t]\, \chi_n + N^{-1/2} Z[t] \left(-i\right)^n \sigma^{2n} t^n - \frac{N^{-1}}{2} Z[t] \left[i\partial_p \right]_{p=0}^n \exp\left(-\sigma^2 p^2 - 2\sigma^2 pt\right). \end{aligned} \tag{D.25}$$

We note that we shall not need the full expansion of the last cumbersome derivative. Then we act on this expression with another logarithmic term (from equation (D.21)), set $t = 0$ and take the m-th derivative in p. After some straightforward algebra we obtain equation (D.3):

$$\langle \hat{\chi}_m \hat{\chi}_n \rangle - \langle \hat{\chi}_m \rangle \langle \hat{\chi}_n \rangle = N^{-1} \left[i\partial_p \right]_{p=0}^m \left(-i\right)^n \sigma^{2n} p^n = N^{-1} n! \sigma^{2n} \delta_{mn} + O\left(N^{-2}\right). \tag{D.26}$$

The same method and equation (D.23) can also be used to compute the expectation values and covariances of cumulant estimators for non-Gaussian distributions, as long as the generating function $Z[p]$ is given. Expectation values of products of two or more cumulants (e.g. $\langle \hat{\chi}_l \hat{\chi}_m \hat{\chi}_n \rangle$) can be found as well, although the calculations are cumbersome. However, in the case of the underlying Gaussian distribution a simple algebraic consideration shows that higher-order correlations between the cumulant estimators $\hat{\chi}_n$ are of order $O\left(N^{-2}\right)$ or smaller, and therefore the $\hat{\chi}_n$ themselves can be approximately treated as independent Gaussian variables. This is found by noticing that the logarithmic derivative operator of equation (D.21) contains terms of order $O\left(N^{-1/2}\right)$ and $O\left(N^{-1}\right)$ while we are only interested in the $O\left(N^{-1}\right)$ terms in the final result. We could introduce a formal operator L_n by

$$L_n \equiv \left[i\partial_p \right]^n \ln \tilde{Z}\left[p, t\right]_{p=0} = \chi_n + N^{-1/2} A_n + N^{-1} B_n + O\left(N^{-3/2}\right). \tag{D.27}$$

The operator coefficients A_n and B_n will later act on $Z[t]$ with their derivatives in t evaluated only at the end; note that $A_n Z[t]_{t=0} = 0$. Then equation (D.17) for the correlation between the cumulant estimators is rewritten as

$$\langle \hat{\chi}_m \hat{\chi}_n \rangle = L_m L_n Z[t]_{t=0} + O\left(N^{-3/2}\right). \tag{D.28}$$

Selecting the $O\left(N^{-1}\right)$ terms in the product gives

$$L_m L_n Z[t]_{t=0} - \left(L_m Z[t]_{t=0} \right) \left(L_n Z[t]_{t=0} \right) = N^{-1} A_m A_n Z[t]_{t=0} + \ldots, \tag{D.29}$$

where all terms containing B_n cancel. The only surviving $O\left(N^{-1}\right)$ term comes from the product of two $O\left(N^{-1/2}\right)$ non-commuting operator terms, while all other terms contain a commuting χ_n and cancel after subtracting the product of the expectation values. Now we notice that the $O\left(N^{-1}\right)$ terms in a product of more than two operators L_n from equation (D.27) must contain at least one factor χ_n. It follows that all "connected correlators" of more than two cumulant estimators, such as

$$\langle \hat{\chi}_l \hat{\chi}_m \hat{\chi}_n \rangle - \langle \hat{\chi}_l \rangle \langle \hat{\chi}_m \hat{\chi}_n \rangle - \langle \hat{\chi}_m \rangle \langle \hat{\chi}_l \hat{\chi}_n \rangle - \langle \hat{\chi}_n \rangle \langle \hat{\chi}_l \hat{\chi}_m \rangle + 2 \langle \hat{\chi}_l \rangle \langle \hat{\chi}_m \rangle \langle \hat{\chi}_n \rangle, \tag{D.30}$$

(for any l, m, n) contain no $O\left(N^{-1}\right)$ terms. Therefore the cumulant estimators $\hat{\chi}_n$ in the limit of a large number of samples N are approximately (up to terms of order N^{-2}) Gaussian distributed and independent variables.

