Measuring the Dark Universe with Gravitational Lensing

Rupert Allison

Department of Physics and New College, Oxford

A thesis submitted for the degree of Doctor of Philosophy
at the University of Oxford.

Hilary Term 2016

Measuring the Dark Universe with Gravitational Lensing

Rupert Allison
New College

Submitted for the Degree of Doctor of Philosophy
Hilary Term 2016

Abstract

This thesis presents the use of gravitational lensing as a measure of the large-scale structure in the Universe. We detail various statistical techniques for performing and interpreting these measurements. We show evidence for the cross-correlation of cosmic microwave background (CMB) lensing with the positions of radio-loud active galactic nuclei (AGN). We demonstrate a weak-lensing cross-correlation pipeline for combining cosmic shear measurements with CMB lensing data. We forecast how cosmological parameters may be constrained through gravitational lensing measurements, with a focus on the sum of the neutrino masses, and discuss practical methods for the statistical extraction of information from data.

We firstly cover the theoretical framework within which this thesis is set. We then make a comparison of sampling techniques for Bayesian parameter estimation. We find that nested sampling delivers high-fidelity estimates for posterior statistics at low computational cost, and should be adopted in favour of Metropolis-Hastings sampling in many cases. Affine-invariant MCMC is competitive when computer clusters can be utilised for massive parallelisation and, along with existing extensions to nested sampling, naturally probes multi-modal and curving distributions.

Next, we correlate the positions of radio sources in the FIRST survey with CMB lensing convergence measured by the Atacama Cosmology Telescope, over 470 deg^2, to determine the bias of these galaxies. We measure the angular cross-power spectrum $C_l^{\kappa g}$ at 4.4σ significance in the multipole range $100 < \ell < 3000$. We fit for the overall bias-model normalisation, finding $b(z_{\rm eff}) = 3.5 \pm 0.8$ for the full sample at an effective redshift $z_{\rm eff} = 1.5$.

We then present comprehensive forecasts for how the sum of the neutrino masses will be constrained with future cosmological observations. We consider prospects for the gravitationally-lensed CMB anisotropies and baryon acoustic oscillations (BAOs) in the galaxy distribution, examining how the projected uncertainty of ≈ 15 meV on the neutrino mass sum (a 4σ detection of the minimal mass) might be reached over the next decade. We find an improved optical depth measurement is important: the projected neutrino mass uncertainty increases to 26 meV if a CMB Stage 4 experiment is limited to $\ell > 20$ and combined with current large-scale polarisation data. Complementary low-redshift probes, including galaxy lensing, will play a role in distinguishing between massive neutrinos and a departure from a $w = -1$, flat geometry.

Finally, we present a forecasting and analysis pipeline for weak-lensing cross-correlations. We introduce and motivate the joint analysis of CMB lensing and galaxy lensing, including a discussion of measurements to-date and key systematic effects in lensing data. We argue that simultaneous analysis provides a robust approach to calibration of these effects. We review the theory of weak-lensing cross-correlations, and illustrate the properties of the observables. We detail how the shear signal is extracted from the shapes of galaxies and show how this is combined with CMB lensing data to measure their cross power spectrum. We demonstrate the pipeline we have developed – from map making to parameter estimation – on real and simulated data from ACT, CFHTLenS and KiDS. We conclude by forecasting the ability of future experiments to detect the lensing cross-correlation.

Declaration

I declare that no part of this thesis has been accepted, or is currently being submitted, for any degree or diploma or certificate or any other qualification in this University or elsewhere. Except where explicit reference is made to the work of others, the work contained in this thesis is my own.

Chapter 2 is based on work I led that was published in the Monthly Notices of the Royal Astronomical Society: *Comparison of sampling techniques for Bayesian parameter estimation*, Allison & Dunkley (2014). Minor edits reflect comments and suggestions from the co-author and referee.

Chapter 3 is based on work I led that was published in the Monthly Notices of the Royal Astronomical Society: *The Atacama Cosmology Telescope: measuring radio galaxy bias through cross-correlation with lensing*, Allison et al. (2015). Minor edits reflect comments and suggestions from co-authors and the referee.

