High-resolution CMB bispectrum estimator

Wu Hyun Sohn

Department of Applied Mathematics and Theoretical Physics
University of Cambridge

This dissertation is submitted for the degree of
Doctor of Philosophy

Trinity College July 2021

Table of contents

List of figures

List of tables

Chapter 1

Introduction

1.1 From background to foreground

New outline: - GR. Statistical homogeneity and isotropy -> FLRW metric. No static solution.

- Hubble's results. Expanding universe -> Big Bang.

- Discovery of CMB. CMB is a key prediction of Big Bang theory

- Big Bang theory's problem with initial conditions. Inflation. Quantum fluctuations as seeds for inhomogeneities of the universe.

- CMB anisotropy, 1 in 10^5. Consistent with inflationary predictions.

- Observational evidence for dark energy and dark matter. Emergence of inflationary LCDM.

- LCDM successful. Planck CMB power spectrum -> consistent and measures parameters to percent level. Era of precision cosmology.

- Quick summary of CMB's usefulness! Its existence, anisotropy and spectrum all contributed... CMB probes initial conditions, growth of the universe and matter distribution (through lensing).

- Next big question - inflation's mechanism? Primordial non-Gaussianity is a robust prediction of many models. Three-point correlation function as a key statistic for differentiation.

- CMB bispectrum. Planck gave a lot of insight. No detection, but oscillatory models are of interest. Main challenge - computational complexity.

- Next generation of CMB experiments upcoming. Polarisation sensitivity improved, higher resolution. Need for development of new, improved pipeline capable of dealing with high resolution data.

- Thesis organisation.

Radio astronomers Arno Penzias and Robert Wilson were calibrating their 50-feet-long horn antenna when they found a mysterious background noise. The measurements were independent of time and location in the sky, and persisted after the removal of various potential contaminants. After the theoretical work of Robert Dicke, Jim Peebles, and David Wilkinson was brought forward [13], Penzias and Wilson identified the noise as cosmic microwave background radiation (CMB): ancient light from the early universe reaching us after billions of years [36]. The discovery provided us with one of the most valuable probes of the physical universe, leading to major development in observational cosmology.

On the theoretical side, modern mathematical formulation of cosmology owes to Einstein's work on general relativity in 1915. Using his framework, Friedmann, Lemaître, Robertson, and Walker contributed to writing down the unique metric for spatially homogeneous and isotropic universe. The FLRW metric dictates growth of the universe from the Big Bang to present day. Such expansion of the universe was supported by Edwin Hubble's measurements of Cepheid variables and redshift (add year), as well as the aforementioned discovery of CMB. What is widely accepted to be the standard model of modern cosmology, the ΛCDM model, appeared only in the late 1990s. The six-parameter model assumes presence of cold dark matter and dark energy, in addition to baryons and radiation, as main contributors to the total energy density of the universe.

The ΛCDM model has been extremely successful in explaining modern cosmological observations. CMB measurements from *Planck* satellite, in particular, show exceptional agreement with the model. Planck was a space observatory developed by the European Space Agency. The Planck satellite observed the CMB in nine frequency bands from 2009 to 2013, with resolution and sensitivity substantially improved compared to its predecessor - the Wilkinson Microwave Anisotropy Probe (WMAP). Able to resolve CMB anisotropy in much smaller scale, Planck placed one of the most stringent bounds on the theoretical parameters of ΛCDM so far.

How does CMB contain so much information about the universe? The answer is twofold. First is due to the fact that CMB anisotropy originates from primordial perturbations. Statistical properties of the initial fluctuations can be deduced from analysing correlation functions of the CMB, letting us constrain the early universe physics. Second reason is that CMB tracks history of the universe as it travels from the background to our foreground. CMB photons scatter with baryons before free-streaming all the way to our foreground, which then experience both growth of the universe and gravitational potential of matter perturbations. These signatures are engraved in CMB anisotropy spectrum, redshift, and weak lensing.

The CMB anisotropy is observed to be nearly Gaussian distributed. Statistical characteristics of a Gaussian random field can be summarised entirely using two-point correlation functions, or their Fourier counterpart: power spectrum. The CMB power spectra have been thoroughly studied to constrain various cosmological parameters. Meanwhile, higher-order statistics such as three-point correlation functions (bispectrum in Fourier space) also contain valuable information about our universe. They probe non-Gaussian statistics of the CMB arising from both primordial origin and late-time effects.

Primordial non-Gaussianity is a key statistic for studying physics of the early universe. The theory of inflation has been successful in describing the observed data, but its exact mechanism is yet undetermined. Currently there are numerous viable inflationary models with well-founded physical motivations. Non-Gaussian signatures of primordial fluctuations are robust predictions of various models, and measuring their shape and amplitude allows us to constrain inflationary scenario. CMB bispectrum analysis from Planck yielded the most precise measurements of primordial non-Gaussianity to date. So far, no statistically significant amount of non-Gaussianity has been detected.

In the near future, we expect several new major CMB experiments. Simons Observatory (SO) is a ground-based experiment currently under construction in the Atacama Desert of Chile. SO is expected to measure both CMB temperature and polarisation to unprecedented precision, largely improved compared to Planck. The first light from SO is planned to be observed in early 2022. Many more CMB Stage-4 (CMB-S4) experiments are proposed, brightening future prospects for CMB. In particular, the upcoming measurements will allow us to constrain primordial non-Gaussianity further, providing discovery potential.

This thesis is organised as follows. In Chapter 1, we review the standard formulation of cosmology, deriving the form of scale factor in homogeneous universe. Motivation and formalism of slow-roll inflation is also presented here. Chapter 2 details cosmological perturbation theory, with a focus on the CMB anisotropy. We summarise methods for computing the transfer function of CMB, and introduce the concept of CMB polarisation. Next, in Chapter 3 we define bispectrum and discuss how it can be used to probe primordial non-Gaussianity. An example of single-field inflation with non-canonical kinetic term will be provided to demonstrate computation of non-Gaussianity using the in-in formalism.

Chapter 4 and 5 contain my original work, based on research conducted in collaboration with my supervisor James Fergusson. Chapter 4 contains the forecast for future CMB-S4 surveys on the primordial non-Gaussianity parameter f_{NL}. SO experiment specifications and expected CMB-S4 setup were used to predict their improved constraints via Fisher information analysis. We focussed on models with oscillatory features, where steep enhancement in polarisation sensitivity greatly benefit constraining power.

Motivated by the positive prospects from forecasts detailed in Chapter 4, we worked on developing a high-resolution bispectrum estimation pipeline suitable for future surveys. Chapter 5 contains formulation and development details of the developed program, as well as consistency checks from the thorough verification process. We outline the benefits of new pipeline compared to conventional methods, and present some working examples. Lastly, Chapter 6 concludes the thesis by summarising and presenting plans for future research.

1.2 The homogeneous universe

In this section we review the standard cosmological formulation for the homogeneous universe, neglecting any perturbations. What we derive here will serve as a background solution for the full perturbative result discussed in the next chapter. We assume general relativity to be an accurate theory of gravity for relevant scales.

1.2.1 Geometry

In general relativity, spacetime is represented by a 4-dimensional Lorentzian manifold equipped with a metric. Distance measure in curved spacetime is given by the metric tensor g;

$$ds^2 = g_{\mu\nu}\, dx^\mu dx^\nu, \tag{1.1}$$

where the Greek letters $\mu, \nu = 0, 1, 2, 3$ represent time (0) and spatial $(1, 2, 3)$ indices of local coordinates. Flat spacetime has metric $g_{\mu\nu} = \eta_{\mu\nu} = \mathrm{diag}\{-1, 1, 1, 1\}$, also known as the Minkowski metric. Throughout this thesis we adopt the sign convention $(-, +, +, +)$ and work in units where $c = 1$. Unless specified otherwise, the Einstein summation convention is assumed.

In curved spacetime, free particles follow a trajectory given by the geodesic equation;

$$\frac{d^2x^\mu}{ds^2} + \Gamma^\mu_{\nu\rho}\frac{dx^\nu}{ds}\frac{dx^\rho}{ds} = 0, \tag{1.2}$$

with s an affine parameter parametrising the trajectory, and $\Gamma^\mu_{\nu\rho}$ the Christoffel symbol representing metric connection. Its value is given in terms of the metric tensor by

$$\Gamma^\mu_{\nu\rho} = \frac{1}{2}\, g^{\mu\sigma}\left(\partial_\rho g_{\nu\sigma} + \partial_\nu g_{\rho\sigma} - \partial_\sigma g_{\nu\rho}\right). \tag{1.3}$$

Here and throughout this thesis, ∂_μ denote the partial derivative with respect to local coordinate x^μ. Note $g^{\nu\sigma}$ is the inverse metric satisfying $g^{\mu\nu} g_{\nu\rho} = \delta^\nu_\rho$.

Defining tangent vector as $U^\mu = dx^\mu/ds$, the equation can be rewritten in a covariant form given by

$$(\nabla_U U)^a = U^b \nabla_b U^a = 0. \tag{1.4}$$

We follow the convention where Roman letters are used for abstract indices. Note that in terms of local coordinates, the covariant derivative of a vector field is defined as $\nabla_\nu U^\mu = \partial_\nu U^\mu + \Gamma^\mu_{\nu\rho} U^\rho$.

Distance between two geodesics that are initially parallel may change in curved spacetime. Such geometric information is encapsulated within the Riemann curvature tensor R^a_{bcd}. [1] From R^a_{bcd} we can evaluate the Ricci curvature tensor R_{ab}, the Ricci scalar R, and finally the Einstein tensor G_{ab}. They are defined as follows.

$$\begin{aligned}
R^\mu_{\nu\rho\sigma} &:= \partial_\rho \Gamma^\mu_{\nu\sigma} - \partial_\sigma \Gamma^\mu_{\nu\rho} + \Gamma^\tau_{\nu\sigma}\Gamma^\mu_{\tau\rho} - \Gamma^\tau_{\nu\rho}\Gamma^\mu_{\tau\sigma} && (1.5)\\
R_{\mu\nu} &:= R^\rho_{\mu\rho\nu} && (1.6)\\
R &:= g^{\mu\nu} R_{\mu\nu} && (1.7)\\
G_{\mu\nu} &:= R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R && (1.8)
\end{aligned}$$

The Einstein tensor is symmetric, i.e. $G_{\mu\nu} = G_{\nu\mu}$. It is also important to note that its divergence vanishes; $\nabla^\mu G_{\mu\nu} = 0$, which can be proven using the contracted Bianchi identity.

1.2.2 The FLRW universe

On very large scales, our universe is observed to be uniform in space (homogeneous) and not have a favoured direction (isotropicc). Spatial part of the homogeneous and isotropicc metric has constant curvature and can be categorised into three: spherical ($\mathbb{S}^3$), Euclidean ($\mathbb{E}^3$), and hyperbolic ($\mathbb{H}^3$). They are induced from embedding $\mathbb{R}^3$ into submanifolds of $\mathbb{R}^4$ equipped with the Euclidean metric, defined as $K|\mathbf{x}|^2 + u^2 = 1$. Here $K = 1, 0, -1$ for $\mathbb{S}^3$, $\mathbb{E}^3$, and $\mathbb{H}^3$, respectively. Writing the embedding as $f : x^i = (x,y,z) \mapsto X^I = (x,y,z,\sqrt{1-K(x^2+y^2+z^2)})$, the induced metric

$$\gamma_{ij} := \frac{\partial X^I}{\partial x^i}\frac{\partial X^J}{\partial x^j}\delta_{IJ} = \delta_{ij} + \frac{x_i x_j}{1 - K x_k x^k}. \tag{1.9}$$

[1] Consider a 1-parameter family of geodesics $\gamma(s,t)$, where t is an affine parameter. The geodesic deviation equation states $T^\rho \nabla_\rho (T^\nu \nabla_\nu S^\mu) = R^\mu_{\nu\rho\sigma} T^\nu T^\rho S^\sigma$, where tangent vectors $T = \partial/\partial t$, $S = \partial/\partial s$.

The spatial line element is therefore given by

$$dl^2 = \gamma_{ij}dx^i dx^j = d\mathbf{x}\cdot d\mathbf{x} + \frac{K(\mathbf{x}\cdot d\mathbf{x})^2}{1-K(\mathbf{x}\cdot\mathbf{x})} \tag{1.10}$$

$$= \frac{1}{1-Kr^2}dr^2 + r^2 d\Omega^2, \tag{1.11}$$

where the angular line element $d\Omega^2 = d\theta^2 + \sin^2\theta d\phi^2$.

We may now write down the form of the metric describing our universe in large scales;

$$ds^2 = -dt^2 + a(t)^2\left(\frac{1}{1-Kr^2}dr^2 + r^2 d\Omega^2\right). \tag{1.12}$$

This is known as the FLRW metric, named after independent researchers who worked on the topicc. Function $a(t)$ is called the scale factor and it dictates the growth of universe over time. Note that the metric is invariant under rescaling $a \to \lambda a$, $r \to r/\lambda$, and $K \to k := \lambda^2 K$. Hence we may set the scale factor to be $a(t_0) = 1$ at present time, at the cost of replacing $K \in \{-1,0,1\}$ by $k \in \mathbb{R}$.

Levi-Civita connection corresponding to the FLRW metric can be computed using the definition (1.3). Its non-zero components are given as follows.

$$\Gamma^0_{ij} = \frac{\dot{a}}{a}\gamma_{ij}, \tag{1.13}$$

$$\Gamma^i_{j0} = \Gamma^i_{0j} = \frac{\dot{a}}{a}\delta^i_j, \tag{1.14}$$

$$\Gamma^i_{jk} = \frac{1}{2a^2}\gamma^{il}\left(\partial_k\gamma_{jl} + \partial_j\gamma_{kl} - \partial_l\gamma_{jk}\right). \tag{1.15}$$

Overdot denotes time derivative $(\dot{\ }) := \partial/\partial t$ here and for the rest of this thesis. Indices for γ are raised and lowered using γ, not g.

Note that a path defined by $t(\tau) = \tau$ and $\mathbf{x}(\tau) = const$ is a timelike geodesic satisfying the geodesic equations (1.2). *Comoving* observers who follow these paths continue to perceive the expanding universe to be isotropic. Meanwhile, they find themselves drift apart, as the physical distance $r_{phys} = a(t)r$ grows in time.

Ricci curvature and Einstein tensor of the FLRW metric follows from definitions (1.5-1.8);

$$R_{00} = -\frac{3\ddot{a}}{a} \tag{1.16}$$

$$R_{ij} = \left[\frac{\ddot{a}}{a} + 2\left(\frac{\dot{a}}{a}\right)^2 + \frac{2k}{a^2}\right] a^2\gamma_{ij}, \tag{1.17}$$

$$R = 6\left[\frac{\ddot{a}}{a} + \left(\frac{\dot{a}}{a}\right)^2 + \frac{k}{a^2}\right], \tag{1.18}$$

$$G_{00} = 3\left[\left(\frac{\dot{a}}{a}\right)^2 + \frac{k}{a^2}\right], \tag{1.19}$$

$$G_{ij} = \left[-\frac{2\ddot{a}}{a} - \left(\frac{\dot{a}}{a}\right)^2 - \frac{k}{a^2}\right] a^2\gamma_{ij}. \tag{1.20}$$

While deriving (1.17) we used the fact that the Ricci tensor of three-dimensional spatial metric γ is equal to $2k\gamma_{ij}$. [2] Also note that components G_{0i} vanish and $G_{ij} \propto g_{ij}$, which is expected for a spatially homogeneous and isotropic spacetime.

1.2.3 Cosmic inventory

According to general relativity, spacetime is curved by its contents. Particles interact with gravity through the energy-momentum tensor $T_{\mu\nu}$, which encapsulates their energy, momentum flux, and stress. Thanks to spatial homogeneity and isotropy, components of the homogenous universe can be modelled as *perfect* fluids; they are completely characterised by their rest frame energy density and isotropic pressure. Defining the 4-velocity to be $U^\mu = dx^\mu/ds$ [3], the energy-momentum tensor of a perfect fluid is given by

$$T_{\mu\nu} = (\rho + P)U_\mu U_\nu + Pg_{\mu\nu}, \tag{1.21}$$

where the energy density ρ and pressure P only depends on time. For an observer comoving with the fluid, $T = \mathrm{diag}\{\rho, P, P, P\}$.

[2] In general, the Ricci tensor of any n-dimensional constant-curvature space with metric g_{ij} is given by $R_{ij} = (n-1)\kappa g_{ij}$. Here κ denotes sectional curvature of the space, which is equal to k for γ defined in (1.9).