Appendix E

Properties of Multivariate Cumulants

Here we examine the estimators $\hat{\chi}_{mn}$ of cumulants of a distribution of two Gaussian variables (x, y) which are independent and have variances σ_x^2 and σ_y^2. We show that the cumulant estimators are approximately independent in the limit of large sample size N, namely that their covariances are

$$\langle \hat{\chi}_{kl} \hat{\chi}_{mn} \rangle - \langle \hat{\chi}_{kl} \rangle \langle \hat{\chi}_{mn} \rangle = N^{-1} m! n! \delta_{km} \delta_{ln} \sigma_x^{2m} \sigma_y^{2n} + O\left(N^{-2}\right), \tag{E.1}$$

while the means are biased by a quantity of order N^{-1},

$$\langle \hat{\chi}_{kl} \rangle = \chi_{kl} + O\left(N^{-1}\right). \tag{E.2}$$

Cumulants of a distribution of two variables (x, y) are defined by analogy with equation (D.13),

$$\chi_{mn} = [i\partial_p]^m [i\partial_q]^n \ln Z[p, q]_{p=q=0}, \tag{E.3}$$

where $Z[p, q]$ is the generating function of moments of the distribution,

$$\mu_{mn} \equiv \langle x^m y^n \rangle = [i\partial_p]^m [i\partial_q]^n Z[p, q]_{p=q=0}. \tag{E.4}$$

A Gaussian distribution is characterised by

$$Z[p, q] = \exp\left[-\frac{1}{2}(p, q)\, \mathbf{C}\, (p, q)^T\right], \tag{E.5}$$

where we multiplied the row and column 2-vectors (p, q) by the appropriate 2 by 2 correlation matrix $\mathbf{C}$. This matrix is diagonal if the variables (x, y) are independent, but we shall not need this for most of the derivation.

We follow the same method of calculation as in Appendix D and introduce a modified (operator-valued) generating function

$$\tilde{Z}[p, q; t, u] \equiv Z[p, q] + N^{-1/2} \exp\left[p\partial_t + q\partial_u\right], \tag{E.6}$$

to be used as in equations (D.16) and (D.17). We shall omit the calculations since they are very similar to those in Appendix D and only cite the results. The bias in the cumulant estimators $\hat{\chi}_{mn}$ is described by a formula analogous to equation (D.24),

$$\langle \hat{\chi}_{mn} \rangle = \chi_{mn} - \frac{N^{-1}}{2} [i\partial_p]^m [i\partial_q]^n \exp\left[-(p, q)\, \mathbf{C}\, (p, q)^T\right]_{p=q=0}. \tag{E.7}$$

The covariance of two cumulant estimators $\hat{\chi}_{kl}$ and $\hat{\chi}_{mn}$ vanishes unless the cumulants are of the same order, $k + l = m + n$, in which case it is given by

$$\langle \hat{\chi}_{kl} \hat{\chi}_{mn} \rangle - \langle \hat{\chi}_{kl} \rangle \langle \hat{\chi}_{mn} \rangle = N^{-1} [i\partial_t]^k [i\partial_u]^l [i\partial_p]^m [i\partial_q]^n \exp\left[-(t, u)\, \mathbf{C}\, (p, q)^T\right]_{p=q=t=u=0}. \tag{E.8}$$

In our case of interest, the covariance matrix $\mathbf{C}$ is diagonal,

$$C = \begin{pmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{pmatrix}, \tag{E.9}$$

and then equations (E.7)–(E.8) simplify to

$$\langle \hat{\chi}_{mn} \rangle = \chi_{mn} - \frac{N^{-1}}{2} [i\partial_p]^m [i\partial_q]^n \exp\left[-p^2\sigma_x^2 - q^2\sigma_y^2\right]_{p=q=0} + O\left(N^{-2}\right), \tag{E.10}$$

$$\mathrm{var}\left[\hat{\chi}_{kl}\right] \equiv \left\langle \hat{\chi}_{kl}\hat{\chi}_{mn}\right\rangle - \left\langle \hat{\chi}_{kl}\right\rangle \left\langle \hat{\chi}_{mn}\right\rangle = N^{-1}m!n!\delta_{km}\delta_{ln}\sigma_x^{2m}\sigma_y^{2n} + O\left(N^{-2}\right). \tag{E.11}$$

The expressions of equations (E.10)–(E.11) are straightforwardly generalised for distributions of three and more variables, for example

$$\mathrm{var}\left[\hat{\chi}_{klm}\right] = N^{-1}k!\,l!\,m!\,\sigma_x^{2k}\sigma_y^{2l}\sigma_z^{2m} + O\left(N^{-2}\right). \tag{E.12}$$