Chapter 4 is based on work published in Physical Review D: *Towards a cosmological neutrino mass detection*, Allison et al. (2015). I led all aspects of this work except the implementation of the BAO Fisher matrix code, which was done by Paul Caucal. Minor edits reflect comments and suggestions from co-authors and the referee.

Chapter 5 is based on work I solely produced which is to appear in the *Proceedings for Hot Topics in General Relativity and Gravitation 2*, Rencontres du Vietnam, August 2015.

Rupert Allison

(February 2016)

Acknowledgements

Firstly, and most importantly, I thank Jo Dunkley, who has been the most supportive and understanding supervisor that I could have asked for. For the last three and half years I have learned so much from her and have been consistently inspired by her intelligence and ideas. Jo has always had time for me – to discuss work, and to guide me through academia – and she has unfailingly forgiven my absences when away pursuing my hobbies. I will always be grateful for her enthusiasm and kindness.

Also within the group at Oxford, thank you to Danielle, David, Elisa, Graeme and Pedro, and of course to Erminia – *grazie mille*; to Sigurd – *tusen takk*. Thanks to Paul Caucal for his work and generally being a top chap. Thibaut, also, has always been there to discuss science and life over lunch, a lap of the DWB, or (most enthusiastically) in the nearest ale house. He has been a constant friend and the source of many a good laugh – thank you. Thanks variously to Phil Bull, Blake Sherwin, Nick Hand and the ACT collaboration. I am grateful to Paul Hewett, who set me on the road of scientific research, and who has always had a kind and wise word for me.

Thanks to my fellow office 512 inhabitants over the years: Adam, Silvio, Phil, Holly, Macarena, Simon, Jamie, Ben, Marisa, the Dragon and Lionel Richie, who have kept me sane, caffeine-fuelled and entertained, even in the darkest moments. I am also indebted to Ashling, who has been a champion – in all senses of the word – and a friendly face from the moment I started in Oxford.

To all my friends from *l'École de Physique des Houches 2013*, particularly Agnès and *mon frère* Julian, thanks for the summer of a lifetime and your continuing friendship. Thanks also to my friends over the years in OUHC who have made my time in Oxford so much richer, and to Elle and Richard for sharing the ups and the downs and for putting up with me.

From my school days, special thanks to Mr. Joyce, Mr. Oldale and Mr. Senior, who helped to encourage my interest in maths and physics, and to Jude and Mr. Wright, who were of no help at all in that department, but did a lot else for me besides.

Thanks to Lizzy for being there for me, for listening to me blab about blobs in the CMB, and for adding so much to my universe outside of physics. Finally, thank you – although it will never be enough – to Mum and Dad, who have always encouraged, supported and strived for me and Joe.

The Road goes ever on and on
Down from the door where it began.
Now far ahead the Road has gone,
And I must follow, if I can,
Pursuing it with eager feet,
Until it joins some larger way
Where many paths and errands meet.
And whither then? I cannot say.

– *J. R. R. Tolkien*

Contents

List of Figures

List of Tables

List of Abbreviations

The following abbreviations are used throughout the text:

ACT	Atacama Cosmology Telescope
ACTPol	Atacama Cosmology Telescope Polarimeter
AdvACT	Advanced ACTPol
CMB	cosmic microwave background
CV	cosmic variance
d.o.f.	degrees of freedom
Eq.	equation
Fig.	figure
FRW	Friedman-Robertson-Walker
GR	General Relativity
ISW	integrated Sachs-Wolfe
PTE	probability to exceed
rms	root mean-square
s.c.	summation convention
SPT	South Pole Telescope

Conventions

The following conventions and notation are used in this thesis:

- We adopt a background Friedmann-Robertson-Walker (FRW) metric throughout. The line element is given by

$$ds^2 = c^2 dt^2 - a^2(t) \left(d\chi^2 + f_k^2(\chi)[d\theta^2 + \sin^2\theta d\varphi^2] \right), \tag{1}$$

  for cosmological time t, scale factor a (dimensionless with $a = 1$ today), radial comoving distance χ, spherical polar angles θ and φ, Gaussian curvature k, and transverse comoving distance

$$f_k(\chi) = \begin{cases} k^{-1/2} \sin(k^{1/2}\chi) & \text{for } k > 0, \\ \chi & \text{for } k = 0, \\ |k|^{-1/2} \sinh(|k|^{1/2}\chi) & \text{for } k < 0. \end{cases} \tag{2}$$

- Considering scalar perturbations to the background metric, these can be written in the conformal Newtonian gauge as

$$ds^2 = a^2(\eta) \left[(1 + 2\psi) d\eta^2 - (1 - 2\phi) \gamma_{ij} dx^i dx^j \right], \tag{3}$$

  where η is conformal time, defined such that $a d\eta = c dt$, which has units of length, and $\gamma_{ij} dx^i dx^j = d\chi^2 + f_k^2(\chi)[d\theta^2 + \sin^2\theta d\varphi^2]$. We denote the conformal time today by η_0.

- The *Weyl potential* – relevant for gravitational lensing – is

$$\Psi = \frac{\psi + \phi}{2}. \tag{4}$$

- Many of the plots and numerical values given throughout this thesis are calculated under the assumption of a given cosmological model. Unless otherwise specified, the fiducial cosmological model adopted is ΛCDM with the following parameters: physical baryon density $\Omega_b h^2 = 0.02222$, physical cold-dark-matter density $\Omega_c h^2 = 0.1197$, Hubble constant $H_0 = 67.31$ km s^{-1} Mpc^{-1}, optical depth to reionisation $\tau = 0.078$, amplitude of primordial scalar perturbations $A_s = 2.1955 \times 10^{-9}$, tilt of the primordial scalar power $n_s = 0.9655$ at pivot scale $k_0 = 0.05$ Mpc^{-1}, no tensor-mode contribution ($r = 0$), a single massive neutrino of mass $\Sigma m_\nu = 60$ meV and an effective 2.046 massless neutrinos. These parameters correspond to the best-fit *Planck* 2015 TT+lowP cosmology (*Planck* Collaboration, 2015c). The angular size of the sound horizon at recombination is a derived parameter: $\theta_{MC} = 1.04084$.

- For three-dimensional fields defined in real-space $f(\mathbf{x})$ we define the Fourier transform $f(\mathbf{k})$ as follows:

$$f(\mathbf{k}) = \mathcal{FT}[f](\mathbf{k}) \equiv \int \frac{d^3\mathbf{x}}{(2\pi)^{3/2}} f(\mathbf{x}) e^{-i\mathbf{k}\cdot\mathbf{x}}. \tag{5}$$

  The three-dimensional Fourier wavevector is denoted by $\mathbf{k}$, with amplitude $k = |\mathbf{k}|$. Abusing notation, the argument distinguishes the Fourier transform $f(\mathbf{k})$ from the function $f(\mathbf{x})$ itself. The inverse Fourier transform may then be written

$$f(\mathbf{x}) = \mathcal{FT}^{-1}[f](\mathbf{x}) \equiv \int \frac{d^3\mathbf{k}}{(2\pi)^{3/2}} f(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}}, \tag{6}$$

  using the following identity for the three-dimensional Dirac delta function:

$$(2\pi)^3 \delta(\mathbf{x} - \mathbf{x}') = \int d^3\mathbf{k}\; e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')}. \tag{7}$$

  With this normalisation convention the convolution theorem is given by

$$\mathcal{FT}[f * g](\mathbf{k}) = (2\pi)^{3/2} f(\mathbf{k}) g(\mathbf{k}). \tag{8}$$