[3] Here s is an affine parametrisation of geodesic followed by the fluid. It is equal to proper time τ for massive particles geodesics, and $g_{\mu\nu}U^\mu U^\nu = -1$. Massless particles such as photons follow null trajectory, and s is chosen so that $g_{\mu\nu}U^\mu U^\nu = 0$.

Divergence of the energy-momentum tensor vanishes for a perfect fluid;

$$\nabla_\nu T^{\mu\nu} = (\rho + P)\nabla_\nu \left(U^\mu U^\nu\right) + P\nabla_\mu g^{\mu\nu} \tag{1.22}$$
$$= (\rho + P)\left(U^\mu \nabla_\nu U^\nu + U^\nu \nabla_\mu U^\nu\right) + \nabla_\mu g^{\mu\nu} \tag{1.23}$$
$$= 0. \tag{1.24}$$

The first term in (1.23) vanishes because the fluid's 4-velocity satisfies the geodesic equation (1.2). Incompressibility implies that $\nabla_\nu U^\nu = 0$, and we see the second term also vanishes. The last term is zero since the Levi-Civita connections are *metric*; i.e., $\nabla g = 0$.

Using connections of the FLRW metric computed in the previous section, the $\mu = 0$ component of (1.24) yields the continuity equation;

$$\dot{\rho} + \frac{3\dot{a}}{a}(\rho + P) = 0. \tag{1.25}$$

Further imposing a constant equation of state $w = P/\rho$,

$$\frac{\dot{\rho}}{\rho} + 3(1+w)\frac{\dot{a}}{a} = 0, \tag{1.26}$$
$$\rho \propto a^{-3(1+w)}. \tag{1.27}$$

The universe contains a number of different components, but all known particles can be broadly categorised into three: radiation, matter, and dark energy.

Radiation consists of photons and neutrinos. The energy-momentum tensor of radiation is traceless, fixing the equation of state to be $w = 1/3$. While the number density of photons decrease as $\propto a^{-3}$, their energy density scales as $\propto a^{-4}$ instead because their wavelength gets stretched out as the universe expands. We define the *redshift* z to quantify this effect;

$$1 + z := \frac{1}{a}. \tag{1.28}$$

Hence, a photon with original wavelength λ_0 gets redshifted by $\Delta\lambda = z\lambda_0$. Note that the redshift is directly related to the scale factor $a(t)$. It can be used to parametrise time, as well as distance to a light source. Neutrinos show similar behaviours to photons since they remain ultra-relativistic.

Matter includes cold dark matter, electrons and protons. The latter two are often grouped as baryons, even though electrons are not technically baryonic. Pressure from non-relativistic matter is negligible, and $w = 0$. Their number density scales $\propto a^{-3}$ as the universe expands, and so does their energy density. Cold dark matter constitutes about 85% of the matter and a significant proportion of the total energy density today.

Dark energy is perhaps the most mysterious of the three, despite having the largest contribution to the total energy density at present time. First evidence of its existence came from Type Ia supernovae measurements which implied that the universe's expansion is accelerating. Subsequent observations of CMB and baryonic acoustic oscillations provided further proof. Exact physical mechanism for dark energy is not yet known, but potential explanations include the cosmological constant and quintessence. For purposes of the ΛCDM model, dark energy has negative pressure ($w = -1$), hence its energy density is independent of the scale factor.

Table 1.1 summarises the species of the universe and their properties discussed above. The last column contains fractional energy density of particles and shows how abundant each species are at present time. Precise definition of fractional density is to follow in the next section.

Table 1.1 Cosmic inventory. The fractional density values are quoted from Planck CMB analysis [39].

	Examples	Equation of State	Density Growth	Fractional Density Toda
Radiation (r)	Photon (γ)	$w = 1/3$	$\rho \propto a^{-4}$	$\Omega_\gamma \approx 1 \times 10^{-4}$
	Neutrino (ν)			$\Omega_\nu < 2 \times 10^{-2}$
Matter (m)	Cold dark matter (c)	$w = 0$	$\rho \propto a^{-3}$	$\Omega_c \approx 0.27$
	Baryon (b)			$\Omega_b \approx 0.05$
Dark Energy (Λ)		$w = -1$	$\rho = \text{const}$	$\Omega_\Lambda \approx 0.68$

1.2.4 Evolution of the universe

We are now ready to calculate the time evolution of the homogeneous universe. The Einstein field equation of general relativity reads

$$G_{\mu\nu} = 8\pi G T_{\mu\nu}, \tag{1.29}$$

where G is the Newtonian constant of gravitation. [4] Substituting in the Einstein tensor for the FLRW metric from (1.19-1.20) and the energy-momentum tensor from (1.21), we obtain

$$3\left[\left(\frac{\dot{a}}{a}\right)+\frac{k}{a^2}\right] =8\pi G\rho, \tag{1.30}$$

$$-\frac{2\ddot{a}}{a}-\left(\frac{\dot{a}}{a}\right)^2-\frac{k}{a^2} =8\pi GP. \tag{1.31}$$

Rearranging above yields the Friedmann equations;

$$\left(\frac{\dot{a}}{a}\right)^2 =\frac{8\pi G}{3}\rho-\frac{k}{a^2}, \tag{1.32}$$

$$\frac{\ddot{a}}{a} =-\frac{4\pi G}{3}(\rho+3P). \tag{1.33}$$

Note that the continuity equation (1.25) can be obtained from (1.33) and the time derivative of (1.32).

The Hubble parameter is defined as $H := \dot{a}/a$. From (1.32) we may compute the critical energy density for which the curvature k vanishes;

$$\rho_{\text{crit},0} := \frac{3H_0^2}{8\pi G}. \tag{1.34}$$

Subscripts 0 indicate that they are evaluated at present time $t = t_0$, where $a(t_0) = 1$.

In reality, energy density and pressure appearing in the Friedmann equations are sums of contributions from different fluid components. Fractional density of a given fluid X is defined as

$$\Omega_X := \frac{\rho_X}{\rho_{\text{crit},0}}. \tag{1.35}$$

In previous section we derived how each fluid's energy density depends on the scale factor. Quoting results summarised in Table 1.1, the Friedmann equations can be rewritten as follows;

$$\dot{a}^2 =H_0^2\left(\frac{\Omega_{r,0}}{a^2}+\frac{\Omega_{m,0}}{a}+\Omega_{\Lambda,0}\,a^2\right)-k, \tag{1.36}$$

$$\ddot{a} =H_0^2\left(-\frac{\Omega_{r,0}}{a^3}-\frac{\Omega_{m,0}}{2a^2}+\Omega_{\Lambda,0}\,a\right). \tag{1.37}$$

[4] Constant G is not to be confused with the Einstein tensor $G_{\mu\nu}$ on the left hand side.

These equations dictate the growth (or shrinking) of the universe given curvature and energy density composition today. For $k <= 0$, the right hand side of (1.36) is always positive regardless of fractional density value. In this case, the fact that universe is currently expanding suffices to show that scale factor a has been increasing monotonically. The universe began with the Big Bang at $a = 0$.

When $k > 0$, $\dot{a}$ vanishes at one or two values of a. There are multiple scenarios in this case, including the Einstein's Static Universe (ESU) where $\dot{a} = \ddot{a} = 0$. The ESU is however unstable; perturbing around the static solution as $a(t) = a_{\text{ESU}}(1 + \xi(t))$ in (1.37) gives $\ddot{\xi} > 0$ to leading order, which implies that there exists a growing solution for ξ. In fact, there are no stable static solutions to the Friedmann equations, rendering such models implausible. Another possibility is a closed universe, where the scale factor grows until it hits the maximum and then decreases. This scenario requires $\ddot{a} < 0$ at all times. Measurements of Type Ia supernovae strongly suggest that the universe is in an accelerating phase, ruling out this option as well. Lastly, there is the bouncing universe model. The scale factor starts large, drops to a minimum value, and bounces back to an accelerating growth. This model has been disregarded due to the need for introduction of various new physics in the early universe, but has recently regained popularity as an alternative to inflation.

Constraints from modern cosmological observations indicate that our universe is extremely flat, with $k \approx 0$. For the rest of this thesis we set the curvature $k = 0$ and follow a standard Big Bang theory.

Note the different powers of the scale factor a are associated with each component of the universe in (1.36). According to the CMB measurements $\Omega_{r,0} \ll \Omega_{m,0} < \Omega_{\Lambda,0}$. The energy density is therefore dominated by a single component at a time, resulting in three different eras:

- Radiation domination (RD) where $0 < a < \Omega_{r,0}/\Omega_{m,0}$,
- Matter domination (MD) with $\Omega_{r,0}/\Omega_{m,0} < a < (\Omega_{m,0}/\Omega_{\Lambda,0})^{1/3}$, and
- Dark energy domination (ΛD) for $a > (\Omega_{m,0}/\Omega_{\Lambda,0})^{1/3}$.

When the universe consists mainly of a single fluid component X, we can simplify the Friedmann equations as follows;

$$\dot{a}^2 = H_0^2 \Omega_{X,0}\, a^{-1-3w_X}, \tag{1.38}$$

$$a \propto t^{\frac{2}{3(1+w_X)}}, \quad \text{if } w_X \neq -1. \tag{1.39}$$

Note that the scale factor grows exponentially for ΛD as $w_\Lambda = -1$; $a \propto \exp(Ht)$.

It is often convenient to consider *conformal* time defined as $\tau := \int_{t_i}^{t}(1/a(t'))dt'$, for some initial time reference t_i. As $dt = ad\tau$, the flat FLRW metric is given in terms of conformal time by

$$ds^2 = -dt^2 + a(t)^2 d\mathbf{x}^2 = a(t)^2 \left[-d\tau^2 + d\mathbf{x}^2\right]. \tag{1.40}$$

Rewriting (1.38) with $da/d\tau = a\dot{a}$, we obtain

$$a \propto \tau^{\frac{2}{3w_X+1}}. \tag{1.41}$$

Calculations in this section are summarised in Table 1.2.

Table 1.2 Evolution of the universe. Ranges for the scale factor are computed from the cosmological parameters estimated in [39].

Era	Scale Factor	Growth (comoving)	Growth (conformal)
Radiation Domination (RD)	$a < 2.9 \times 10^{-4}$	$a \propto t^{1/2}$	$a \propto \tau$
Matter Domination (MD)	$2.9 \times 10^{-4} < a < 0.77$	$a \propto t^{2/3}$	$a \propto \tau^2$
Dark Energy Domination (ΛD)	$a > 0.77$	$a \propto e^{Ht}$	$a \propto -1/\tau$

1.3 Inflation

Soon after the quantitative formulation of Big Bang cosmology, several issues with the initial conditions were raised. The CMB was observed to be nearly homogeneous, even though many parts of it should have been causally disconnected at the time. Curvature of the universe is extremely close to zero, while $k = 0$ is an unstable stationary point. The standard Big Bang cosmology provided little justification for such initial smoothness and fine-tuning.

The theory of cosmic inflation not only resolved most of these problems successfully, but also provided a physical mechanism for the generation of inhomogeneities in the universe. Quantum fluctuations of the inflationary field seeds the initial conditions, whose statistical properties are consistent with current observations. Inflation has therefore become the most widely accepted theory of the early universe to date.

In this section we formulate the puzzles which led to the introduction of inflation (1.3.1) and outline the basic inflationary paradigm (1.3.2).

1.3.1 The horizon problem

According to relativity, information cannot travel faster than the speed of light. It is therefore possible to have two different points in spacetime that are causally disconnected; their lightcones do not intersect, so no events since the Big Bang could have affected both. Physical properties at such two points are independent of each other. The aim of this section is to compute the size of causally disjoint regions at the epoch of recombination, when most of the CMB photons start free-streaming.

Consider a photon travelling in a straight line. Writing the radial part of $d\mathbf{x}^2$ as $d\chi^2$, our spacetime metric (1.40) becomes

$$ds^2 = a(\tau)^2(-d\tau^2 + d\chi^2). \tag{1.42}$$

Photons follow null geodesics, meaning $ds^2 = 0$ along their trajectories. Thus $\chi(\tau) = \tau + const$ or $\chi(\tau) = -\tau + const$. They appear to be straight and diagonal lines on the χ-τ plane, as shown in Figure 1.1. The distance light travels starting from some initial time τ_i to τ_f is then given by

$$\chi_{\mathrm{PH}} := \tau_f - \tau_i = \int_{a_i}^{a_f} \frac{d\tau}{da} da = \int_{a_i}^{a_f} \frac{1}{a\dot{a}} da. \tag{1.43}$$

Here, χ_{PH} is called the particle horizon. No particles could have travelled further than this distance since the initial time τ_i.

Suppose that the universe is dominated by a single perfect fluid X in between a_i and a_f. The simplified Friedmann equation (1.38) then gives

$$\chi_{\mathrm{PH}} = \int_{a_i}^{a_f} \frac{a^{(3w_X-1)/2}}{H_0\sqrt{\Omega_{X,0}}} da = \frac{2}{(3w_X+1)H_0\sqrt{\Omega_{X,0}}} \left(a_f^{(3w_X+1)/2} - a_i^{(3w_X+1)/2} \right). \tag{1.44}$$

Note that the particle horizon is bounded as $a_i \to 0$ if and only if $3w_X + 1 > 0$.

According to the conventional big bang cosmology, the universe begins at $a_i = 0$ with radiation contributing the most to energy density. The CMB last scattering surface lies around redshift $z \sim 1090$ (or $a_{\mathrm{rec}} = 9.17 \times 10^{-4}$) in matter domination era. Quoting cosmological parameters from tables 1.1 and 1.2, as well as treating radiation and matter domination separately in the integral, we get $\chi_{\mathrm{CMB}} \approx 340$ Mpc.

Conformal distance to the last scattering surface χ_* can also be computed using the same formula (1.44), now integrating from a_{rec} to $a_0 = 1$. Approximating again by separating matter and dark energy domination era, $\chi_* \approx 15000$ Mpc. This is much larger than χ_{PH}!

As shown in Figure 1.1, CMB photons at two different opposite sides of the sky had no causal contact at all; their particle horizons have zero overlap. Furthermore, $\chi_{\text{CMB}}/\chi_{*} \approx 1.3$ degrees. Every disjoint 1.3 deg patch in the sky were causally unrelated at the time of recombination. There is no obvious reason for cosmological parameters in these patches to be similar. Despite this fact, the observed CMB is isotropic everywhere in temperature to order $O(10^{-5})$. This is the horizon problem; the universe at recombination is too homogeneous considering the small particle horizon then.

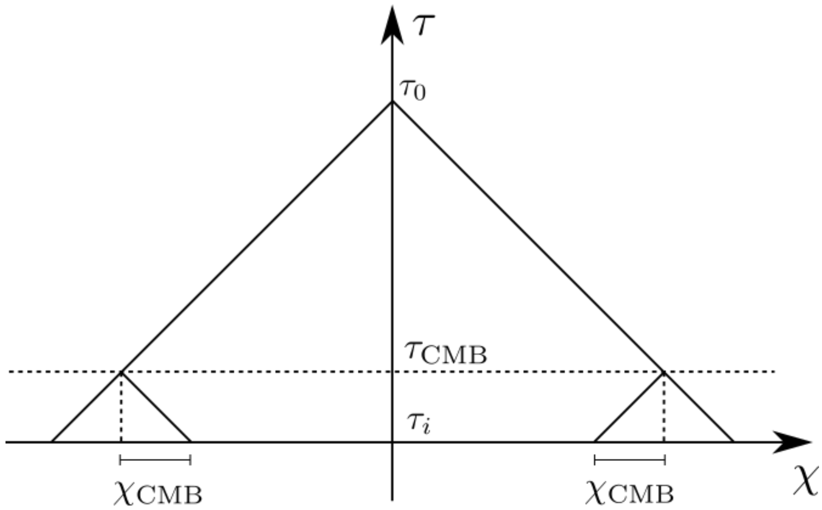

Fig. 1.1 The horizon problem. According to the conventional big bang cosmology, different regions of the CMB we observe today have had no overlap in their particle horizon. Yet, the CMB is measured close to uniform everywhere.

To resolve this issue, we need the particle horizon at recombination to be larger. The theory of cosmic inflation achieves this by having a period in the early universe where $3w+1<0$. In this case, we see from (1.44) that χ_{PH} is unbounded as $a_i \to 0$. The initial conformal time $\tau_i \to -\inf$, allowing enough proper time for the particle horizon growth. Then even the two opposite regions of the CMB we see can have had causal contact in the past, as depicted in Figure 1.2.