For the case of cumulants of Fourier components of a Gaussian random field, the individual Fourier modes are independently distributed with variances equal to the power spectrum $P(k)$. Therefore the expressions we derived are applicable directly with the substitution $\sigma_i^2 = P(k_i)$, for the appropriate modes. A general Fourier cumulant $\tilde{C}^{(n)}(k_1, ..., k_n)$ of equation (11.6) with all n vectors k_i distinct will be estimated from N samples obtained by rotating the set of vectors $\{k_i\}$ and is, in our notation, a cumulant $\chi_{11\ldots1}$ of the n-variate distribution of the n Fourier modes a_{k_i}. Its sample variance is therefore

$$\mathrm{var}\left[\tilde{C}^{(n)}(k_1, ..., k_n)\right] = N^{-1}\left[\prod_{i=1}^{n} P(k_i)\right] + O\left(N^{-2}\right). \tag{E.13}$$

Bibliography

[1] P. P. Avelino, E. P. S. Shellard, J. H. P. Wu & B. Allen, Phys. Rev. Lett. **81**, 2008 (1998)

[2] P. P. Avelino, E. P. S. Shellard, J. H. P. Wu & B. Allen, Ap. J. **507**, L101 (1998)

[3] P. P. Avelino, E. P. S. Shellard, J. H. P. Wu & B. Allen, Phys. Rev. **D60**, 23511 (1999)

[4] P. P. Avelino, J. H. P. Wu & E. P. S. Shellard, *'The cluster abundance in cosmic string models for structure formation'* (see astro-ph/9906313)

[5] J. H. P. Wu, P. P. Avelino, E. P. S. Shellard & B. Allen, *'Cosmic Strings, Loops, and Linear Growth of Matter perturbations'* (see astro-ph/9812156)

[6] J. H. P. Wu, *'Causal Compensated Matter Perturbations'*, in preparation (1999)

[7] E. P. S. Shellard, J. H. P. Wu & B. Allen, *'CMB Anisotropies by Cosmic Strings—a Small Angle Approach'*, in preparation (1999)

[8] J. H. P. Wu, *'Extracting the Non-Gaussian Signal'*, in preparation (1999)

[9] S. Winitzki & J. H. P. Wu, *'Inter-scale Correlations as Measure of CMB Gaussianity'*, in preparation (1999)

[10] P. P. Avelino, E. P. S. Shellard, J. H. P. Wu & B. Allen, work in progress.

[11] S. Weinberg, *'Gravitation and Cosmology'*, Wiley, New York (1972)

[12] P. J. E. Peebles, *'Principles of Physical Cosmology'*, Princeton University Press (1993)

[13] E. W. Kolb & M. S. Turner, *'The Early Universe'*, Addison-Wesley, Redwood City, California (1990)

[14] A. A. Penzias & R. W. Wilson, Ap. J. **142**, 419 (1965)

[15] J. Mather *et al.*, Ap. J. **354**, L287 (1990)

[16] H. Guth, M. Halpern & M. Wishnaw, Phys. Rev. Lett. **65**, 937 (1990)

[17] J. Mather *et al.*, Ap. J. **420**, 439 (1994)

[18] D. J. Fixsen *et al.*, Ap. J. **456**, L71 (1996)

[19] D. J. Fixsen *et al.*, Ap. J. **473**, 576 (1996)

[20] C. J. Copi, D. N. Schramm & M. S. Turner, Science **267**, 192 (1995)

[21] D. N. Schramm & M. S. Turner, Rev. Mod. Phys. **70**, 303 (1998)

[22] G. Efstathiou, S. L. Bridle, A. N. Lasenby, M. P. Hobson & R. S. Ellis, *'Constraints on Ω_Λ and $\Omega_{\rm m}$ from Distant Type 1a Supernovae and Cosmic Microwave Background Anisotropies'* (see astro-ph/9812226; MNRAS in press)

[23] R. B. Tully, *'The Cosmological Parameters H_0 and Ω_0'* (see astro-ph/9802026)

[24] P. M. Garnavich *et al.*, Ap. J. **493**, L53 (1998)

[25] S. Perlmutter *et al.*, *'Measurements of Omega and Lambda from 42 High-Redshift Supernovae'* (see astro-ph/9812133; Ap. J. in press)

[26] C. S. Kochanek, Ap. J. **466**, 638 (1996)

[27] M. Webster, S. L. Bridle, M. P. Hobson, A. N. Lasenby, O. Lahav & G. Rocha, *'Joint estimation of cosmological parameters from CMB and IRAS data'* (see astro-ph/9802109)

[28] V. R. Eke, S. Cole & C. S. Frenk, Mon. Not. R. Astron. Soc. **282**, 263 (1996)

[29] N. A. Bahcall, *'Cosmology with Clusters of Galaxies'* (see astro-ph/9901076)