- In two-dimensions these definitions are modified to:

$$f(\boldsymbol{\ell}) = \mathcal{FT}[f](\boldsymbol{\ell}) \equiv \int \frac{d^2\mathbf{x}}{2\pi} f(\mathbf{x}) e^{-i\boldsymbol{\ell}\cdot\mathbf{x}} \tag{9}$$

and

$$f(\mathbf{x}) = \mathcal{FT}^{-1}[f](\mathbf{x}) \equiv \int \frac{d^2\boldsymbol{\ell}}{2\pi} f(\boldsymbol{\ell}) e^{i\boldsymbol{\ell}\cdot\mathbf{x}}. \tag{10}$$

The two-dimensional Fourier wavevector is denoted by $\boldsymbol{\ell}$, with amplitude $\ell = |\boldsymbol{\ell}|$. The analogue of Eq. 7, in two dimensions, has the replacement $(2\pi)^3 \to (2\pi)^2$ and with the adopted normalisation convention the convolution theorem reads

$$\mathcal{FT}[f * g](\boldsymbol{\ell}) = 2\pi f(\boldsymbol{\ell}) g(\boldsymbol{\ell}). \tag{11}$$

Chapter 1

Introduction

In this chapter we cover the theoretical framework within which this thesis is set. We introduce many of the concepts and notation which will be used subsequently throughout the text. In Sec. 1.1 we begin with a mathematical description of the cosmological principle, random fields, power spectra and correlation functions. We then present the standard picture of a ΛCDM universe – the currently favoured cosmological model – and the related topics of inflation and structure formation (Sec. 1.2). The cosmic microwave background is central to this thesis; its origin and a quantitative description of its properties are detailed in Sec. 1.3. The concept of gravitational lensing is introduced in Sec. 1.4 and the idea of galaxies as biased tracers of the underlying matter field is explained in Sec. 1.5. We discuss the use of multiple dark matter tracers and cross-correlations in Sec. 1.6. We then cover the basic statistical methods used in this thesis to extract or forecast the useful information in data (Sec. 1.7–1.8). In Sec. 1.9 we describe in some detail many of the relevant contemporary and near-future experiments, their observations and scientific output. We finish with a (non-exhaustive) list and discussion of some interesting outstanding problems in cosmology (Sec. 1.10).

1.1 Random fields, homogeneity and isotropy

In cosmology we deal extensively with the concept of a *random field*, which is defined by a set of random variables indexed at every point in some N-dimensional space. For instance, the variation of the matter density in the universe can be modelled as a scalar (single-valued) random

field defined everywhere in three-dimensional space (indexed, for example, by comoving position $\mathbf{x}$) . Importantly, cosmological theory cannot predict the exact realisation of large-scale structure that we might observe, but it does make concrete statements about the *statistics* of the matter field and how they evolve over time.

A given theoretical model defines a *probability distribution* for the random field, $\mathbb{P}[f]$, from which all its statistics (e.g., its mean and covariance) are derivable, and these can then be compared against observations to test the viability of that model. For example, the *concordance* ΛCDM model includes the ansatz that the initial density perturbations were Gaussian, with a near-scale-invariant power spectrum (Sec. 1.2.2). On cosmological scales, therefore, this model specifies that the matter density can be described a Gaussian random field, with a power spectrum prescribed by linear perturbation theory. At late times (and on small-scales) this assumption breaks down, as structure begins to grow under the non-linear action of gravity (Sec 1.2.5).

The cosmological principle states that the universe is statistically *homogeneous* and *isotropic* (Peacock, 1998). Homogeneity of a random field is defined such that the statistical properties of the field are invariant under arbitrary translations. Isotropy of a random field is defined such that its statistical properties are invariant under arbitrary rotations. We now define more concretely some of the concepts introduced above.