Inflation can alternatively be characterised using the comoving hubble radius defined as

$$\mathscr{H}^{-1} := \frac{1}{aH}. \tag{1.45}$$

Note that the particle horizon can then be expressed in terms of the comoving hubble radius as

$$\chi_{\text{PH}} = \int_{a_i}^{a_f} \frac{1}{a\dot{a}} da = \int_{\ln a_i}^{\ln a_f} \mathscr{H}^{-1} d\ln a. \tag{1.46}$$

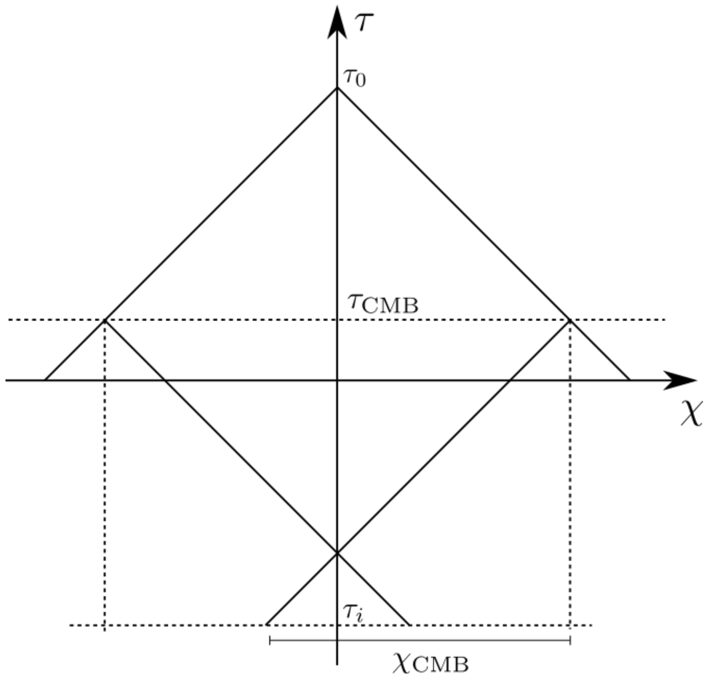

Fig. 1.2 Solution to the horizon problem. Inflation allows more conformal time for different regions to have been in causal contact before recombination.

The particle horizon represents the distance where objects could have *ever* talked to each other. On the other hand, the comoving hubble radius is a scale for how far information can reach *now*. [5] As can be seen from (1.44), $\mathscr{H}^{-1} \propto a^{(3w_X+1)/2}$, which is $\propto a$ for radiation domination and $\propto a^{1/2}$ during matter domination.

Inflation explains homogeneity of the observed CMB by requiring $\mathscr{H}^{-1}$ to have shrunk rapidly in early times; $d\mathscr{H}^{-1}/d\ln a < 0$. We find that $\ddot{a} > 0$, so the universe undergoes accelerated expansion. Steep decrease in $\mathscr{H}^{-1}$ is expressed using a 'slow-roll' parameter

$$\varepsilon := -\frac{d\ln H}{d\ln a} = -\frac{\dot{H}}{H^2} \ll 1. \tag{1.47}$$

Inflation also needs to last long enough for the particle horizon to grow sufficiently large. We define another parameter to denote this constraint;

$$\eta := \frac{d\ln \varepsilon}{d\ln a} = \frac{\dot{\varepsilon}}{\varepsilon H} \ll 1. \tag{1.48}$$

[5] Two points $\mathscr{H}^{-1}$ apart drifts away with relative physical velocity $v_{\text{phys}} = \dot{a}\mathscr{H}^{-1} = 1$, which is equal to c in our units. It is difficult for such points to have causal interaction right now, especially within Hubble time H^{-1}.

1.3.2 Slow-roll inflation

The simplest model of inflation consists of a single scalar field ϕ. The action for a real scalar field with canonical kinetic term and potential $V(\phi)$ is given by

$$S_\phi = \int dt d^3\mathbf{x}\, \sqrt{-g}\left[-\frac{1}{2}g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - V(\phi)\right], \tag{1.49}$$

where $g := \det g$. Denoting the integrand as the Lagrangian density $\mathscr{L}[t,\mathbf{x},\phi,\partial_\mu\phi]$, the energy-momentum tensor can be expressed using functional derivatives as

$$T_{\mu\nu} = \frac{-2}{\sqrt{-g}}\frac{\delta\mathscr{L}}{\delta g^{\mu\nu}} = -2\frac{\delta\mathscr{L}}{\delta g^{\mu\nu}} + g_{\mu\nu}\mathscr{L}. \tag{1.50}$$

Here, we used the identity $\delta\sqrt{-g} = -(1/2)\sqrt{-g}g_{\mu\nu}\delta g^{\mu\nu}$. Substituting (1.49),

$$T_{\mu\nu} = \left(\partial_\mu\phi\right)\left(\partial_\nu\phi\right) + g_{\mu\nu}\left(-\frac{1}{2}g^{\rho\sigma}\partial_\rho\phi\partial_\sigma\phi - V(\phi)\right). \tag{1.51}$$

Now suppose that ϕ drives inflation in the FLRW background. Due to the symmetries present in homogeneous and isotropic metric, the inflation field can only depend on time; $\phi(\mathbf{x},t) = \bar{\phi}(t)$. We can read off the inflation field's energy density and pressure from the energy-momentum tensor.

$$\rho_{\bar{\phi}} = -T^0_{\ 0} = \frac{1}{2}\dot{\bar{\phi}}^2 + V(\bar{\phi}), \tag{1.52}$$

$$P_{\bar{\phi}}\delta^i_j = T^i_{\ j} = \delta^i_j\left(\frac{1}{2}\dot{\bar{\phi}}^2 - V(\bar{\phi})\right). \tag{1.53}$$

The two terms $\frac{1}{2}\dot{\bar{\phi}}^2$ and $V(\bar{\phi})$ can be interpreted as the kinetic and potential energy of inflation field, respectively. The equation of state is also expressed in terms of the two;

$$w_{\bar{\phi}} = \frac{P_{\bar{\phi}}}{\rho_{\bar{\phi}}} = \frac{\frac{1}{2}\dot{\bar{\phi}}^2 + V(\bar{\phi})}{\frac{1}{2}\dot{\bar{\phi}}^2 - V(\bar{\phi})}. \tag{1.54}$$

It follows that $\bar{\phi}$ satisfies the condition $3w_{\bar{\phi}} + 1 < 0$ required for inflation, as long as the potential energy dominates over kinetic energy.

The classical equations of motion follows from the Euler-Lagrange equation. After some calculations we obtain

$$\ddot{\bar{\phi}} + 3H\dot{\bar{\phi}} = -V'(\bar{\phi}). \tag{1.55}$$

Even though $\bar{\phi}$ is a field, its dynamics given in (1.55) are identical to those of a particle rolling down an one-dimensional potential. Its movement should be slow for the kinetic energy to be much smaller than the potential, and hence the name 'slow-roll' inflation.

The Friedmann equations (1.32-1.33) can now be expressed in terms of the background inflationary field.

$$H^2 = \frac{8\pi G}{3}\left(\frac{1}{2}\dot{\bar{\phi}}^2 + V(\bar{\phi})\right) \approx \frac{8\pi G}{3} V(\bar{\phi}), \tag{1.56}$$

$$\frac{\ddot{a}}{a} = -\frac{8\pi G}{3}\left(\dot{\bar{\phi}}^2 - \frac{1}{2}V(\bar{\phi})\right) \approx \frac{8\pi G}{3} V(\bar{\phi}), \tag{1.57}$$

where the slow-roll approximations have been used in the last step. Taking derivative of (1.56) gives us $3H\dot{\bar{\phi}} \approx -V'$. The parameters defined as

$$\varepsilon_V := \frac{1}{16\pi G}\left(\frac{V'}{V}\right)^2, \quad \eta_V := \frac{1}{8\pi G}\frac{V''}{V}, \tag{1.58}$$

can then shown to be small. Therefore the slow roll parameters from (1.47-1.48) satisfy

$$\varepsilon \approx \varepsilon_V \ll 1, \quad \eta \approx -2\eta_V + 4\varepsilon_V \ll 1. \tag{1.59}$$

We see that as long as the potential $V(\phi)$ is chosen such that $\varepsilon_V, \eta_V \ll 1$, the field ϕ can drive a period of accelerated expansion. Here, $\bar{\phi}(t)$ acts as a clock; it measures the progress of inflation, until ε eventually grows comparable to 1 and inflation ends.

1.3.3 Quantum fluctuations

So far our considerations on the inflation field ϕ has been entirely classical. Moving on to quantum theory, field values are no longer fixed at each point in spacetime. The goal of this section is to quantify statistical properties of these quantum fluctuations.

For simplicity, we assume that the metric remains unperturbed. In reality, gravity is coupled with perturbations of the inflation field, but this approximation still allows us to derive most of the crucial results. We further neglect any terms suppressed by slow-roll parameters.

We write the inflation field as a sum of the classical solution and perturbation;

$$\phi(\mathbf{x},t) = \bar{\phi}(t) + \frac{v(\mathbf{x},t)}{a(t)}, \tag{1.60}$$

where the factor of $1/a$ has been introduced for later convenience. First rewriting the scalar field action (1.49) in terms of conformal time,

$$S_\phi = \int d\tau d^3\mathbf{x} a(\tau)^2 \left[\frac{1}{2}(\phi')^2 - \frac{1}{2}(\nabla\phi)^2 - a(\tau)^2 V(\phi)\right]. \tag{1.61}$$

When we include perturbations, terms linear in v vanish from the equations of motion of $\bar{\phi}$. Further removing terms with derivatives of $V(\phi)$ using slow-roll condition,

$$\delta S_\phi = \int d\tau d^3\mathbf{x} \left[\frac{1}{2}\left(v' - \frac{a'}{a}v\right)^2 - \frac{1}{2}(\nabla v)^2\right] \tag{1.62}$$

$$= \int d\tau d^3\mathbf{x} \left[\frac{1}{2}(v')^2 + \frac{1}{2}\frac{a''}{a}v^2 - \frac{1}{2}(\nabla v)^2\right]. \tag{1.63}$$

Integration by parts has been used to obtain the last line. The equations of motion for v follows;

$$v'' - \frac{a''}{a}v - \nabla^2 v = 0. \tag{1.64}$$

Defining Fourier transforms of v as

$$v(\mathbf{x}, \tau) = \int \frac{d^3\mathbf{k}}{(2\pi)^3} e^{i\mathbf{k}\cdot\mathbf{x}} \tilde{v}(\mathbf{k}, \tau), \tag{1.65}$$

we obtain the Mukhanov-Sasaki equation;

$$v'' + (k^2 - \frac{a''}{a})v = 0 \tag{1.66}$$

Tildes above v have been omitted for brevity. Each $\mathbf{k}$ mode of the perturbative field $v(\mathbf{k}, \tau)$ evolves independently from each other. During slow-roll inflation $a \propto -1/\tau$, and $a''/a = 2/\tau^2$. General form of the solution is given by the *mode functions* $v_k(\tau)$.

$$v_k(\tau) = c_+ \left(1 - \frac{i}{k\tau}\right) e^{-ik\tau} + c_- \left(1 + \frac{i}{k\tau}\right) e^{ik\tau}. \tag{1.67}$$

We would like to canonically quantise the field $v(\mathbf{k},\tau)$. To achieve this goal, we first convert from the Lagrangian to Hamiltonian formalism.

$$\pi := \frac{\partial \mathscr{L}}{\partial v'} = v', \tag{1.68}$$

$$\mathscr{H} := v'\frac{\partial \mathscr{L}}{\partial v'} - \mathscr{L} = \frac{1}{2}\pi^2 + \frac{1}{2}(\nabla v)^2 - \frac{1}{2}\frac{a''}{a}v^2. \tag{1.69}$$

We now promote classical fields $v(\mathbf{k},\tau)$, $\pi(\mathbf{k},\tau)$ to operators $\hat{v}_{\mathbf{k}}(\tau)$, $\hat{\pi}_{\mathbf{k}}(\tau)$ satisfying equal-time commutation relations [6]

$$[\hat{v}_{\mathbf{k}_1}(\tau), \hat{\pi}_{\mathbf{k}_2}(\tau)] = (2\pi)^3\delta^{(3)}(\mathbf{k}_1+\mathbf{k}_2), \tag{1.70}$$

$$[\hat{v}_{\mathbf{k}_1}(\tau), \hat{v}_{\mathbf{k}_2}(\tau)] = [\hat{\pi}_{\mathbf{k}_1}(\tau), \hat{\pi}_{\mathbf{k}_2}(\tau)] = 0. \tag{1.71}$$

Defining operators $\hat{a}_{\mathbf{k}}$ and $\hat{a}^\dagger_{\mathbf{k}}$ appropriately, we may write

$$\hat{v}_{\mathbf{k}}(\tau) = v_k(\tau)\hat{a}_{\mathbf{k}} + v_k(\tau)^*\hat{a}^\dagger_{-\mathbf{k}}, \tag{1.72}$$

$$\hat{\pi}_{\mathbf{k}}(\tau) = v'_k(\tau)\hat{a}_{\mathbf{k}} + v'_k(\tau)^*\hat{a}^\dagger_{-\mathbf{k}}. \tag{1.73}$$

As long as we normalise the mode functions $v_k(\tau)$ so that its Wronskian $W := v_k v_k'^* - v_k^* v_k' = i$ (purely imaginary since v_k is real), we obtain

$$[\hat{v}_{\mathbf{k}_1}(\tau), \hat{\pi}_{\mathbf{k}_2}(\tau)] = i\left[\hat{a}_{\mathbf{k}_1}, \hat{a}^\dagger_{\mathbf{k}_2}\right], \tag{1.74}$$

$$\left[\hat{a}_{\mathbf{k}_1}, \hat{a}^\dagger_{\mathbf{k}_2}\right] = (2\pi)^3\delta^{(3)}(\mathbf{k}_1-\mathbf{k}_2). \tag{1.75}$$

Constructed $\hat{a}$ and $\hat{a}^\dagger$ are analogous to the creation and annihilation operators of quantum harmonic oscillator. Our next step is to compute the Hamiltonian operator;

$$\hat{H} = \int d^3\mathbf{x}\left[\frac{1}{2}\hat{\pi}^2 + \frac{1}{2}(\nabla\hat{v})^2 - \frac{1}{2}\frac{a''}{a}\hat{v}^2\right] \tag{1.76}$$

$$= \int d^3\mathbf{x}\frac{d^3\mathbf{k}_1}{(2\pi)^3}\frac{d^3\mathbf{k}_2}{(2\pi)^3}\frac{1}{2}e^{i(\mathbf{k}_1+\mathbf{k}_2)\cdot\mathbf{x}}\left[\hat{\pi}_{\mathbf{k}_1}\hat{\pi}_{\mathbf{k}_2} - (\mathbf{k_1}\cdot\mathbf{k_2})\hat{v}_{\mathbf{k}_1}\hat{v}_{\mathbf{k}_2} - \frac{a''}{a}\hat{v}_{\mathbf{k}_1}\hat{v}_{\mathbf{k}_2}\right] \tag{1.77}$$

$$= \int \frac{d^3\mathbf{k}}{(2\pi)^3}\frac{1}{2}\left[\hat{\pi}_{\mathbf{k}}\hat{\pi}_{-\mathbf{k}} + k^2\hat{v}_{\mathbf{k}}\hat{v}_{-\mathbf{k}} - \frac{a''}{a}\hat{v}_{\mathbf{k}}\hat{v}_{-\mathbf{k}}\right] \tag{1.78}$$

$$= \int \frac{d^3\mathbf{k}}{(2\pi)^3}\left[E_k\left(\hat{a}_{\mathbf{k}}\hat{a}^\dagger_{\mathbf{k}} + \hat{a}^\dagger_{-\mathbf{k}}\hat{a}_{-\mathbf{k}}\right) + F_k\hat{a}_{\mathbf{k}}\hat{a}_{-\mathbf{k}} + F_k^*\hat{a}^\dagger_{\mathbf{k}}\hat{a}^\dagger_{-\mathbf{k}}\right], \tag{1.79}$$

[6] This follows from its Fourier equivalent: $[\hat{v}(\mathbf{x}_1,\tau), \hat{\pi}(\mathbf{x}_2,\tau)] = i\delta^{(3)}(\mathbf{x}_1-\mathbf{x}_2)$.

where

$$\omega_k^2 := k^2 - \frac{a''}{a}, \quad E_k := \frac{1}{2}(|v_k'|^2 + \omega_k^2|v_k|^2), \quad F_k := \frac{1}{2}(v_k'^2 + \omega_k^2 v_k^2). \tag{1.80}$$

Defining the vacuum state

Note that we are currently in Heisenberg picture where the operators depend on time. Our mode functions have two degrees of freedom: c_+ and c_-. One of them has been fixed by the normalisation condition $W = 2ik(|c_+|^2 - |c_-|^2) = i$. The other degree of freedom remains, availing us one-parameter family of possible initial mode functions. Note that $\hat{a}_{\mathbf{k}}^\dagger$ are defined in terms of $\hat{v}_{\mathbf{k}}$ and $\hat{\pi}_{\mathbf{k}}$. Fixing $\hat{a}_{\mathbf{k}}$ is thus equivalent to choosing $\hat{v}_{\mathbf{k}}$.