[30] M. Fukugita, C. J. Hogan & P. J. E. Peebles, *'The Cosmic Baryon Budget'* (see astro-ph/9712020)

[31] W. L. Freedman, J. R. Mould, R. C. Kennicutt & B. F. Madore, *'The Hubble Space Telescope Key Project to Measure the Hubble Constant'* (see astro-ph/9801080)

[32] C. H. Lineweaver & D. Barbosa, Ap. J. **496**, 624 (1998)

[33] J. A. Peacock & S. J. Dodds, Mon. Not. R. Astron. Soc. **267**, 1020 (1994)

[34] G. F. Smoot *et al.*, Ap. J. **396**, L1 (1992)

[35] J. R. Bond & G. Efstathiou, Ap. J. **285**, L45 (1984)

[36] G. Efstathiou, *'The Physics of The Early Universe'*, edited by A. Heavens, A. Davies & J. Peacock (Scottish Universities Summer School Publications, Edinburgh, 1990)

[37] A. R. Liddle & D. H. Lyth, Phys. Rep. **231**, 1 (1993)

[38] For a review see A. Vilenkin & E. P. S. Shellard, *'Cosmic strings and other topological defects'*, Cambridge University Press, Cambridge (1994)

[39] J. H. P. Wu, *'Analytical and Numerical Approaches to Inflationary Cosmology'* (MSc thesis, University of Sussex, 1996)

[40] A. H. Guth, Phys. Rev. **D23**, 347 (1981)

[41] Los Alamos National Laboratory e-Print Archive, `http://xxx.lanl.gov`

[42] A. Albrecht, P. J. Steinhardt, M. S. Turner & F. Wilczek, Phys. Rev. Lett. **48**, 1437 (1982)

[43] L. F. Abbott, E. Fahri & M. B. Wise, Phys. Lett. **117B**, 29 (1982)

[44] A. Dolgov & A. D. Linde, Phys. Lett. **116B**, 329 (1982)

[45] A. R. Liddle, P. Parsons & J. D. Barrow, Phys. Rev. **D50**, 7222 (1994)

[46] D. S. Salopek & J. R. Bond, Phys. Rev. **D42**, 3936 (1990)

[47] A. G. Muslimov, Class. Quantum Grav. **7**, 231 (1990)

[48] J. E. Lidsey, Phys. Lett. **B273**, 42 (1991)

[49] D. H. Lyth, Phys. Rev. **D31**, 1792 (1985)

[50] E. D. Stewart & D. H. Lyth, Phys. Lett. **B302**, 171 (1993)

[51] R. Scaramella & N. Vittorio, Ap. J. **353**, 372 (1990)

[52] A. A. Starobinsky, Sov. Astron. Lett. **11**, 133 (1985)

[53] E. F. Bunn, A. R. Liddle & M. White, Phys. Rev. **D54**, 5917 (1996)

[54] C. L. Bennett *et al.*, Ap. J. **464**, L1 (1996)

[55] J. Magueijo, A. Albrecht, D. Coulson & P. Ferreira, Phys. Rev. Lett. **76**, 2617 (1996)

[56] U. L. Pen, U. Seljak & N. Turok, Phys. Rev. Lett. **79**, 1615 (1997)

[57] A. R. Liddle, *'Aspects of inflationary reconstruction'* (see astro-ph/9903195) Proceedings for 19th Texas Symposium on Relativistic Astrophysics, Paris (1999)

[58] S. W. Hawking & N. Turok, Phys. Lett. **B425**, 25 (1998)

[59] N. Turok & S. W. Hawking, Phys. Lett. **B432**, 271 (1998)

[60] J. Goldstone, Nuovo Cim. **19**, 154 (1961)

[61] T. W. B. Kibble, J. Phys. **A9**, 1387 (1976)

[62] N. G. Turok & J. Zadrozny, Phys. Rev. Lett. **65**, 2331 (1990)

[63] Ya. B. Zel'dovich, I. Kobzarev & L. Okun, Zh. Eksp. Theor. Fiz. **67**, 3 (1974)

[64] Ya. B. Zel'dovich & M. Khlopov, Phys. Lett. **B79**, 239 (1978)

[65] J. Preskill, Phys. Rev. Lett. **43**, 365 (1979)

[66] P. W. Higgs, Phys. Lett. **12**, 132 (1964)

[67] N. Turok & P. Bhattacharjee, Phys. Rev. **D29**, 1557 (1984)

[68] E. P. S. Shellard, Nucl. Phys. **B283**, 624 (1987)

[69] E. Bertschinger, Ap. J. **316**, 489 (1987)