1.1.1 Three-dimensional random fields

We consider a (scalar) random field f defined at each point $\mathbf{x}$ in three-dimensional space with a probability distribution $\mathbb{P}[f]$. Here $\mathbb{P}[f]$ is a functional: a probability density over functions f such that $\int df\ \mathbb{P}[f] = 1$ and $\forall f,\ \mathbb{P}[f] \in [0, \infty)$. The n-point correlation function of f is defined by

$$\langle f(\mathbf{x}_1) f(\mathbf{x}_2) ... f(\mathbf{x}_n) \rangle \equiv \int df\ \mathbb{P}[f] f(\mathbf{x}_1) f(\mathbf{x}_2) ... f(\mathbf{x}_n). \tag{1.1}$$

The translation operator $T_{\mathbf{r}}$ and rotation operator R are defined such that:

$$T_{\mathbf{r}} f(\mathbf{x}) \equiv f(\mathbf{x} - \mathbf{r}), \tag{1.2}$$

$$R f(\mathbf{x}) \equiv f(\mathsf{R}^{-1}\mathbf{x}), \tag{1.3}$$

where $\mathbf{r}$ is a displacement vector and R a rotation matrix. Mathematically, we say a random field is *homogeneous* (H) if $\mathbb{P}[f] = \mathbb{P}[T_{\mathbf{r}} f]\ \forall \mathbf{r}$ (e.g., Challinor, 2009). Similarly, a random field is said to be *isotropic* (I) if $\mathbb{P}[f] = \mathbb{P}[Rf]\ \forall R$. It is easy to show (the expected result that) the 1-point function – the *mean* over realisations f – of a homogenous field is the same everywhere:

$$\begin{aligned}
\langle f(\mathbf{x}-\mathbf{r})\rangle &= \int df\ \mathbb{P}[f] f(\mathbf{x}-\mathbf{r}) \\
&= \int df\ \mathbb{P}[f] T_{\mathbf{r}} f(\mathbf{x}) \\
&= \int df\ \mathbb{P}[T_{\mathbf{r}} f] T_{\mathbf{r}} f(\mathbf{x}) \qquad \text{(by homogeneity)} \\
&= \int df\ \mathbb{P}[f] f(\mathbf{x}) \qquad \text{(relabelling } T_{\mathbf{r}} f \to f) \\
&= \langle f(\mathbf{x})\rangle \quad \forall \mathbf{r}.
\end{aligned} \tag{1.4}$$

In a similar fashion, one can show that the 2-point correlation function $\xi(\mathbf{x}, \mathbf{x}') \equiv \langle f(\mathbf{x}) f(\mathbf{x}')\rangle$ of a homogeneous field f depends only on the vector separation between the arguments, $\mathbf{x} - \mathbf{x}'$. Furthermore, if additionally f is isotropic then the correlation function depends only on the *distance* between the two points, $|\mathbf{x} - \mathbf{x}'|$, i.e.,

$$\xi(\mathbf{x}, \mathbf{x}') = \xi(|\mathbf{x} - \mathbf{x}'|) \qquad \text{(given } H \ \&\ I) \tag{1.5}$$

only. For a homogeneous field the Fourier-space counterpart of the 2-point correlation function is

$$\begin{aligned}
\langle f(\boldsymbol{k}) f(\boldsymbol{k}')\rangle &= \int \frac{d^3\mathbf{x}}{(2\pi)^{3/2}} \frac{d^3\mathbf{x}'}{(2\pi)^{3/2}} \langle f(\mathbf{x}) f(\mathbf{x}')\rangle e^{-i\mathbf{k}\cdot\mathbf{x}} e^{-i\mathbf{k}'\cdot\mathbf{x}'} \\
&= \int \frac{d^3\mathbf{x}}{(2\pi)^{3/2}} \frac{d^3\mathbf{x}'}{(2\pi)^{3/2}} \xi(\mathbf{x}-\mathbf{x}') e^{-i\mathbf{k}\cdot\mathbf{x}} e^{-i\mathbf{k}'\cdot\mathbf{x}'} \qquad \text{(by homogeneity)} \\
&= \int \frac{d^3\mathbf{x}}{(2\pi)^{3/2}} \frac{d^3\mathbf{r}}{(2\pi)^{3/2}} \xi(\mathbf{r}) e^{-i\mathbf{x}\cdot(\mathbf{k}+\mathbf{k}')} e^{i\mathbf{k}'\cdot\mathbf{r}} \qquad \text{(replacing } \mathbf{x}' \to \mathbf{x} - \mathbf{r}) \\
&= \delta(\mathbf{k}+\mathbf{k}') \int d^3\mathbf{r}\ \xi(\mathbf{r}) e^{-i\mathbf{k}\cdot\mathbf{r}} \\
&= \delta(\mathbf{k}+\mathbf{k}') P(\mathbf{k}).
\end{aligned} \tag{1.6}$$