We define the vacuum $|\mathbf{0}\rangle$ to be the state satisfying $\hat{a}_{\mathbf{k}}|\mathbf{0}\rangle = 0$ for all $\mathbf{k} \in \mathbb{R}^3$. The expected energy of the vacuum state is given by

$$\langle \mathbf{0}|\hat{H}|\mathbf{0}\rangle = \int \frac{d^3\mathbf{k}}{(2\pi)^3} E_k \langle \mathbf{0}|[\hat{a}_{\mathbf{k}}, \hat{a}_{\mathbf{k}}^\dagger]|\mathbf{0}\rangle \tag{1.81}$$

$$= \int d^3\mathbf{k} E_k \delta^{(3)}(0), \tag{1.82}$$

where the divergence $\delta^{(3)}(0)$ arises only because we are integrating over the whole space. Removing this factor, we may interpret E_k as the vacuum energy density for mode k.

We now ask the vacuum state to be a ground state of the Hamiltonian. Minimising the energy density E_k while keeping normalisation condition $W = i$, we obtain the *Bunch-Davies* mode function;

$$v_{\mathbf{k}}(\tau) = \frac{1}{\sqrt{2k}} \left(1 - \frac{i}{k\tau}\right) e^{-ik\tau}. \tag{1.83}$$

For this choice of mode function and vacuum, $E_k \to \hbar\omega_k/2$ as $\tau \to -\infty$. [7] This is analogous to the case of quantum harmonic oscillator.

Lastly, we compute the zero-point fluctuation of the inflation field. By definition,

$$\hat{v}(\mathbf{x}, \tau) = \int \frac{d^3\mathbf{k}}{(2\pi)^3} \left[v_k \hat{a}_{\mathbf{k}} + v_k^* \hat{a}_{\mathbf{k}}^\dagger\right] e^{i\mathbf{k}\cdot\mathbf{x}}, \tag{1.84}$$

[7] $\hbar$ has been reinstated here for clarity.

and we have

$$\langle|\hat{v}(\mathbf{x},\tau)|^2\rangle = \int \frac{d^3\mathbf{k}_1}{(2\pi)^3}\frac{d^3\mathbf{k}_2}{(2\pi)^3}\langle\mathbf{0}|v_{k_1}\hat{a}_{\mathbf{k}_1}v^*_{k_2}\hat{a}^\dagger_{\mathbf{k}_2}|\mathbf{0}\rangle \tag{1.85}$$

$$= \int \frac{d^3\mathbf{k}_1}{(2\pi)^3}\frac{d^3\mathbf{k}_2}{(2\pi)^3}v_{k_1}v^*_{k_2}\langle\mathbf{0}|\left[\hat{a}_{\mathbf{k}_1},\hat{a}^\dagger_{\mathbf{k}_2}\right]|\mathbf{0}\rangle \tag{1.86}$$

$$= \int \frac{d^3\mathbf{k}}{(2\pi)^3}|v_k|^2 \tag{1.87}$$

$$= \int d(\ln k)P_v(k,\tau). \tag{1.88}$$

The dimensionless power spectrum is defined as

$$P_v(k,\tau) := \frac{k^3}{2\pi^2}|v_k(\tau)|^2. \tag{1.89}$$

Recall that perturbations in ϕ is given by $\delta\phi = v/a$ from (1.60). The *dimensionless power spectrum* for fluctuations in ϕ is therefore

$$P_{\delta\phi}(k,\tau) = \frac{P_v(k,\tau)}{a(\tau)^2} = \left(\frac{H}{2\pi}\right)^2\left(1+(k\tau)^2\right), \tag{1.90}$$

where we used the fact that $a(\tau) = -1/H\tau$ during slow-roll inflation. For scales larger than the comoving Hubble radius we have $k\tau \ll 1$. In this limit, $P_{\delta\phi} \to (H/2\pi)^2$ which is nearly constant. This is a key prediction from our simplistic model of inflation; we expect a near scale-invariant power spectrum of perturbations.

Chapter 2

Cosmic Microwave Background Anisotropy

2.1 Cosmic microwave background

2.2 The inhomogeneous universe

2.3 CMB power spectrum

2.4 CMB polarisation

Chapter 3

Bispectrum and Primordial Non-Gaussianity

3.1 Bispectrum

3.2 Primordial non-Gaussianity

Chapter 4

CMB Stage-4 Forecast

4.1 Abstract

We present forecasts on the primordial non-Gaussianity parameter $f_{\rm NL}$ of feature models for the future Cosmic Microwave Background Stage-4 (CMB-S4) experiments. The Fisher matrix of the bispectrum estimator was computed using noise covariances expected for preliminary CMB-S4 specifications including ones for the Simons Observatory. We introduce a novel method that improves the computation by orthonormalising the covariance matrix. The most sensitive CMB-S4 experiment with 1' beam and $1\mu K$-arcmin noise would yield a factor of 1.7-2.2 times more stringent constraints compared to Planck. Under the Simons Observatory baseline conditions the improvement would be about 1.3-1.6 times to Planck. We also thoroughly studied the effects of various model and experimental parameters on the forecast. Detailed analysis on the constraints coming from temperature and E-mode polarisation, in particular, provided some insight into detecting oscillatory features in the CMB bispectrum.

4.2 Introduction

The Cosmic Microwave Background (CMB) radiation is one of our most valuable probes of the primordial universe. The temperature and polarisation of this ancient light contains rich statistical information both about the primordial perturbations created during inflation and also their subsequent evolution until now. This allows us to test our inflationary theories and also the history of our universe. The recent Planck CMB experiments have provided stringent tests on various models of inflation through the estimation of cosmological parameters and via primordial non-Gaussianity [37, 38].

The simplest model of inflation involves a single scalar field slowly rolling down a smooth potential. In this case the CMB temperature fluctuations are expected to be Gaussian distributed with only tiny deviations (e.g. [27]). However, many other physically well-motivated models generate larger non-Gaussian signatures at the end of inflation (see reviews of [9]). Such primordial non-Gaussianities are well constrained by three-point correlation functions of the CMB anisotropies or their Fourier transform, the CMB bispectrum. Different inflationary models predict bispectra with different momentum dependence, or 'shapes'. We constrain these models by using an optimal estimator for their amplitude parameter, f_{NL}, for each specific bispectrum shape (see, e.g, [23, 26] for reviews).

Although all observations to date are consistent with vanishing non-Gaussianity, the models most favoured by the latest Planck CMB analysis were the ones with oscillations in the primordial power spectrum [37]. Among them are feature models, where the oscillations are caused by a sharp feature in either the inflationary potential [41, 2, 11, 3, 21, 14], sound speed [33, 5], or multi-field potentials [1] (see [9, 12] for reviews). The primordial power spectrum then becomes scale dependent, displaying sinusoidal oscillations that are linearly spaced in momentum space. The resulting bispectrum also oscillates and is highly uncorrelated with other popular bispectrum templates [31], therefore allowing us to constrain them independently.

Planck constrained f_{NL} for feature models from CMB bispectra, but no signal above 3σ significance were found after accounting for the 'look elsewhere effect' as introduced in [17]. The multi-peak statistic analysis, however, revealed some non-standard signals up to 4σ level that deserves attention [37]. There have been many other searches on signatures of oscillations. Constraints also come from the CMB power spectrum [28, 7, 32, 29, 30, 18], the large scale structure [8, 4], and a combination of the two [22, 6]. We expect stronger constraints on feature models from future LSS experiments [10]. This paper covers the prospects of upcoming CMB experiments in constraining f_{NL} for feature models.

Currently there are two implementations of the optimal estimator for constraining f_{NL} for feature type models. The Planck analysis adopted the Modal estimator for which the given bispectrum is expanded using a separable basis [19, 16]. This method is efficient and can flexibly account for various oscillatory shapes and can easily constrain all frequencies simultaneously. However, when the oscillation frequency is large the modal basis fails to converge within reasonable number of basis elements, making the method impractical. The other approach using the Komatasu-Spergel-Wandelt (KSW) estimator is viable for various shapes including the feature model [24, 34]. Although this method only applies to models with separable bispectra, even highly oscillatory templates can reliably be computed. This method is however more expensive as each frequency must be dealt with separately. We present further optimisations to the fast KSW estimator introduced in [44] and apply it on feature models for forecasts in this paper.

The next generation of CMB experiments, CMB Stage-4, consists of many exciting proposed experiments located at the South Pole, the Atacama Desert in Chile, and perhaps space [15, 43, 42]. One of the main goals of these experiments is to measure the polarisation signal in the CMB to the cosmic variance limit. Preliminary specifications have been released for these experiments [15, 43] and these have been used to produce some forecasts for the standard f_{NL} templates but not yet for feature type models. In this paper we address this by presenting the Fisher forecasts on f_{NL} for feature models based on these specifications and observe that feature type models receive larger improvements from the extra polarisation information than the standard templates justifying this analysis.

The paper is organised as follows. First we briefly review the theory of CMB bispectrum in Section 4.3. Bispectrum template for the feature model is defined and computed here. In Section 4.4 we formulate the bispectrum estimator and introduce a new method to further optimise its computation. The technique is applied to the case of feature model to yield equations for the Fisher forecast of f_{NL}. We also briefly discuss implementation details. In Section 4.5 we present our forecast results and their dependence on model and experimental parameters. In particular, forecasts for the Simons observatory are compared with the Planck results. The results are summarised in Section 4.6.

4.3 Feature model bispectrum

4.3.1 CMB bispectrum

One of the main subjects of primordial NG studies is the 3-point correlation function of the primordial perturbations which is defined by;

$$\langle \Phi(\mathbf{k}_1)\Phi(\mathbf{k}_2)\Phi(\mathbf{k}_3)\rangle = (2\pi)^3\delta^{(3)}(\mathbf{k}_1+\mathbf{k}_2+\mathbf{k}_3)B_\Phi(k_1,k_2,k_3), \tag{4.1}$$

where we have assumed statistical homogeneity and isotropy. The primordial bispectrum B_Φ vanishes for Gaussian perturbations, but more general inflation models predict non-zero bispectra with various shapes. In order to constrain these models we re-parameterise the bispectrum into a amplitude parameter and a normalised shape part;

$$B_\Phi(k_1,k_2,k_3) = f_{NL}B_\Phi^{(f_{NL}=1)}(k_1,k_2,k_3). \tag{4.2}$$

Constraining f_{NL} from the CMB measurements allows to determine how well the particular shape under consideration aligns with the data which we can translate into constrains on the model itself.

In order to compare the theory with measurements we first need to relate the primordial perturbations to spherical multipole modes of the late-time CMB anisotropies.

$$a^X_{lm} = 4\pi(-i)^l \int \frac{d^3\mathbf{k}}{(2\pi)^3}\Phi(\mathbf{k})\Delta^X_l(k)Y_{lm}(\hat{\mathbf{k}}). \tag{4.3}$$

Here the index X is either T or E, representing CMB temperature and E-mode polarisation, respectively. The linear CMB radiation transfer function $\Delta^X_l(k)$ can be computed from the Boltzmann solvers like CAMB [25].

Three point correlation function of a^X_{lm}'s yield the reduced bispectrum $b_{l_1l_2l_3}$ times a geometrical factor $\mathscr{G}^{l_1l_2l_3}_{m_1m_2m_3}$ named the Gaunt integral. After some algebraic manipulations we obtain the following useful formula for the reduced bispectrum;

$$b^{X_1X_2X_3}_{l_1l_2l_3} = \left(\frac{2}{\pi}\right)^3 \int_0^\infty r^2 dr \int_{\mathscr{V}_k} d^3\mathbf{k}\ (k_1k_2k_3)^2 B_\Phi(k_1,k_2,k_3) \prod_{i=1}^{3}\left[j_{l_i}(k_ir)\,\Delta^{X_i}_{l_i}(k_i)\right], \tag{4.4}$$

where j_l is the spherical Bessel function arising from the Rayleigh expansion formula. Using this equation, we can compute the projected bispectrum from any given primordial bispectrum. Direct computation of this four-dimensional integral for every l combination, however, is practically impossible. Not only is the integral in 4D but also the oscillatory integrand requires a large number of sample points in each of k_i, making the full calculation for every l_i triple prohibitively expensive. All bispectrum estimators get around this problem by expanding B_Φ as a sum of *separable* terms. This will be explained in more detail later using the feature model template as an example.

4.3.2 Feature model

We follow the works of [34, 17, 18, 37] and assume the following template for the bispectrum of feature models;

$$B^{\text{feat}}_\Phi(k_1,k_2,k_3) = \frac{6A^2}{(k_1k_2k_3)^2}\sin(\omega K+\phi), \tag{4.5}$$

where $K = k_1+k_2+k_3$, A represents the primordial power spectrum amplitude, and ϕ is a phase. The oscillation 'frequency' ω is associated with the location and scale of feature in the inflationary potential. It is often written in terms of the oscillation scale k_c as $\omega = 2\pi/3k_c$. ω is measured in Mpc but we omit the unit for notational conveniences.

The feature model template has two free parameters that need to be fixed before we can constrain the model: the frequency ω and phase ϕ. The phase can be easily dealt with from observing that

$$B_{\Phi}^{\text{feat}}(k_1,k_2,k_3) = \cos\phi \, B_{\Phi}^{\sin}(k_1,k_2,k_3) + \sin\phi \, B_{\Phi}^{\cos}(k_1,k_2,k_3). \tag{4.6}$$

Here $B_{\Phi}^{\sin}$ and $B_{\Phi}^{\cos}$ correspond to feature models with $\phi = 0, \pi/2$ respectively. Non-zero phase simply corresponds to a linear combination of the sine and cosine templates. As we will see later these two shapes are in fact highly uncorrelated, and therefore they can be constrained independently from each other.

On the other hand, one still has a complete freedom of choice on the oscillation frequency. Such freedom dramatically expands size of the parameter space. In practice we constrain f_{NL} for each fixed value of oscillation frequency, which yields hundreds of estimates. Since there are so many estimates we are looking at, there is a good chance that we find notable signals by sheer luck. Accounting for this 'look elsewhere effect' has been resolved using methods in [17] and subsequently applied to the Planck analysis [37, 18]. The look-elsewhere-adjusted statistics used in the literature can be employed for the future CMB-S4 data analysis. This work, however, focuses on forecasting the 'raw' estimates and comparing them with those of Planck.