[70] N. Turok, Phys. Lett. **B123**, 387 (1983)

[71] A. Vilenkin, Phys. Rev. **D41**, 3038 (1990)

[72] B. Carter, Phys. Rev. **D41**, 3869 (1990)

[73] T. Vachaspati & A. Vilenkin, Phys. Rev. Lett. **67**, 1057 (1991)

[74] C. J. A. P. Martins & E. P. S. Shellard, Phys. Rev. **D53**, R575 (1996)

[75] C. J. A. P. Martins & E. P. S. Shellard, Phys. Rev. **D54**, 2535 (1996)

[76] A. Albrecht & N. Turok, Phys. Rev. Lett. **54**, 1868 (1985)

[77] A. Albrecht & N. Turok, Phys. Rev. **D40**, 973 (1989)

[78] D. P. Bennett & F. R. Bouchet, Phys. Rev. **D41**, 2408 (1990); D. P. Bennett & F. R. Bouchet, Phys. Rev. Lett. **60**, 257 (1988)

[79] D. P. Bennett, *'High Resolution Simulations of Cosmic String Evolution: Numerics and Long String Evolution'* (1990); F. R. Bouchet, *'High Resolution Simulations of Cosmic String Evolution: Small Scale Structure and Loops'* in *'The Formation and Evolution of Cosmic Strings'*, Cambridge University Press, Cambridge (1990)

[80] B. Allen & E. P. S. Shellard Phys. Rev. **D45**, 1898 (1992)

[81] S. Veeraraghavan & A. Stebbins, Ap. J. **365**, 37 (1990)

[82] U. Pen, D. Spergel & N. Turok, Phys. Rev. **D49**, 692 (1994)

[83] W. Hu, U. Seljak, M. White & M. Zaldarriaga, Phys. Rev. **D57**, 3290 (1998)

[84] W. Hu & M. White, Phys. Rev. **D56**, 596 (1997)

[85] U. Seljak & M. Zaldarriaga, Ap. J. **469**, 437 (1996)

[86] K. C. Jacobs, Nature **215**, 1156 (1967)

[87] J. M. Cohen, Nature **216**, 249 (1967)

[88] A. D. Chernin, Nature **220**, 250 (1968)

[89] C. B. G. McIntosh, Mon. Not. R. Astron. Soc. **140**, 461 (1968)

[90] We thank N. Turok for providing the idea of spliting the photon and baryon components here.

[91] D. J. Eisenstein & W. Hu, Ap. J. **496**, 605 (1998)

[92] W. Hu & N. Sugiyama, Ap. J. **444**, 489 (1995)

[93] W. Hu & N. Sugiyama, Ap. J. **471**, 542 (1996)

[94] J. Silk, Ap. J. **151**, 459 (1968)

[95] R. A. Battye, Phys. Rev. **D55**, 7361 (1997)

[96] R. Sunyaev & Ya. B. Zel'dovich, Ap&SS, **7**, 3 (1970)

[97] W. Press & E. Vishniac, Ap. J. **236**, 323 (1980)

[98] S. M. Carroll, W. H. Press & E. L. Turner, Annu. Rev. Astron. Astrophys. **30**, 499 (1992)

[99] D. J. Eisenstein, *'An Analytic Expression for the Growth Function in a Flat Universe with a Cosmological Constant'* (see astro-ph/9709054)

[100] C. Cheung & J. Magueijo, Phys. Rev. **D56**, 1982 (1997)

[101] J. Robinson & B. D. Wandelt, Phys. Rev. **D53**, 618 (1996)

[102] A. Stebbins, *'A New Picture for Cosmic String seeded structure formation'* in *'The Formation and Evolution of Cosmic Strings'*, Cambridge University Press, Cambridge (1990)

[103] B. Allen & E. P. S. Shellard, Phys. Rev. Lett. **64**, 119 (1990); E. P. S. Shellard & B. Allen, *'On the Evolution of Cosmic Strings'* in *'The Formation and Evolution of Cosmic Strings'*, Cambridge University Press, Cambridge (1990)

[104] We thank R. R. Caldwell for his contribution here (see Ref. [119])

[105] J. V. Villumsen, Ap. J. Suppl. Ser. **71**, 407 (1989)

[106] A. Albrecht, R. A. Battye & J. Robinson, Phys. Rev. Lett. **79**, 4736 (1997)

[107] R. A. Battye, A. Albrecht & J. Robinson, Phys. Rev. Lett. **80**, 4847 (1998)

[108] A. Vilenkin, Phys. Rev. Lett. **46**, 1169 (1981)