We see that homogeneity implies that the two-point function in Fourier space is *diagonal*, so that distinct modes are uncorrelated, and the mode variance depends in general on all components of

the wavenumber $\mathbf{k}$. Here we have defined the *power spectrum*

$$P(\mathbf{k}) \equiv \int d^3\mathbf{r}\ \xi(\mathbf{r}) e^{-i\mathbf{k}\cdot\mathbf{r}}, \tag{1.7}$$

which is the Fourier transform of the 2-point correlation function. This quantity is well-defined for all three-dimensional fields, but for homogeneous fields it is particularly useful since it fully quantifies the mode covariance (through Eq. 1.6). Further assuming that the field is isotropic, one can perform the angular integration in Eq. 1.7; the power spectrum P is then a function only of the *magnitude* of the wavenumber k:

$$P(k) = 4\pi \int_0^\infty dr\ r^2 \xi(r) j_0(kr), \tag{1.8}$$

so that

$$\langle f(\boldsymbol{k}) f(\boldsymbol{k}') \rangle = P(k)\delta(\mathbf{k} + \mathbf{k}') \qquad \text{(given } H \text{ \& } I\text{)}. \tag{1.9}$$

Using the orthogonality relation for the spherical Bessel functions j_ℓ, we can invert Eq. 1.8 to write

$$\xi(r) = \frac{1}{2\pi^2} \int_0^\infty dk\ k^2 P(k) j_0(kr). \tag{1.10}$$

Equations 1.8 and 1.10 represent a fundamental connection between the power spectrum and 2-point correlation function of a homogenous and isotropic three-dimensional random field.

1.1.2 Random fields on the sphere

The primary CMB anisotropies, and many other cosmological observables, can be accurately modelled as statistically-isotropic random fields defined on the sphere S^2. A complete basis for (square-integrable) functions f on the sphere is given by the spherical harmonics $\{Y_{lm}\}$:

$$f(\mathbf{n}) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} f_{lm} Y_{lm}(\mathbf{n}), \tag{1.11}$$

where $\mathbf{n} = (\theta, \varphi)$ specifies angular position on the sky. Using the orthogonality relations of the spherical harmonics Y_{lm}, the spherical-harmonic coefficients can be written

$$f_{lm} = \int d^2\mathbf{n} f(\mathbf{n}) Y^*_{lm}(\mathbf{n}). \tag{1.12}$$

Statistical isotropy of f ensures that its 2-point correlation function in spherical-harmonic space is diagonal, i.e., the harmonic coefficients are orthogonal (e.g., Baldi & Marinucci, 2006):

$$\langle f_{lm} f^*_{l'm'} \rangle = C_l \delta_{ll'} \delta_{mm'}, \tag{1.13}$$

where we call C_l the *angular power spectrum*. Using Eq. 1.11 and Eq. 1.13, we can write the angular (2-point) correlation function as

$$\xi(\theta) = \langle f(\mathbf{n}) f(\mathbf{n}') \rangle = \sum_{l=0}^{\infty} \frac{2l+1}{4\pi} C_l P_l(\cos\theta), \tag{1.14}$$

where $\theta = \cos^{-1}(\mathbf{n}\cdot\mathbf{n}')$ is the angular separation on the sky and $\{P_l\}$ are the Legendre polynomials. In particular, the *variance* of f is

$$\begin{aligned} \mathrm{Var}_{\mathbb{P}[f]}[f] &\equiv \langle f(\mathbf{n}) f(\mathbf{n}) \rangle - \langle f(\mathbf{n}) \rangle^2 \\ &= \xi(0) - \bar{f}^2 \\ &= \sum_{l=1}^{\infty} \frac{2l+1}{4\pi} C_l, \end{aligned} \tag{1.15}$$