4.3.3 Separability

The bispectrum template of feature models (4.5) is an example of separable shape. It can be expressed as a sum of terms in the form $f(k_1)g(k_2)h(k_3)$ for some functions f, g and h, dramatically simplifying the computation of reduced bispectrum $b_{l_1 l_2 l_3}$. The three-dimensional integral over the k space in (4.4) splits into three individual one-dimensional integrals for separable shapes. Feature models for example has

$$\begin{aligned} b_{l_1 l_2 l_3}^{X_1 X_2 X_3,\text{feat}} &= 6A^2 \left(\frac{2}{\pi}\right)^3 \int_0^\infty r^2 dr \int_{\mathcal{V}_k} d^3\mathbf{k}\, e^{i\omega(k_1+k_2+k_3)} \prod_{i=1}^{3} \left[j_{l_i}(k_i r) \Delta_{l_i}^{X_i}(k_i) \right] \\ &= 6A^2 \left(\frac{2}{\pi}\right)^3 \int_0^\infty r^2 dr \prod_{i=1}^{3} \left[\int_0^\infty dk_i\, e^{i\omega k_i} j_{l_i}(k_i r) \Delta_{l_i}^{X_i}(k_i) \right]. \end{aligned} \tag{4.7}$$

Here the real and imaginary parts of b^{feat} correspond to the bispectra of cosine and sine feature models, respectively. Now define

$$s_l^X(r) := \frac{2A^{2/3}}{\pi}\int_0^\infty dk\,\sin(\omega k)\,j_l(kr)\Delta_l^X(k) \tag{4.8}$$

$$c_l^X(r) := \frac{2A^{2/3}}{\pi}\int_0^\infty dk\,\cos(\omega k)\,j_l(kr)\Delta_l^X(k). \tag{4.9}$$

These are analogous to $\alpha_l^X(r)$ and $\beta_l^X(r)$ in the usual Komatsu-Spergel-Wandelt (KSW) estimator for local non-Gaussianity. Then (4.7) reduces to

$$\begin{aligned} b_{l_1l_2l_3}^{X_1X_2X_3,\text{feat}} &= 6\int_0^\infty r^2dr\left(c_{l_1}^{X_1}c_{l_2}^{X_2}c_{l_3}^{X_3} - c_{l_1}^{X_1}s_{l_2}^{X_2}s_{l_3}^{X_3} - s_{l_1}^{X_1}c_{l_2}^{X_2}s_{l_3}^{X_3} - s_{l_1}^{X_1}s_{l_2}^{X_2}c_{l_3}^{X_3}\right) \\ &\quad + 6i\int_0^\infty r^2dr\left(s_{l_1}^{X_1}c_{l_2}^{X_2}c_{l_3}^{X_3} + c_{l_1}^{X_1}s_{l_2}^{X_2}c_{l_3}^{X_3} + c_{l_1}^{X_1}c_{l_2}^{X_2}s_{l_3}^{X_3} - s_{l_1}^{X_1}s_{l_2}^{X_2}s_{l_3}^{X_3}\right). \\ &= b_{l_1l_2l_3}^{X_1X_2X_3,\cos} + i\,b_{l_1l_2l_3}^{X_1X_2X_3,\sin} \end{aligned} \tag{4.10}$$

4.4 Efficient computation of the estimator with polarisation

4.4.1 Estimator

The optimal estimator for a given bispectrum in the weak non-Gaussian limit is [24, 23];

$$S_i = \frac{1}{6}\sum_{l_j,m_j}\sum_{X_j}\mathscr{G}^{l_1l_2l_3}_{m_1m_2m_3}b_{l_1l_2l_3}^{X_1X_2X_3,(i)}(C^{-1}_{l_1m_1,l_4m_4})^{X_1X_4}(C^{-1}_{l_2m_2,l_5m_5})^{X_2X_5}(C^{-1}_{l_3m_3,l_6m_6})^{X_3X_6} \left[a_{l_4m_4}^{X_4}a_{l_5m_5}^{X_5}a_{l_6m_6}^{X_6} - \left(C_{l_4m_4,l_5m_5}a_{l_6m_6}^{X_6} + 2\text{ cyclic}\right)\right]. \tag{4.11}$$

Computing this form involves an inversion of the full covariance matrix which is very computationally expensive. As a result we will follow the diagonal covariance approximation in [44] for the inverse covariances, so $C^{-1}_{l_1l_4m_1m_4} \approx 1/C_{l_1}\,\delta^D_{l_1l_4}\,\delta^D_{m_1-m_4}$; and approximate the covariance in the linear term by an ensemble average of realistic simulations $C^{X_1X_2}_{l_4l_5m_4m_5} \approx \left\langle a_{l_1m_1}^{X_1}a_{l_2m_2}^{X_2}\right\rangle$. With these the estimator takes the form;

$$\hat{f}_i = \sum_j (F^{-1})_{ij}S_j, \tag{4.12}$$

where

$$S_i = \frac{1}{6} \sum_{l_j, m_j} \sum_{X_j, X'_j} \mathscr{G}^{l_1 l_2 l_3}_{m_1 m_2 m_3} b^{X_1 X_2 X_3, (i)}_{l_1 l_2 l_3} (C^{-1}_{l_1})^{X_1 X'_1} (C^{-1}_{l_2})^{X_2 X'_2} (C^{-1}_{l_3})^{X_3 X'_3}$$
$$\left[a^{X'_1}_{l_1 m_1} a^{X'_2}_{l_2 m_2} a^{X'_3}_{l_3 m_3} - \left(\left\langle a^{X'_1}_{l_1 m_1} a^{X'_2}_{l_2 m_2} \right\rangle a^{X'_3}_{l_3 m_3} + 2 \text{ cyclic} \right) \right]. \tag{4.13}$$

Here summations are over l_j, m_j, X_j and X'_j for each $j = 1,2,3$. The spherical multipole moments a^X_{lm}'s are computed from either observations or simulations. The covariance matrix C_l is a 2×2 matrix consisting of values C^{TT}_l, C^{TE}_l, C^{ET}_l and C^{EE}_l. [1] The linear terms (the second in square brackets) are required to account for anisotropies induced by masking and anisotropic noise. The bracket $\langle \cdot \rangle$ denotes averaging over Monte Carlo simulations of Gaussian realisations.

The Fisher matrix of the estimator is given by

$$F_{ij} = \frac{f_{\text{sky}}}{6} \sum_{\text{all } X, X'} \sum_{\text{all } l} h^2_{l_1 l_2 l_3} \, b^{X_1 X_2 X_3, (i)}_{l_1 l_2 l_3} (C^{-1}_{l_1})^{X_1 X'_1} (C^{-1}_{l_2})^{X_2 X'_2} (C^{-1}_{l_3})^{X_3 X'_3} \, b^{X'_1 X'_2 X'_3, (j)}_{l_1 l_2 l_3}. \tag{4.14}$$

Here f_{sky} denotes the fraction of the sky covered by the experiment, and $h^2_{l_1 l_2 l_3} := \sum_{m_j} \left(\mathscr{G}^{l_1 l_2 l_3}_{m_1 m_2 m_3} \right)^2$. Since the estimator $\hat{f}_i$ in (4.12) is nearly optimal, its 68% confidence (1σ) interval can be computed from the Fisher matrix as $\sigma_i := \Delta f^{(i)}_{NL} = (F^{-1})_{ii}$.

Note that most CMB-S4 experiments are ground-based, so they can probe smaller fraction of the sky compared to Planck. Having a smaller fraction of the sky leads to increased uncertainties for the estimator. Current estimate is that the new experiments will cover 40% of the sky, significantly less than the 74% of Planck. The error bars will thus increase by a factor of 1.38 from the decrease in f_{sky} alone. This will can be reduced by combining Planck data for unobserved pixels in these experiments

4.4.2 Orthonormalising the covariance matrix

In [16] it was noted that orthogonalising the multipoles of temperature and polarisation maps dramatically reduces the number of terms in computation of the Modal estimators. This technique can also be applied to KSW estimators, or indeed any optimal bispectrum estimator, which is yet to be done to the authors' knowledge.

In both (4.13) and (4.14) there are summations over indices X and X' to account for correlations between the CMB temperature and E-mode polarisation. This can be simplified

[1] Note that this is equivalent to having a $2l \times 2l$ matrix with diagonal $l \times l$ block matrices C^{TT}, C^{TE}, C^{ET} and C^{EE} as in other literatures including [16].

by essentially making a change of basis in X space for each l so that every C_l becomes orthonormal. Perform a Cholesky decomposition on C_l and invert the matrix. Then $C_l^{-1} = L_l^T L_l$, where L_l is a lower triangular matrix given by

$$L_l = \begin{pmatrix} \frac{1}{\sqrt{C_l^{TT}}} & 0 \\ \frac{-C_l^{TE}}{\sqrt{C_l^{TT}}\sqrt{C_l^{TT}C_l^{EE}-C_l^{TE2}}} & \frac{C_l^{TT}}{\sqrt{C_l^{TT}}\sqrt{C_l^{TT}C_l^{EE}-C_l^{TE2}}} \end{pmatrix} \tag{4.15}$$

Now let

$$\tilde{\Delta}_l^X(k) = \sum_{X'} L_l^{XX'} \Delta_l^{X'}(k), \quad \text{and} \quad \tilde{a}_{lm}^X = \sum_{X'} L_l^{XX'} a_{lm}^{X'}. \tag{4.16}$$

When a_{lm}'s are generated from simulations, the second transformation is not required as long as the new transfer function $\tilde{\Delta}_l^X$ is used in the process.

Defining $\tilde{b}_{l_1l_2l_3}$ to be the corresponding reduced bispectrum, (4.13) and (4.14) simplify to

$$S_i = \frac{1}{6} \sum_{l_j,m_j} \sum_{X_j} \mathscr{G}^{l_1l_2l_3}_{m_1m_2m_3} \tilde{b}^{X_1X_2X_3,(i)}_{l_1l_2l_3} \left[\tilde{a}^{X_1}_{l_1m_1} \tilde{a}^{X_2}_{l_2m_2} \tilde{a}^{X_3}_{l_3m_3} - \left(\left\langle \tilde{a}^{X_1}_{l_1m_1} \tilde{a}^{X_2}_{l_2m_2} \right\rangle \tilde{a}^{X_3}_{l_3m_3} + 2 \text{ cyclic} \right) \right], \tag{4.17}$$

$$F_{ij} = \frac{f_{\text{sky}}}{6} \sum_{\text{all } X} \sum_{\text{all } l} h^2_{l_1l_2l_3} \, \tilde{b}^{X_1X_2X_3,(i)}_{l_1l_2l_3} \, \tilde{b}^{X_1X_2X_3,(j)}_{l_1l_2l_3}. \tag{4.18}$$

Using this method not only is mathematically concise, but also halves the number of terms involved in the summation. Linear transformations (4.16) only need to be done once in the beginning of the program and cost little compared to the main computation. We also found it easier to optimise the code using instruction level vectorisations after this simplification.

The only downside of this method is that we no longer can get breakdowns of signal from each of TTT, TTE, TEE and EEE bispectrum since our new modes are linear combinations of T and E modes. However in most cases we are interested in either T-only or $T+E$ results, and this method works perfectly well in these cases.

4.4.3 Estimator for feature models

We apply general estimator to the case of feature models. The method is similar to the one seen in [34] except that now the polarisation is included and the covariance matrices are trivial thanks to the orthonormalisation process outlined above.

Consider the bispectrum shape of

$$B_\Phi(k_1,k_2,k_3) = f_{NL}^{\sin} B^{\sin}(k_1,k_2,k_3) + f_{NL}^{\cos} B^{\cos}(k_1,k_2,k_3), \tag{4.19}$$

for a fixed value of oscillation frequency ω. Here $B^{\sin}$ and $B^{\cos}$ correspond to reduced bispectra $b^{\sin}$ and $b^{\cos}$ defined in (4.10). The Fisher matrix F is 2×2 but its off-diagonal entries are 2-3 orders of magnitude smaller than diagonal ones in most cases as will be presented in the next section. Thus, the two shapes are assumed to be uncorrelated and constrained individually. Here we present detailed computations for $f_{NL}^{\sin}$ only but the cosine one can be computed similarly.

From (4.10) and the definition of Gaunt integral $\mathscr{G}^{l_1 l_2 l_3}_{m_1 m_2 m_3} = \int d\hat{\mathbf{n}}\, Y_{l_1 m_1}(\hat{\mathbf{n}}) Y_{l_2 m_2}(\hat{\mathbf{n}}) Y_{l_3 m_3}(\hat{\mathbf{n}})$ it follows that

$$S^{\mathrm{cub}} = \int_0^\infty r^2 dr \int d^2\hat{\mathbf{n}} \left[-{M_s}^3 + 3 M_s {M_c}^2 \right] \quad \text{and} \tag{4.20}$$

$$S^{\mathrm{lin}} = -3 \int_0^\infty r^2 dr \int d^2\hat{\mathbf{n}} \left[-M_s \left\langle {M_s}^2 \right\rangle + M_s \left\langle {M_c}^2 \right\rangle + 2 M_c \left\langle M_s M_c \right\rangle \right], \tag{4.21}$$

where

$$\begin{aligned} M_s(r, \hat{\mathbf{n}}) &= \sum_X \sum_{lm} \tilde{s}^X_l(r)\, \tilde{a}^X_{lm} Y_{lm}(\hat{\mathbf{n}}), \\ M_c(r, \hat{\mathbf{n}}) &= \sum_X \sum_{lm} \tilde{c}^X_l(r)\, \tilde{a}^X_{lm} Y_{lm}(\hat{\mathbf{n}}). \end{aligned} \tag{4.22}$$

The sum of S^{cub} and S^{lin} gives the final value of S for sine feature model.

For efficient Fisher matrix calculation we follow [40] and deploy the identity

$$h^2_{l_1 l_2 l_3} = \frac{(2l_1+1)(2l_2+1)(2l_3+1)}{8\pi} \int_{-1}^1 d\mu\, P_{l_1}(\mu) P_{l_2}(\mu) P_{l_3}(\mu), \tag{4.23}$$

where $P_l(\mu)$ represents the Legendre polynomial. Then,

$$F = \frac{3}{4\pi} \int r^2 dr \int r'^2 dr' \int d\mu \left[P^3_{ss} + 3 P_{ss} P^2_{cc} - 3 P^2_{cs} P_{ss} - 3 P^2_{sc} P_{ss} + 6 P_{cs} P_{sc} P_{cc} \right]. \tag{4.24}$$

where we have defined

$$\begin{aligned} P_{ss}(r, r', \mu) &:= \sum_X \sum_l (2l+1)\, \tilde{s}^X_l(r) \tilde{s}^X_l(r') P_l(\mu) \\ P_{sc}(r, r', \mu) &:= \sum_X \sum_l (2l+1)\, \tilde{s}^X_l(r) \tilde{c}^X_l(r') P_l(\mu). \end{aligned} \tag{4.25}$$

and similarly P_{cs} and P_{cc}.

Calculation of (4.22) and (4.25) are two of the most computationally expensive steps. If we did not orthonormalise the covariance matrix, there would be extra summations over X' and some 2×2 matrix computations involving $(C_l^{-1})^{XX'}$ in these steps.

4.4.4 Probing beam and instrumental noise

In an ideal experiment where measurements are made on each point of the sky perfectly, the covariance matrices $C_l^{XX'}$ in (4.13) and (4.14) consists purely of the signal. In reality, however, the probing beam has finite width and the sensors are noisy. These effects can be incorporated by modifying the covariance matrices and bispectra as follows.

$$C_l^{X_1X_2} \to W_l^{X_1}W_l^{X_2}C_l^{X_1X_2} + N_l^{X_1X_2}, \qquad b_{l_1l_2l_3}^{X_1X_2X_3} \to W_{l_1}^{X_1}W_{l_2}^{X_2}W_{l_3}^{X_3}b_{l_1l_2l_3}^{X_1X_2X_3}, \tag{4.26}$$

where W_l^X and $N_l^{X_1X_2}$ represent the beam window function and the noise covariance matrix, respectively. When substituted into KSW estimator, these changes are equivalent to modifying

$$\begin{aligned} C_l^{X_1X_2} &\to C_l^{X_1X_2} + \left(W_l^{X_1}W_l^{X_2}\right)^{-1} N_l^{X_1X_2} \\ &= (C_l^{\rm sig})^{X_1X_2} + (C_l^{\rm noise})^{X_1X_2}, \end{aligned} \tag{4.27}$$

while keeping the bispectra same. Here we have defined the effective (beam-corrected) noise covariance matrix $C_l^{\rm noise}$. Modes for which $C_l^{\rm noise}$ is much larger than $C_l^{\rm sig}$ contribute little to the f_{NL} estimator.

For forecasting purposes we assume Gaussian beam and white uncorrelated noise until more detailed experiment specifications become available. Under these assumptions, the effective noise covariances reduce to [35]

$$C_l^{{\rm noise},TT} = \exp\left(l(l+1)\sigma_{\rm beam}^2\right)N_{\rm white}, \quad C_l^{{\rm noise},EE} = 2\,C_l^{{\rm noise},TT}, \quad C_l^{{\rm noise},TE} = 0. \tag{4.28}$$

The factor of two for EE mode is comes from measuring two Stokes parameters Q and U. The Gaussian beam profile is usually specified by its FWHM (full width at half maximum) in *arcmin*, which is then converted to standard deviations in radians for $\sigma_{\rm beam}$. The noise level often comes in the units of $\mu K \cdot arcmin$. This is then divided by $T_{\rm CMB} = 2.725K$, converted to radians and squared to get $N_{\rm white}$.