[109] J. Silk & A. Vilenkin, Phys. Rev. Lett. **53**, 1700 (1984); Ya. B. Zel'dovich, Mon. Not. R. Astron. Soc. **192**, 663 (1980)

[110] R. Basu, A. H. Guth & A. Vilenkin, Phys. Rev. **D44**, 340 (1991)

[111] J. Borrill, Phys. Rev. Lett. **76**, 3255 (1996)

[112] E. J. Copeland, T. W. B. Kibble & D. A. Steer, Phys. Rev. **D58**, 43508 (1998)

[113] C. Contaldi, M. Hindmarsh & J. Magueijo, Phys. Rev. Lett. **82**, 679 (1999)

[114] G. Vincent, N. D. Antunes & M. Hindmarsh, Phys. Rev. Lett. **80**, 2277 (1998)

[115] J. N. Moore & E. P. S. Shellard, *'On the evolution of abelian-Higgs string networks'* (see hep-ph/9808336)

[116] R. J. Scherrer, J. M. Quashnock, D. N. Spergel & W. H. Press, Phys. Rev. **D39**, 371 (1989)

[117] T. Vachaspati & A. Vilenkin, Phys. Rev. **D31**, 3052 (1985)

[118] P. P. Avelino, R. R. Caldwell & C. J. A. P. Martins, Phys. Rev. **D56**, 4568 (1997)

[119] B. Allen, R. R. Caldwell, S. Dodelson, L. Knox, E. P. S. Shellard & A. Stebbins, Phys. Rev. Lett. **79**, 2624 (1997)

[120] R. Durrer, M. Kunz & A. Melchiorri, Phys. Rev. **D59**, 123005 (1999)

[121] A. Albrecht & A. Stebbins, Phys. Rev. Lett. **68**, 2121 (1992)

[122] A. Albrecht & A. Stebbins, Phys. Rev. Lett. **69**, 2615 (1992)

[123] A. R. Liddle, D. H. Lyth, D. Roberts & P. T. P. Viana, Mon. Not. R. Astron. Soc. **278**, 644 (1996)

[124] A. Jenkins, C. S. Frenk, F. R. Pearce, P. A. Thomas, J. M. Colberg, S. D. M. White, H. M. P. Couchman, J. A. Peacock, G. Efstathiou & A. H. Nelson, Ap. J. **499**, 20 (1998)

[125] P. P. Avelino & J. P. M. de Carvalho, *'The growth of linear perturbations in generic defect models for structure formation'* (see astro-ph/9810364)

[126] P. P. Avelino, PhD thesis, DAMTP, University of Cambridge (1996)

[127] P. P. Avelino & R. R. Caldwell, Phys. Rev. **D53**, 5339 (1996)

[128] A. Sornborger, R. Brandenberger, B. Fryxell & K. Olson, *'The Structure of Cosmic String Wakes'* (see astro-ph/9704145) Proceedings for 18th Texas Symposium on Relativistic Astrophysics, Paris (1997)

[129] P. T. P. Viana & A. R. Liddle, Mon. Not. R. Astron. Soc. **303**, 535 (1999)

[130] S. D. M. White, G. Efstathiou & C. S. Frenk, Mon. Not. R. Astron. Soc. **262**, 1023 (1993)

[131] P. T. P. Viana & A. R. Liddle, Mon. Not. R. Astron. Soc. **281**, 323 (1996)

[132] P. Ferreira, Phys. Rev. Lett. **74**, 3522 (1995)

[133] A. Stebbins, private communication. (Unpublished work by S. Colombi, B. Bennett, F. Bouchet and A. Stebbins, 1994)

[134] B. Allen, R. R. Caldwell, E. P. S. Shellard, A. Stebbins & S. Veeraraghavan, Phys. Rev. Lett. **77**, 3061 (1996)

[135] N. Sugiyama, *'Cosmic Background Anisotropies in CDM Cosmology'* (see astro-ph/9412025)

[136] T. Vachaspati, Phys. Lett. **B282**, 305 (1992)

[137] R. Moessner 1995, Mon. Not. R. Astron. Soc. **277**, 927 (1995)

[138] C. Park, D. N. Spergel & N. Turok, Ap. J. **372**, L53 (1991)

[139] R. H. Brandenberger, D. M. Kaplan & S. A. Ramsey, *'Some Statistics for Measuring Large-Scale Structure'* (see astro-ph/9310004)

[140] J. Robinson & A. Albrecht, *'A Statistic for identifying cosmic string wakes and other sheet-like structure'* (see astro-ph/9505123)

[141] P. P. Avelino, Ap. J. **487**, 18 (1997)