For the Planck experiment, using 5 arcmin FWHM beam and the 47 μK·arcmin noise level gives good approximations to the post-component-separation noise covariances. For CMB-S4 experiments the details are not confirmed, but the beam FWHM is expected to lie between 1-5 arcmin, while the noise level will range from 1 to 9 μK·arcmin. [15]

In real measurements there exist extra contaminations in large angular scales due to 1/f noises and the component separation process. Though most of our analysis assumes simpler form of noise covariances elaborated above, for the Simons Observatory forecasts we follow [43] and model 1/f noise as $N_l = N_{red}(l/l_{knee})^{\alpha_{knee}} + N_{white}$. The noise curves from each channel were then put together using the inverse variance method. This is a good approximation for the E mode polarisation but not for temperature, since extra degradations occur during the component separation process. Still, because dominant contributions to the feature model signal comes from polarisation data, this would be a reasonable approximation for our forecast. For Planck the full post-component-separation noise curves are available and hence used for computations.

4.4.5 Implementation and validation

We implemented the pipeline outlined above using the C programming language and parallelised using hybrid MPI + openMP. The code was then run in the COSMOS super-computing system.

The transfer functions are generated from the CAMB code [25]. Bessel function values were pre-computed using recursion relations and stored in a file, while the Legendre function values were computed on the fly using the GNU scientific library. The angular power spectrum data was generated from ΛCDM parameters estimated in the Planck 2015 results.

Numerical integration for variables k, r and r' were done using simple trapezoidal methods, as they can be easily vectorised for optimisation. On the other hand, integration of μ required more care because the Legendre polynomials are highly oscillatory. We adopted the Gauss-Legendre quadrature rule with $1.5\, l_{\max} + 1$ points which can integrate polynomials up to order $3\, l_{\max}$ exactly. The weights and nodes were computed in the beginning using the QUADPTS code [20].

Various checks have been done to ensure that the code runs correctly. First we used the code to reproduce the Planck results, which agreed within 3% error. The code was then used to compute bispectrum for the constant model, corresponding to the case where $\omega = \phi = 0$. There exists an approximate analytic form in this case [19] which we were able to reproduce accurately. We also performed convergence tests on r and r' integration by doubling the number of points for each of them. The grid was chosen to be very dense around recombination and quite dense near reionisation. We confirmed that changes in the integral are less than 0.5% for each value of ω.

4.5 CMB-S4 forecast results

4.5.1 Phase dependence

We now present the CMB-S4 forecast on the error bars of primordial non-Gaussianity parameter for feature models. For notational convenience we denote the error bars for sine and cosine feature models by $\sigma_{\sin}$ and $\sigma_{\cos}$. Superscripts T and $T+E$ are also put to distinguish temperature-only analysis from the full polarisation ones.

First of all, we check that the sine and cosine bispectrum templates defined in 4.7 are indeed uncorrelated and can be constrained separately. Equivalently, we can see if the Fisher matrix of feature models is robust to changes in the phase for different ω values of interest. Feature model bispectra with a specific phase ϕ can be represented as a sum of sine and cosine ones as in (4.6). Hence, its Fisher matrix is given by

$$F(\omega,\phi) = \cos^2\phi\ F_{ss}(\omega) + \sin^2\phi\ F_{cc}(\omega) + 2\cos\phi\sin\phi\ F_{sc}(\omega), \tag{4.29}$$

where F_{ss} is the element F_{ij} of the Fisher matrix in (4.14) with reduced bispectra $b^{(i)} = b^{(j)} = b^{\sin}$, and so on. Correlation between sine and cosine templates can be expressed as $F_{sc}/(F_{ss}F_{cc})^{1/2}$, and this value can be learned from analysing the ϕ dependence of $F(\omega,\phi)$.

Figure 4.1 shows forecast error bars for the full phase range $[0,\pi]$ in the most sensitive experiment specification of 1' beam and $1\mu K\cdot arcmin$ noise. The forecast σ varies within 1% level for every $\omega \geq 20$. In terms of the Fisher matrix, the cross term F_{sc} was 2-3 orders of magnitude smaller than F_{ss} and F_{cc} for all cases. In other words, correlation between the sine and cosine templates was smaller than 1%. This justifies our previous choice of constraining $f_{NL}^{\sin}$ and $f_{NL}^{\cos}$ separately. We now focus our attention to $\sigma_{\sin}$ in future discussions.

For smaller values of ω, the phase affects the error bar primarily through modulating the amplitude if the acoustic oscillations of the CMB itself. The radiation transfer functions are non-zero for k values in $0-0.8$ Mpc^{-1}. The argument ωk covers less than two full periods in this k range if $\omega \leq 10$ Mpc, and phase has direct influence on the amplitude of the acoustic peaks. In the extreme case of $\omega = 0$, bispectrum vanishes completely for the sin feature model. Variations in the overall bispectrum amplitude therefore result in varying Fisher information for low frequencies.

4.5.2 l_{max} dependence

Figure 4.2 shows the graph of forecast error bar $\sigma_{\sin}^{T+E}$ as we increase l_{max}. The forecasts were done within angular scale range $2 \leq l \leq l_{max}$, the oscillation frequency ω set to 100, and assuming 1' beam and $1\mu K$·arcmin noise. The Planck noise curves were approximated

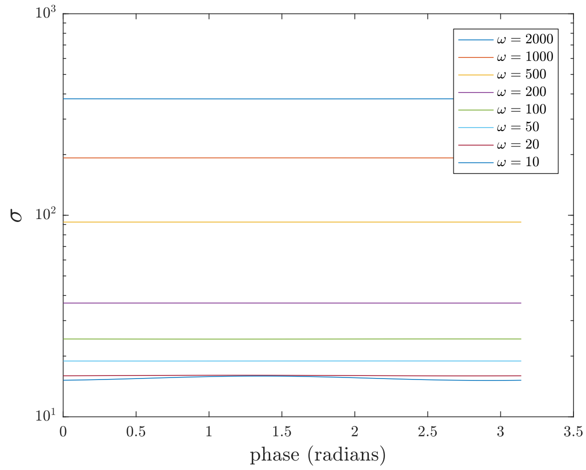

Fig. 4.1 Forecast error bars σ^{T+E} versus the phase ϕ. Apart from the smallest frequency $\omega = 10$, the error bar remains almost constant. This implies that the sine ($\phi = 0$) and cosine ($\phi = \pi/2$) feature models can be constrained independently.

by ones for 5' beam and 47 μK·arcmin noise for this plot only, since we extend l_{max} to 4000 here.

The Planck error bar essentially stalls out when l_{max} reaches 2000. The forecast error bar, on the other hand, keeps decreasing until $l_{max} = 4000$ thanks to the improved sensitivity in measuring small scales (and large l's). Despite the information loss due to smaller sky coverage f_{sky}, the forecast error bar reduces to about 42% of Planck by $l_{max} = 4000$. This corresponds to a factor of 2.4 times improvement to measurement precision on $f_{\rm NL}$.

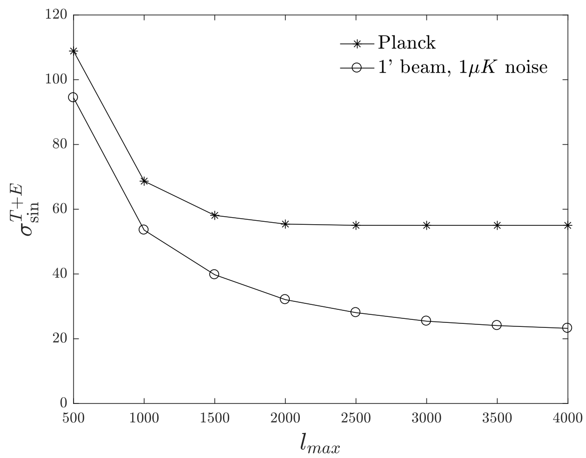

Fig. 4.2 Forecast error bars $\sigma^{T+E}_{\rm sin}$ when multipoles $2 \le l \le l_{max}$ are included, in comparison with Planck. The oscillation frequency ω is set to 100 Mpc in all cases. Planck did not have access to the information from modes $l \ge 2000$, but the CMB-S4 experiments are expected to be able to explore modes up to $l = 4000$.

4.5.3 Beam and noise dependence

Now we explore the effects of different beam width and noise levels on the forecast error bars. Figure 4.3 shows forecast $\sigma_{\text{sin}}^{T+E}$ for ranges of beam and noise levels. Their oscillation frequencies are also varied, but only two representatives $\omega = 20$ and 2000 are chosen here. Forecasts for the other values of ω also display similar dependences on beam width and noise levels.

First of all, note that all estimated error bars in the plot are smaller than Planck, for which $\sigma_{\text{sin}}^{T+E} = 34$ when $\omega = 20$ and $\sigma_{\text{sin}}^{T+E} = 610$ when $\omega = 2000$. In fact even the least sensitive CMB-S4 specification of 5' beam and $9\mu K$·arcmin noise is expected to put better bounds on feature models.

Wider beams and noisier detectors provide less signal and thus larger error bars, as expected. In this range of beam width and noise levels, noise has a bigger effect on the forecast; experiments with 1' beam and $5\mu K$·arcmin noise yields larger error bars than the ones with 5' beam with $1\mu K$·arcmin noise. Between the most sensitive specification of 1' beam and $1\mu K$·arcmin and the least sensitive one with 5' beam and $9\mu K$·arcmin, σ_{sin} differs by a factor of 1.6.

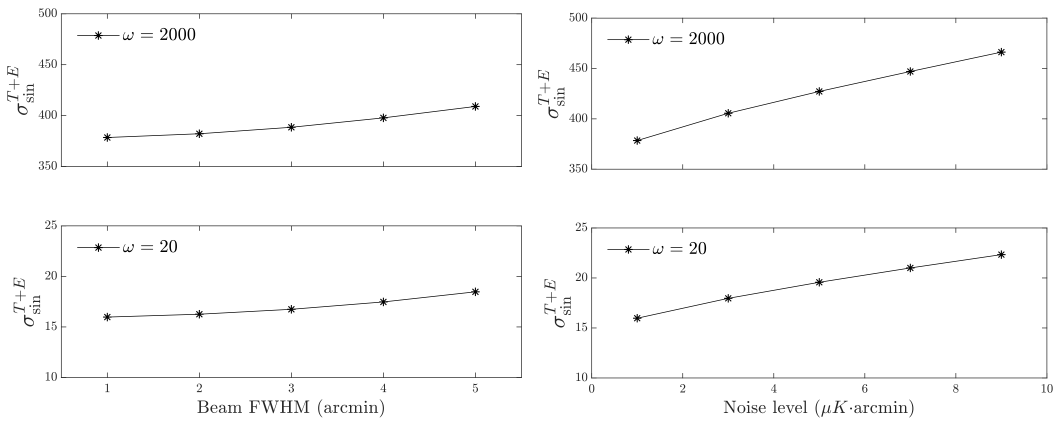

Fig. 4.3 Beam (left) and noise (right) dependences of the forecast error $\sigma_{\text{sin}}^{T+E}$ for $\omega = 2000$ (top) and $\omega = 20$ (bottom). The noise level was set as $1\mu K$·arcmin for the first plot, while the second plot had fixed beam FWHM of 1'. We obtain less information from using wider beam and noisier sensors, as expected.

4.5.4 Oscillation frequency dependence

We present the main results of the forecast. Figure 4.4 summarises the σ_{sin} forecasts for several different CMB-S4 preliminary specifications, including the Simons Observatory (SO) baseline and goal. Note that the 1/f noise effects are incorporated in SO forecasts but not in other ones. We also provide 1σ errors for joint estimators, for which Planck signals from the fraction of the sky not covered by CMB-S4 are combined via $\sigma_{\text{joint}}^{-2} = \sigma_{\text{CMB-S4}}^{-2} + \sigma_{\text{Planck}}^{-2}$. This method is not statistically optimal but sufficient to give an idea of the joint estimation power.

The most sensitive setup with 1' beam and $1\mu K$·arcmin noise would yield error bars that are 47-62% of Planck. These correspond to a factor of 1.6-2.1 improvements. Here relatively smaller improvements are made for high oscillation frequencies. They correspond to smaller momentum scales $k_* = 2\pi/3\omega$, or larger angular scales, which benefit less from the increased sensitivity of CMB-S4 experiments. When the results are combined with Planck, the error bar reduces to 45-57% of Planck, or a factor of 1.7-2.2 improvement.

Forecast error bars from the SO baseline specification and the more ambitious one do not differ very much. Quoting in terms of the baseline values, σ_{sin} lies about 68-86% of that of Planck or equivalently, 1.2-1.5 times smaller than Planck. Numbers change to 62-74% when combined with Planck, so that the overall improvement ratio is about 1.3-1.6.

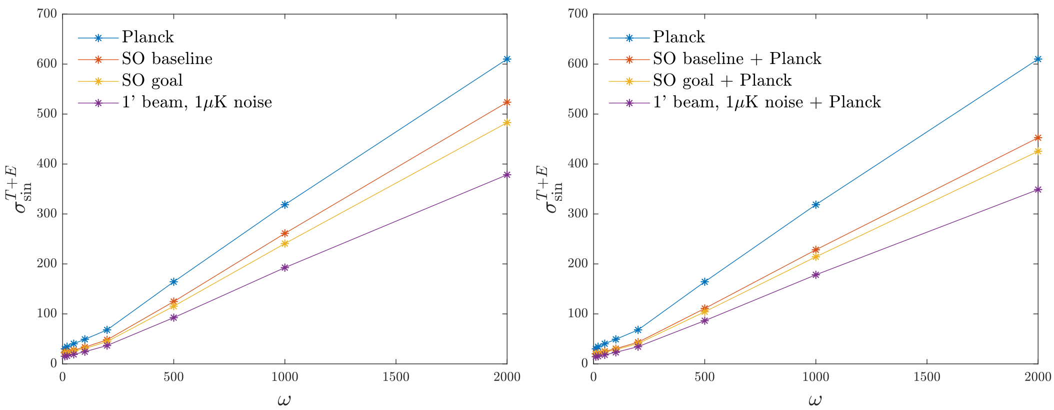

Fig. 4.4 Frequency dependence of the forecast error in comparison to Planck (left). All CMB-S4 specifications would improve constraints on feature models. The most sensitive setup with 1' beam and $1\mu K$·arcmin noise is expected to yield error bars that are 1.6-2.1 times smaller than Planck. When the Planck results are combined with CMB-S4, we get even stronger constraints (right).

Figure 4.5 shows the results when only the CMB temperature data are used in the forecast. CMB-S4 would in fact be worse than Planck in terms of constraining f_{NL}^{feat} for this case. The loss in information due to less sky coverage overwhelms the increased sensitivity. We see again that the real strength of CMB-S4 experiments lies in measuring CMB polarisation.

Then how much information do we actually gain from adding E-mode polarisation? Figure 4.6 shows the ratio of σ_{sin}s between the temperature-only and polarisation-included analyses. The forecast error bars reduces up to 4.6 times smaller when polarisation information is added, which is much larger than the corresponding Planck value of 2.2. The ratio

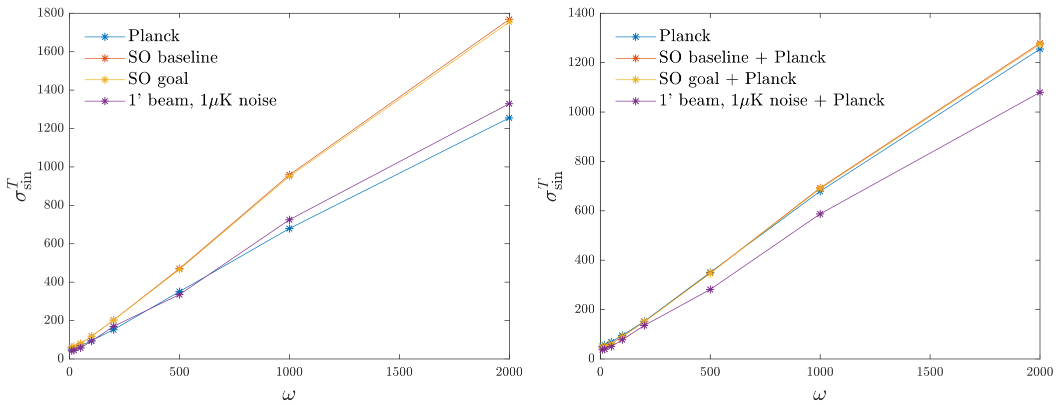

Fig. 4.5 Frequency dependence of the forecast error from temperature data only, in comparison to Planck (left). The CMB-S4 experiments would perform worse than Planck when only the temperature map is concerned. After the addition of Planck data the error bars improve only marginally (right). This shows that polarisation data is crucial for constraining feature models.

decreases overall when the joint statistics with Planck are considered. An intriguing feature of this plot is that the ratio is maximised around $\omega = 200$ before it starts dropping again.

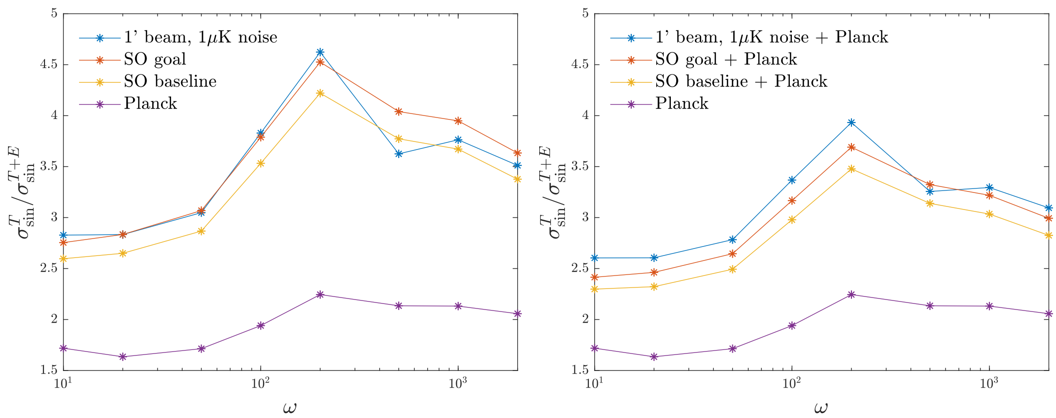

Fig. 4.6 Improvements on the forecast error when including E-mode polarisation data. Constraints from the CMB-S4 experiments would improve significantly from addition of the polarisation data. The improvement is maximised around $\omega \approx 200$ Mpc.

In order to gain insight on this behaviour, we performed some simplified computations using the power spectrum. We imposed oscillations on the primordial power spectrum as $P'(k) = P(k)(1+\sin(2\omega k+\phi))$, which is just like our feature model bispectrum template but with $\omega(k_1+k_2+k_3)$ replaced by $\omega(k+k)$. $P'(k)$ is then projected to the late-time harmonic space using the transfer functions;

$$C_l^{\prime X_1 X_2} = \frac{2}{\pi} \int k^2 dk P'(k) \Delta_l^{X_1}(k) \Delta_l^{X_2}(k). \tag{4.30}$$

We observed that the fractional variation $(C_l' - C_l)/C_l$ displays some oscillations in l, and the largest contribution comes from a term $\propto \sin(2\omega l/\Delta\tau)$ where $\Delta\tau$ represents the conformal distance to last scattering surface. This fact can be explained by approximating the transfer function as $\Delta_l(k) \approx (1/3) j_l(k\Delta\tau)$ and noting that the spherical Bessel function has a sharp peak at l for large l's. The integral in (4.30) therefore picks up a term proportional to $\sin(2\omega l/\Delta\tau)$.

The amplitude of these 'maximal' oscillations in $(C_l' - C_l)/C_l$ were then computed using discrete Fourier transform for different values of oscillation scale ω and two different phases $\phi = 0, \pi/2$ (i.e. sine and cosine). The results are shown in Figure 4.7. Some extra wiggles to the graph come from the phase of oscillations imposed; we indeed see that graphs of sine and cosine oscillate between each other. Some peak features near $\omega \approx 70$ and 140 arise from resonances with Baryonic Acoustic Oscillations.

We can think of the computed amplitude as a measure of information C_l's contain about primordial oscillations. First of all, note that the amplitude in all four plots generally decreases as ω grows. Previously in Figure 4.4 we saw that the amount of information obtained from the CMB is smaller for larger ω's, consistent with what can be said from the amplitude analysis. Moreover, the amplitudes for the EE mode are generally larger than the TT mode ones, and their difference is the largest in the ω range of 70 to 300. This could serve as a heuristic explanation for the improvement in forecast error bars from including polarisation data being maximised around $\omega = 200$, as depicted in Figure 4.6.

4.5.5 Comparison to scale invariant models

Our pipeline for forecasting f_{NL}^{feat} also yields forecasts for f_{NL} of the constant model. Constant models are scale invariant and have trivial shape, so that $B(k_1, k_2, k_3) \propto (k_1 k_2 k_3)^{-2}$. Forecasts on f_{NL}^{const} follow from our pipeline by simply setting the oscillation frequency $\omega = 0$ and phase $\phi = \pi/2$. Table 4.1 summarises the forecast results for several different CMB-S4 specifications mentioned before, using both T and E data and in combination with Planck data from the regions of the sky not covered by CMB-S4. For the 1' beam and $1\mu K$·arcmin noise setup, the error bar is expected to be reduced by a factor of 2.3 compared to Planck.

The latest Planck constraints on f_{NL} of some popular bispectrum templates are given by $f_{\text{NL}}^{\text{local}} = 2.5 \pm 5.7$, $f_{\text{NL}}^{\text{equil}} = -16 \pm 70$, and $f_{\text{NL}}^{\text{ortho}} = -34 \pm 33$ [37]. CMB-S4 experiments are expected to yield better estimates on these as well. Table 4.2 summarises the forecast

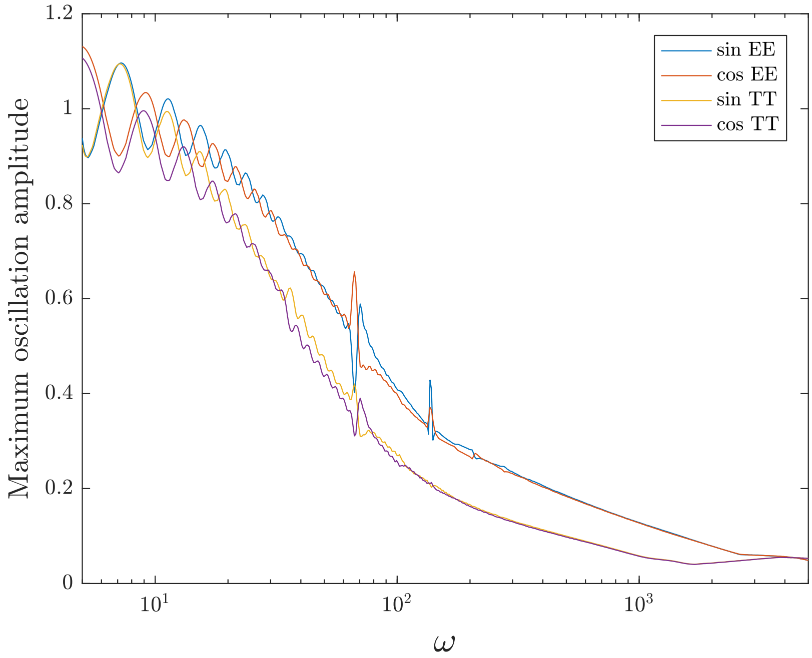

Fig. 4.7 The maximum amplitude of oscillations detected in fractional variations of the projected power spectrum C_l^{TT} and C_l^{EE}, when extra oscillations $\sin(2\omega k)$ and $\cos(2\omega k)$ were imposed on the primordial power spectrum. Heuristically this shows that E-mode polarisation is more sensitive to the primordial oscillations, especially in the ω range of 70 to 300. Some peaks near $\omega = 70$ and 140 arise from resonances with Baryonic Acoustic Oscillations.

	Planck	SO baseline + Planck	SO goal + Planck	1' beam, $1\mu K$ noise + Planck
$\sigma(f_{\mathrm{NL}}^{\mathrm{const}})$	23.4	14.9	14.0	10.4

Table 4.1 Forecasts on the estimation errors of f_{NL} for the constant model

improvement ratio given in [15] together with the constant and feature model ratios computed in this work.

To the authors' surprise, the estimation error for feature models does not improve as much as other templates. Feature models benefit much more from polarisation data than other scale independent shapes; for example, $\sigma^T/\sigma^{T+E} = 4.6$ for the feature models with $\omega = 200$ in CMB-S4 while the value equals 2.8 for the constant models. Because CMB-S4 would have significantly enhanced polarisation measurement sensitivity, we originally expected the feature models to be constrained significantly better than Planck.

In order to investigate this lack of improvement, we performed a breakdown analysis on the improvements gained from CMB-S4 temperature and polarisation; we computed $\sigma(f_{\mathrm{NL}})$ for the constant and feature models using each of the four combinations of Planck / CMB-S4

	Local	Equilateral	Orthogonal	Constant	Feature ($\omega = 200$)
$\sigma^{\text{Planck}}/\sigma^{\text{CMB-S4}}$	2.5	2.1	2.4	2.3	2.0

Table 4.2 Expected improvement ratios of the f_{NL} estimation errors for the CMB-S4 1'beam, $1\mu K$arcmin setup, for various bispectrum templates. The local, equilateral and orthogonal results are quoted from [15].

noise curves for temperature / polarisation (e.g. Planck T + CMB-S4 E). The results are summarised in Table 4.3.

$\sigma(f_{\text{NL}}^{\text{const}})$ improvement		E		$\sigma(f_{\text{NL}}^{\text{feat}})$ improvement		E	
		Planck	CMB-S4			Planck	CMB-S4
T	Planck	1.0	1.6	T	Planck	1.0	1.7
	CMB-S4	1.1	2.2		CMB-S4	0.9	1.9

Table 4.3 Expected improvements on the estimation errors of f_{NL} for each combination of Planck / CMB-S4 temperature (T) and polarisation (E) data. Here the CMB-S4 assumes 1' beam and $1\mu K$arcmin noise. For feature model the oscillation frequency $\omega = 200$ and phase $\phi = 0$. The sky fraction $f_{sky} = 0.4$ for all cases except for Planck T + Planck E, for which $f_{sky} = 0.76$.

We see that the constraints on feature models improve by a factor of 1.7 when swapping Planck polarisation noises with the CMB-S4 ones. This factor is indeed larger than the one for constant model, which equals 1.6. The difference is however not significant. It seems that the amount of feature signals in polarisation data left unexplored by Planck is not tremendously large compared to the constant model. The feature model improves less than the constant model when the temperature measurements are enhanced. In fact, for feature models the signal loss from smaller sky fraction f_{sky} eclipses the signal gain from more sensitive temperature measurements. This lack of improvements from temperature causes the full CMB-S4 constraints on the feature model not to improve as much as the constant model overall.

4.6 Conclusion

Upcoming CMB Stage-4 experiments will provide an opportunity to measure CMB temperature and polarisation with greater precision. The estimation of primordial non-Gaussianity parameters would greatly benefit from the improvement in measurement sensitivity. In this research we made forecasts on f_{NL} for the feature models, which have not been done so far despite the growing interests on inflation models with primordial oscillations. For efficient forecasts we simplified the bispectrum estimator for f_{NL} by orthonormalising the covariance matrix, further optimising the computation. When the most sensitive CMB Stage-4 experiment specification of 1' beam and $1\mu K$arcmin noise is concerned, we expect a factor of 1.7-2.2 times more stringent constraints compared to Planck. Under realistic Simons Observatory conditions the improvement would be about 1.3-1.6 to Planck.

Although this is not a massive boost in the estimation power, we can hope to verify current 4σ-level signals found in the 2015 Planck analysis. It is also worth noting that the CMB-S4 experiments would allow us to explore modes $l > 2000$, especially since localised oscillations in this range are currently unconstrained. Moreover, though we have only considered linearly spaced oscillations in this work, we expect even better improvements on the models inducing logarithmically spaced oscillations. Higher l modes would promote the constraining power, since the oscillation slows down in small scales for this type of models and therefore gets less suppressed by the transfer functions. Lastly, cross-validation using these new statistically independent modes would be useful.

We also extensively studied how the forecasts depend on various parameters. Frequency dependences of the ratio between T and T+E forecasts were particularly illuminating - the improvement from adding polarisation information is maximised around $\omega = 200$. Some simplified calculations were presented to heuristically address this fact. Even though the estimation power on feature models massively benefit from the polarisation data, overall expected improvements compared to Planck are quite underwhelming. Breakdown analysis on temperature and polarisation contribution revealed that the feature models would indeed improve more than other scale-independent models if only the polarisation measurement sensitivity is enhanced to the CMB-S4 standards. However, boosts in the temperature measurements affect scale-independent models more so that they gain more information overall.

Chapter 5

High-Resolution CMB Bispectrum Estimator

Intro: current methods for CMB bispectrum estimation.

5.1 Formalism

5.1.1 Modal estimator?

Review of the Modal estimator formalism

5.1.2 CMB-BEst formalism

Recall that the CMB bispectrum estimator for a given template can be written as

$$\hat{f}_{NL} = \frac{1}{N} \sum_{l_j, m_j} \frac{\mathcal{G}^{l_1 l_2 l_3}_{m_1 m_2 m_3} b_{l_1 l_2 l_3}}{C_{l_1} C_{l_2} C_{l_3}} \left[a_{l_1 m_1} a_{l_2 m_2} a_{l_3 m_3} - \left(\left\langle a^G_{l_1 m_1} a^G_{l_2 m_2} \right\rangle a_{l_3 m_3} + 2 \text{ cyc.} \right) \right]. \tag{5.1}$$

Here we omit superscripts X for temperature and polarisation for notational convenience. Even though the formalism in this section will be presented for CMB temperature data only, the method is general and can easily be extended to include polarisation. For estimation of the full covariance matrix $C_{lm,l'm'}$ needed for the linear term, we use ensemble average from Gaussian simulations, as denoted by superscripts G and the bracket $\langle \cdot \rangle$.

The normalisation factor is given by

$$N = \sum_{l_j} \frac{h^2_{l_1 l_2 l_3} b^2_{l_1 l_2 l_3}}{C_{l_1} C_{l_2} C_{l_3}}. \tag{5.2}$$

The core part of our estimation routine is the separable mode expansion of shape function;

$$S(k_1,k_2,k_3) := (k_1 k_2 k_3)^2 B(k_1,k_2,k_3) = \sum_{p_j} \alpha_{p_1 p_2 p_3} q_{p_1}(k_1) q_{p_2}(k_2) q_{p_3}(k_3). \tag{5.3}$$

Choices for the basis functions $q_p(k)$ are detailed in the next section. Due to the separability, the reduced bispectrum reduces to a compact form of

$$b_{l_1 l_2 l_3} = \sum_{p_j} \alpha_{p_1 p_2 p_3} \int dr \tilde{q}_{p_1}(l_1,r) \tilde{q}_{p_2}(l_2,r) \tilde{q}_{p_3}(l_3,r), \tag{5.4}$$

where the *projected* mode functions are defined as

$$\tilde{q}_p(l,r) := \frac{2r^{\frac{2}{3}}}{\pi} \int dk q_p(k) \Delta_l(k) j_l(kr). \tag{5.5}$$

Radiative transfer functions $\Delta_l(k)$ and spherical Bessel functions $j_l(kr)$ are denoted the same way as the previous chapter.

Every term appearing in (5.1) except the Gaunt integral is now separable. Using the definition $\mathscr{G}^{l_1 l_2 l_3}_{m_1 m_2 m_3} = \int d^2\mathbf{n} Y_{l_1 m_1}(\mathbf{n}) Y_{l_2 m_2}(\mathbf{n}) Y_{l_3 m_3}(\mathbf{n})$, we can render it separable at the cost of introducing an extra integral.

Define the filtered maps as

$$M^{(i)}_p(\mathbf{n},r) := \sum_{l,m} \frac{\tilde{q}_p(l,r)}{C_l} a^{(i)}_{lm} Y_{lm}(\mathbf{n}), \tag{5.6}$$

where $a^{(i)}_{lm}$'s are represent the spherical harmonic transform of the ith CMB map. The convention is so that 0th map corresponds to the observed CMB map, while maps number 1-N_{sims} are Gaussian simulations. Without the factors involving $\tilde{q}$ and C_l's, M is simply equal to the original map in real space. Each mode extracts different anisotropy scales present in the map.

The bispectrum estimator (5.1) reduces to

$$\hat{f}_{NL} = \frac{1}{N} \sum_{p_j} \alpha_{p_1 p_2 p_3} (\beta^{cub}_{p_1 p_2 p_3} - 3\beta^{lin}_{p_1 p_2 p_3}), \tag{5.7}$$

where most of the computation required is now contained in the 'β's, given by

$$\beta^{cub}_{p_1p_2p_3} := \int dr \int d^2\mathbf{n}\, M^{(0)}_{p_1}(\mathbf{n},r) M^{(0)}_{p_2}(\mathbf{n},r) M^{(0)}_{p_3}(\mathbf{n},r), \tag{5.8}$$

$$\beta^{lin}_{p_1p_2p_3} := \frac{1}{N_{sims}} \sum_{j=1}^{N_{sim}} \int dr \int d^2\mathbf{n}\, M^{(j)}_{p_1}(\mathbf{n},r) M^{(j)}_{p_2}(\mathbf{n},r) M^{(0)}_{p_3}(\mathbf{n},r). \tag{5.9}$$

It is important to note that the beta matrices depends only on the choice of mode functions and input map data, and are independent of the theoretical bispectrum considered. Once β^{cub} and β^{lin} are computed and stored, we may constrain any model of interest by decomposing the template to get α, and then simply taking a dot product: $\alpha \cdot \beta / N$.