[142] J. R. Gott, A. L. Melott & M. Dickinson, Ap. J. **306**, 341 (1986)

[143] M. S. Vogeley, C. Park, M. J. Geller, J. P. Huchra & J. R. Gott, Ap. J. **420**, 525 (1994)

[144] W. Saunders *et al.*, Nature **349**, 32 (1991)

[145] A. Canavezes *et al.*, *'The Topology of the IRAS Point Source Catalogue Redshift Survey'* (see astro-ph/9712228; MNRAS in press)

[146] J. N. Fry & R. J. Scherrer, Ap. J. **429**, 36 (1994)

[147] M. J. Chodorowski & F. R. Bouchet, *'Kurtosis in Large-Scale Structure as a Constraint on Non-Gaussian Initial Conditions'* (see astro-ph/9507038)

[148] S. D. M. White, G. Efstathiou & C. S. Frenk, Mon. Not. R. Astron. Soc. **262**, 1023 (1993)

[149] T. Kitayama & Y. Suto, Ap. J. **490**, 557 (1997)

[150] L. Wang & P. J. Steinhardt, Ap. J. **508**, 483 (1998)

[151] W. H. Press & P. Schechter, Ap. J. **187**, 452 (1974)

[152] W. A. Chiu, J. P. Ostriker & M. A. Strauss, Ap. J. **494**, 479 (1997)

[153] J. Robinson & J. E. Baker, *'Evolution of the cluster abundance in non-Gaussian models'* (see astro-ph/9905098)

[154] K. Koyama, J. Soda & A. Taruya, *'Constraints on a non-gaussian (χ^2_m) CDM model'* (see astro-ph/9903027)

[155] J. Robinson, E. Gawiser & J. Silk, *'A Simultaneous Constraint on the Amplitude and Gaussianity of Mass Fluctuations in the Universe'* (see astro-ph/9805181)

[156] J. P. Henry, Ap. J. **489**, L1 (1997)

[157] J. P. Henry, in preparation (1999)

[158] S. Colafrancesco, F. Lucchin & S. Matarrese, Ap. J. **345**, 3 (1989)

[159] F. Bernardeau, Ap. J. **427**, 51 (1994)

[160] C. Lacey & S. Cole, Mon. Not. R. Astron. Soc. **262**, 627 (1993)

[161] C. Lacey & S. Cole, Mon. Not. R. Astron. Soc. **271**, 676 (1994)

[162] J. F. Navarro, C. S. Frenk & S. D. M. White, Mon. Not. R. Astron. Soc. **275**, 720 (1995)

[163] C. van de Bruck, *'Cosmic Strings and Structure Formation'* (see astro-ph/9810409) Proceedings for the Second International Workshop on Dark Matter (DARK98), eds. H. V. Klapdor-Kleingrothaus & L. Baudis (1998)

[164] A. N. Lasenby, A. W. Jones & Y. Dabrowski, Phil. Trans. R. Soc. Lond. A., (1998; in press)

[165] N. Kaiser & A. Stebbins, Nature **310**, 391 (1984)

[166] J. Gott, Ap. J. **288**, 422 (1985)

[167] N. Turok & D. Spergel, Phys. Rev. Lett. **64**, 2736 (1990)

[168] M. White, J. E. Carlstrom, M. Dragovan & W. L. Holzapfel, Ap. J. **514**, 12 (1999)

[169] W. Hu, N. Sugiyama & J. Silk, Nature **386**, 37 (1997)

[170] R. Sachs & A. Wolfe, Ap. J. **147**, 73 (1967)

[171] D. Sáez, E. Holtmann & G. F. Smoot, *'Small Angular Scale Simulations of the Microwave Sky'* (see astro-ph/9606164; Ap. J. in press)

[172] P. H. Press, S. A. Teukolsky, W. T. Vettering & B. P. Flannery, *'Numerical Recipes'*, Cambridge University Press, Cambridge (1994)

[173] I. S. Gradshteyn & I. M. Ryzhik, *'Table of Integrals, Series and Products'*, Academic Press, New York (1980)

[174] W. Hu, D. N. Spergel & M. White, Phys. Rev. **D55**, 3288 (1997)

[175] A. Stebbins, Ap. J. **327**, 584 (1988)

[176] M. Hindmarsh, Ap. J. **431**, 534 (1994)

[177] S. Veeraraghavan & A. Stebbins, Phys. Rev. **D51**, 1465 (1995)

[178] S. Veeraraghavan & A. Stebbins, (unpublished)