The normalisation can also be obtained in a similar fashion;

$$N = \sum_{p_j, p'_j} \alpha_{p_1p_2p_3} \Gamma_{p_1p_2p_3, p'_1, p'_2, p'_3} \alpha_{p'_1, p'_2, p'_3}, \tag{5.10}$$

or equivalently, $N = \alpha^T \Gamma \alpha$. We exploit separability once again to compute the Γ matrix;

$$\Gamma_{p_1p_2p_3, p'_1, p'_2, p'_3} := \int dr \int d\mu\, \mathscr{P}_{p_1p'_1}(\mu, r, r') \mathscr{P}_{p_3p'_3}(\mu, r, r') \mathscr{P}_{p_3p'_3}(\mu, r, r'), \tag{5.11}$$

$$\mathscr{P}_{pp'}(\mu, r, r') := \sum_l \frac{2l+1}{(8\pi)^{1/3} C_l} \tilde{q}'_p(l, r) \tilde{q}_p(l, r') P_l(\mu), \tag{5.12}$$

where $P_l(\mu)$'s are the Legendre polynomials.

In summary, CMB-BEst computes the main quantities: β^{cub}, β^{lin}, and Γ. The most computationally expensive part is the linear term β^{lin} by a couple orders of magnitude in most cases. Considerable effort has been put to optimise corresponding part of the code, which will be detailed in the following sections.

5.1.3 Basis functions

Special case (pmax=2) -> ones used in forecast for constant feature models. KSW choice (pmax=4) Legendre functions basis. Augmentation - 1/k basis.

Decomposition. Separability -> three separate integrations

5.2 Implementation and optimisation

The CMB-BEst formalism significantly reduces the amount of computation needed for the CMB bispectrum estimation. Obtaining the linear term β^{lin} in (5.9), however, is still practically impossible unless thoroughly optimised. In this section, we provide details for various aspects of our optimisation process: algorithm design, parallel computing, and data locality improvements. Final specifications and data files used are outlined at the end.

5.2.1 Algorithm

Basic order of integration, psuedo-codes, justifications Data used, programs used, etc.

Algorithm 1 Computing βs: the naïve method

Allocate $M(i,p,n)$ $\triangleright$ Memory $\sim O(N_{sims} \cdot p_{max} \cdot N_{pix})$

for each map i **do**
    for each mode p **do** $\triangleright$ $O(N_{sims} \cdot p_{max} \cdot N_{pix}^{3/2})$
        compute $M(i,p,n)$ by SHT
    end for
end for $\triangleright$ $M(i,p,n)$ ready

for each map i **do**
    for each set of modes (p_1,p_2,p_3) **do**
        for each pixel n **do** $\triangleright$ $O(N_{sims} \cdot p_{max}^3 \cdot N_{pix})$
            $\beta^{cub}(i,p_1,p_2,p_3) \mathrel{+}= M(i,p_1,n) \cdot M(i,p_2,n) \cdot M(i,p_3,n)$
        end for
    end for
end for

for each map i **do**
    for each map $j \neq i$ **do**
        for each set of modes (p_1,p_2,p_3) **do**
            for each pixel n **do** $\triangleright$ $O(N_{sims}^2 \cdot p_{max}^3 \cdot N_{pix})$
                $\beta^{lin}(i,p_1,p_2,p_3) \mathrel{+}= M(j,p_1,n) \cdot M(j,p_2,n) \cdot M(i,p_3,n)$
            end for
        end for
    end for
end for

(Afterwards, subtract beta cubic from beta linear and divide it by Nsims-1)

5.2.2 Parallel computing

Table: 3 levels of parallelisation. How CMB-BEst exploits each level.

Algorithm 2 Computing βs: optimised for computation

Allocate $M(i,p,n)$ ▷ $O(N_{sims} \cdot p_{max} \cdot N_{pix})$
Allocate $C(p_1,p_2,n)$ ▷ $O(p_{max} \cdot p_{max} \cdot N_{pix})$

for each map i **do**
  for each mode p **do** ▷ $O(N_{sims} \cdot p_{max} \cdot N_{pix}^{3/2})$
    compute $M(i,p,n)$ by SHT
  end for
end for ▷ $M(i,p,n)$ ready

for each map j **do**
  for each pair of modes (p_1,p_2) **do**
    for each pixel n **do** ▷ $O(N_{sims} \cdot p_{max}^2 \cdot N_{pix})$
      $C(p_1,p_2,n) \mathrel{+}= M(j,p_1,n) \cdot M(j,p_2,n)$
    end for
  end for
end for ▷ $C(p_1,p_2,n)$ ready

for each map i **do**
  for each set of modes (p_1,p_2,p_3) **do**
    for each pixel n **do** ▷ $O(N_{sims} \cdot p_{max}^3 \cdot N_{pix})$
      $\beta^{cub}(i,p_1,p_2,p_3) \mathrel{+}= M(i,p_1,n) \cdot M(i,p_2,n) \cdot M(i,p_3,n)$
      $\beta^{lin}(i,p_1,p_2,p_3) \mathrel{+}= C(p_1,p_2,n) \cdot M(i,p_3,n)$
    end for
  end for
end for

Algorithm 3 Computing βs: our final version

Allocate $m(p,n)$ $\triangleright\ O(p_{max} \cdot N_{pix})$
Allocate $C(p_1,p_2,n)$ $\triangleright\ O(p_{max} \cdot p_{max} \cdot N_{pix})$

for each map i **do**
  for each mode p **do** $\triangleright\ O(N_{sims} \cdot p_{max} \cdot N_{pix}^{3/2})$
    compute $M(i,p,n)$ by SHT and store in $m(p,n)$
  end for

  for each pair of modes (p_1,p_2) **do** $\triangleright\ O(N_{sims} \cdot p_{max}^2 \cdot N_{pix})$
    for each pixel n **do**
      $C(p_1,p_2,n) \mathrel{+}= m(p_1,n) \cdot m(p_2,n)$
    end for
  end for
end for $\triangleright\ C(p_1,p_2,n)$ ready

for each map i **do**
  for each of mode p **do** $\triangleright\ O(N_{sims} \cdot p_{max} \cdot N_{pix}^{3/2})$
    compute $M(i,p,n)$ by SHT and store in $m(p,n)$
  end for

  for each set of modes (p_1,p_2,p_3) **do**
    for each pixel n **do** $\triangleright\ O(N_{sims} \cdot p_{max}^3 \cdot N_{pix})$
      $\beta^{cub}(i,p_1,p_2,p_3) \mathrel{+}= m(p_1,n) \cdot m(p_2,n) \cdot m(p_3,n)$
      $\beta^{lin}(i,p_1,p_2,p_3) \mathrel{+}= C(p_1,p_2,n) \cdot m(p_3,n)$
    end for
  end for
end for

5.2.3 Data locality

Cache blocking. Some Amplxe results screenshot to confirm? + Initialising parts of large arrays at processors that actually use them.

5.2.4 Cluster adaptation

Three different versions of the code. Balance between performance / memory.

5.2.5 Final specifications

Code, runtime, etc?

5.3 Verification

5.3.1 Internal consistency checks

Quick derivation of optimality of the bispectrum estimator? Sample versus Theory covariance checks. Expected chi-square distribution etc.

5.3.2 Consistency with Planck

Consistency vs Modal for Planck results. Discussion on kmin and kmax range

References

[1] Achúcarro, A., Gong, J. O., Hardeman, S., Palma, G. A., and Patil, S. P. (2011). Features of heavy physics in the CMB power spectrum. *Journal of Cosmology and Astroparticle Physics*, 2011(1).

[2] Adams, J., Easther, R., and Cresswell, B. (2001). Inflationary perturbations from a potential with a step. *Physical Review D - Particles, Fields, Gravitation and Cosmology*, 64(12):6.

[3] Adshead, P., Dvorkin, C., Hu, W., and Lim, E. A. (2012). Non-Gaussianity from step features in the inflationary potential. *Physical Review D - Particles, Fields, Gravitation and Cosmology*, 85(2):1–26.

[4] Ballardini, M., Finelli, F., Fedeli, C., and Moscardini, L. (2016). Probing primordial features with future galaxy surveys. *Journal of Cosmology and Astroparticle Physics*, 2016(10).

[5] Bartolo, N., Cannone, D., and Matarrese, S. (2013). The effective field theory of inflation models with sharp features. *Journal of Cosmology and Astroparticle Physics*, 2013(10).

[6] Benetti, M. and Alcaniz, J. S. (2016). Bayesian analysis of inflationary features in Planck and SDSS data. *Physical Review D*, 94(2):1–8.

[7] Benetti, M., Lattanzi, M., Calabrese, E., and Melchiorri, A. (2011). Features in the primordial spectrum: New constraints from WMAP7 and ACT data and prospects for the Planck mission. *Physical Review D - Particles, Fields, Gravitation and Cosmology*, 84(6):1–8.

[8] Chantavat, T., Gordon, C., and Silk, J. (2011). Large scale structure forecast constraints on particle production during inflation. *Physical Review D - Particles, Fields, Gravitation and Cosmology*, 83(10):1–9.

[9] Chen, X. (2010). Primordial non-gaussianities from inflation models. *Advances in Astronomy*, 2010(ii).

[10] Chen, X., Dvorkin, C., Huang, Z., Namjoo, M. H., and Verde, L. (2016). The future of primordial features with large-scale structure surveys. *Journal of Cosmology and Astroparticle Physics*, 2016(11).

[11] Chen, X., Easther, R., and Lim, E. A. (2007). Large non-Gaussianities in single-field inflation. *Journal of Cosmology and Astroparticle Physics*, 06(6).

[12] Chluba, J., Hamann, J., and Patil, S. P. (2015). Features and New Physical Scales in Primordial Observables: Theory and Observation. 24(10):1–133.

[13] Dicke, R. H., Peebles, P. J. E., Roll, P. G., and Wilkinson, D. T. (1965). Cosmic black-body radiation. *The Astrophysical Journal*, 142:414–419.

[14] Dvorkin, C. and Hu, W. (2010). Generalized slow roll approximation for large power spectrum features. *Physical Review D - Particles, Fields, Gravitation and Cosmology*, 81(2):1–14.

[15] et al, K. N. A. (2016). CMB-S4 Science Book, First Edition.

[16] Fergusson, J. R. (2014). Efficient optimal non-Gaussian CMB estimators with polarization. *Physical Review D - Particles, Fields, Gravitation and Cosmology*, 90(4).

[17] Fergusson, J. R., Gruetjen, H. F., Shellard, E. P., and Liguori, M. (2015a). Combining power spectrum and bispectrum measurements to detect oscillatory features.

[18] Fergusson, J. R., Gruetjen, H. F., Shellard, E. P., and Wallisch, B. (2015b). Polyspectra searches for sharp oscillatory features in cosmic microwave sky data. *Physical Review D - Particles, Fields, Gravitation and Cosmology*, 91(12).

[19] Fergusson, J. R., Liguori, M., and Shellard, E. P. S. (2012). The CMB bispectrum. *Journal of Cosmology and Astroparticle Physics*, 2012(12).

[20] Hale, N. and Townsend, A. (2013). Fast and Accurate Computation of Gauss–Legendre and Gauss–Jacobi Quadrature Nodes and Weights. *SIAM journal on scientific computing*, 35(2):A652—-A674.

[21] Hazra, D. K., Shafieloo, A., Smoot, G. F., and Starobinsky, A. A. (2014). Wiggly whipped inflation. *Journal of Cosmology and Astroparticle Physics*, 2014(8).

[22] Hu, B. and Torrado, J. (2015). Searching for primordial localized features with CMB and LSS spectra. *Physical Review D - Particles, Fields, Gravitation and Cosmology*, 91(6):1–10.

[23] Komatsu, E. (2010). Hunting for Primordial Non-Gaussianity in the Cosmic Microwave Background. *Classical and Quantum Gravity*, 27.

[24] Komatsu, E., Spergel, D. N., and Wandelt, B. D. (2005). Measuring Primordial Non-Gaussianity in the Cosmic Microwave Background. *The Astrophysical Journal*, 634(1):14–19.

[25] Lewis, A., Challinor, A., and Lasenby, A. (2000). Efficient Computation of Cosmic Microwave Background Anisotropies in Closed Friedmann-Robertson-Walker Models. *The Astrophysical Journal*, 538(2):473–476.

[26] Liguori, M., Sefusatti, E., Fergusson, J. R., and Shellard, E. P. (2010). Primordial non-gaussianity and bispectrum measurements in the cosmic microwave background and large-scale structure. *Advances in Astronomy*, 2010.

[27] Maldacena, J. (2002). Non-Gaussian features of primordial fluctuations in single field inflationary models. *Journal of High Energy Physics*, 2003(05):13.

[28] Martin, J. and Ringeval, C. (2004). Superimposed oscillations in the WMAP data? *Physical Review D - Particles, Fields, Gravitation and Cosmology*, 69(8):9.

[29] Meerburg, P. D., Spergel, D. N., and Wandelt, B. D. (2014a). Searching for oscillations in the primordial power spectrum. I. Perturbative approach. *Physical Review D - Particles, Fields, Gravitation and Cosmology*, 89(6):19–26.

[30] Meerburg, P. D., Spergel, D. N., and Wandelt, B. D. (2014b). Searching for oscillations in the primordial power spectrum. II. Constraints from Planck data. *Physical Review D - Particles, Fields, Gravitation and Cosmology*, 89(6):1–6.

[31] Meerburg, P. D., Van Der Schaar, J. P., and Corasaniti, P. S. (2009). Signatures of initial state modifications on bispectrum statistics. *Journal of Cosmology and Astroparticle Physics*, 2009(5).

[32] Meerburg, P. D., Wijers, R. A., and van der Schaar, J. P. (2012). WMAP7 constraints on oscillations in the primordial power spectrum. *Monthly Notices of the Royal Astronomical Society*, 421(1):369–380.

[33] Miranda, V., Hu, W., and Adshead, P. (2012). Warp features in DBI inflation. *Physical Review D - Particles, Fields, Gravitation and Cosmology*, 86(6):1–10.

[34] Münchmeyer, M., Bouchet, F., Jackson, M. G., and Wandelt, B. (2014). The Komatsu Spergel Wandelt estimator for oscillations in the cosmic microwave background bispectrum. *Astronomy & Astrophysics*, 570:A94.

[35] Ng, K. W. and Liu, G. C. (1999). Correlation Functions of Cmb Anisotropy and Polarization. *International Journal of Modern Physics D*, 08(01):61–83.

[36] Penzias, A. A. and Wilson, R. W. (1965). A measurement of excess antenna temperature at 4080 mc/s. *The Astrophysical Journal*, 142:419–421.

[37] Planck Collaboration (2015). Planck 2015 results. XVII. Constraints on primordial non-Gaussianity. *Astronomy & Astrophysics*, 594.A17.

[38] Planck Collaboration (2020a). Planck 2018 results. ix. constraints on primordial non-gaussianity. *Astronomy & Astrophysics*, 641:A9.

[39] Planck Collaboration (2020b). Planck 2018 results. vi. cosmological parameters. *Astronomy & Astrophysics*, 641:A6.

[40] Smith, K. M. and Zaldarriaga, M. (2011). Algorithms for bispectra: Forecasting, optimal analysis and simulation. *Monthly Notices of the Royal Astronomical Society*, 417(1):2–19.

[41] Starobinsky, A. (1992). Aa starobinsky, jetp lett. 55, 489 (1992). *JETP Lett.*, 55:489.

[42] The COrE collaboration (2015). COrE (Cosmic Origins Explorer) White Paper.

[43] The Simons Observatory Collaboration (2018). The Simons Observatory: Science goals and forecasts.

[44] Yadav, A. P. S., Komatsu, E., and Wandelt, B. D. (2007). Fast Estimator of Primordial Non-Gaussianity from Temperature and Polarization Anisotropies in the Cosmic Microwave Background. pages 1–15.