[179] J. Borrill, E. J. Copeland, A. R. Liddle & A. Stebbins, Phys. Rev. **D50**, 2469 (1994)

[180] F. R. Bouchet, D. P. Bennett & A. Stebbins, Nature **335**, 410 (1988)

[181] P. J. E. Peebles, *'An Isocurvature CDM Cosmogony. I. A Worked Example of Evolution Through Inflation'* (see astro-ph/9805194)

[182] P. J. E. Peebles, *'An Isocurvature CDM Cosmogony. II. Observational Tests'* (see astro-ph/9805212)

[183] L. Amendola, Astro. Lett. and Communications **33**, 63 (1996)

[184] P. Graham, N. Turok, P. M. Lubin & J. A. Schuster, *'A Simple Test for Non-Gaussianity in CMBR Measurements'* (see hep-ph/9307236)

[185] X. Luo, Phys. Rev. **D49**, 3810 (1994)

[186] A. Kogut, A. J. Banday, C. L. Bennett, K. M. Górski, G. Hinshaw, G. F. Smoot & E. L. Wright, Ap. J. **464**, L29 (1996)

[187] S. Winitzki, *'Non-Gaussian Signature of Weak Gravitational Lensing in the CMB'* (see astro-ph/9806105)

[188] M. P. Hobson, A. W. Jones & A. N. Lasenby, *'Wavelet analysis and the detection of non-Gaussianity in the CMB'* (see astro-ph/9810200)

[189] L. Popa, *'Power Fluctuations in the Wavelet Space: Large-Scale CMB Non-Gaussian Statistics'* (see astro-ph/9806086)

[190] P. G. Ferreira, J. Magueijo & J. Silk, Phys. Rev. **D56**, 4592 (1997)

[191] X. Luo & D. N. Schramm, Ap. J. **408**, 33 (1993)

[192] A. F. Heavens, *'Estimating Non-Gaussianity in the Microwave Background'* (see astro-ph/9804222; MNRAS in press)

[193] P. G. Ferreira, J. Magueijo & K. M. Gorski, Ap. J. **503**, L1 (1998)

[194] A. J. Banday, S. Zaroubi & K. M. Gorski, *'On the Non-Gaussianity Observed in the COBE-DMR Sky Maps'* (see astro-ph/9908070)

[195] B. C. Bromley & M. Tegmark, *'Is the Cosmic Microwave Background Really Non-Gaussian?'* (see astro-ph/9904254; Ap. J. in press)

[196] H. Minkowski, Mathematische Annalen **57**, 447 (1903)

[197] S. Winitzki & A. Kosowsky, New Astronomy **3**, No. 2, 75 (1997)

[198] J. Schmalzing & K. M. Górski, Mon. Not. R. Astron. Soc. **297**, 355 (1998)

[199] P. Coles & J. D. Barrow, Mon. Not. R. Astron. Soc. **228**, 407 (1987)

[200] P. Coles, Mon. Not. R. Astron. Soc. **243**, 509 (1988)

[201] G. F. Smoot, L. J. B. A. Tenorio, A. Kogut, E. L. Wright, G. Hinshaw & C. L. Bennett, Ap. J. **437**, 1 (1994)

[202] W. N. Colley, J. R. Gott III & C. Park, Mon. Not. R. Astron. Soc. **281**, 82 (1996)

[203] J. R. Bond & G. Efstathiou, Mon. Not. R. Astron. Soc. **226**, 655 (1987)

[204] N. Vittorio & R. Juskiewicz, Ap. J. **314**, L29 (1987)

[205] N. Turok, Ap. J. **473**, L5 (1996)

[206] R. B. Barreiro, J. L. Sanz, E. Martínez-González & J. Silk, *'Correlation of Excursion Sets for Non-Gaussian CMB Temperature Distributions'* (see astro-ph/9712143; MNRAS in press)

[207] P. G. Ferreira & J. Magueijo, Phys. Rev. **D55**, 3358 (1997)

[208] A. Lewin, A. Albrecht & J. Magueijo, Mon. Not. R. Astron. Soc. **302**, 131 (1999)

[209] J. Pando, D. Valls-Gabaud & L.-Z. Fang, Phys. Rev. Lett. **81**, 4568 (1998)

[210] M. P. Hobson *et al.*, *'The Effect of Point Sources on Satellite Observations of the Cosmic Microwave Background'* (see astro-ph/9810241)

[211] E. W. Greisen & M. Calabretta, Bull. American Astron. Soc. **25**, No. 2, 808 (1993)

[212] R. G. Crittenden & N. G. Turok, *'Exactly Azimuthal Pixelizations of the Sky'* (see astro-ph/9806374)