where, without loss of generality, we identify $C_0 \equiv 4\pi\bar{f}^2$, and $\bar{f} = \langle f(\mathbf{n}) \rangle$ is the mean value of the field. We can invert Eq. 1.14, using orthogonality of the Legendre polynomials, to write the angular power spectrum in term of the angular correlation function:

$$C_l = 2\pi \int_0^{\pi} d\theta \sin(\theta) \xi(\theta) P_l(\cos\theta). \tag{1.16}$$

We show the cosmic microwave background (CMB) temperature correlation function in Fig. 1.1, and the corresponding power spectrum in Fig. 1.2. The expressions 1.14 and 1.16 show that, for a statistically isotropic field on the sphere, the angular power spectrum and correlation function

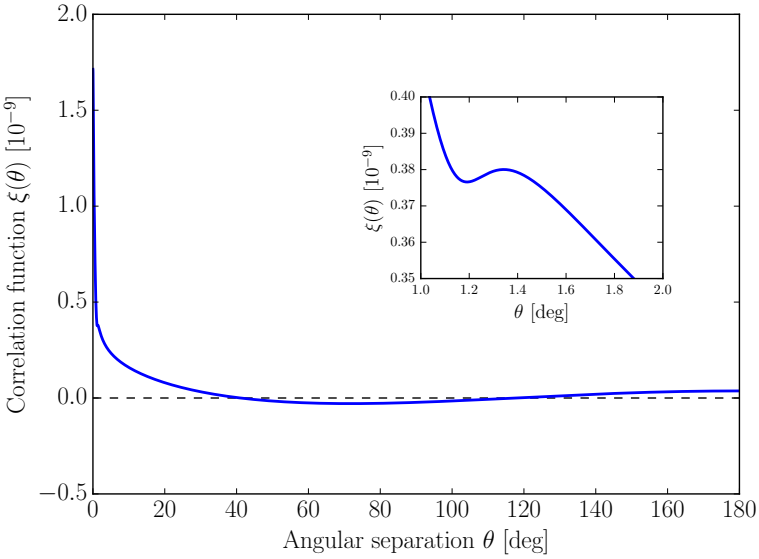

Figure 1.1: Lensed CMB temperature correlation function in the fiducial ΛCDM cosmology. We highlight (inset panel) the baryon acoustic oscillation (BAO) feature around $\theta = 1.4$ deg.

contain the same information.

1.1.3 Projection onto the sphere

Since projection is a linear operation, a homogeneous and isotropic three-dimensional field projects to an isotropic field on the sphere. Consider a field f_i on the sphere that is the projection of a homogeneous and isotropic 3D field F_i (i is a label for the field of interest):

$$f_i(\mathbf{n}) = \int d\chi \; W^i(\chi) F_i(\chi, \mathbf{n}), \qquad \text{(no s.c.)} \tag{1.17}$$

where χ is the radial comoving distance and $\mathbf{n}$ the angular position, for some *projection kernel* $W^i(\chi)$, also called the *window function*. Then, using the spherical-harmonic decomposition (Eq. 1.12), homogeneity and isotropy, the angular cross-spectrum between two fields i, j is

$$C_\ell^{ij} = \frac{2}{\pi} \int dk d\chi d\chi' \; k^2 P_{ij}(k; \eta_0 - \chi, \eta_0 - \chi') W^i(\chi) W^j(\chi) j_\ell(k\chi) j_\ell(k\chi'), \tag{1.18}$$

where j_ℓ is the spherical Bessel function of the first kind (e.g., LoVerde & Afshordi, 2008). We denote by $P_{ij}(k; \eta_1, \eta_2)$ the cross-power spectrum between fields F_i and F_j evaluated at conformal time η_1, η_2 respectively.

We now specialise to the case, as is common in cosmology, where projected fields f_i are simply

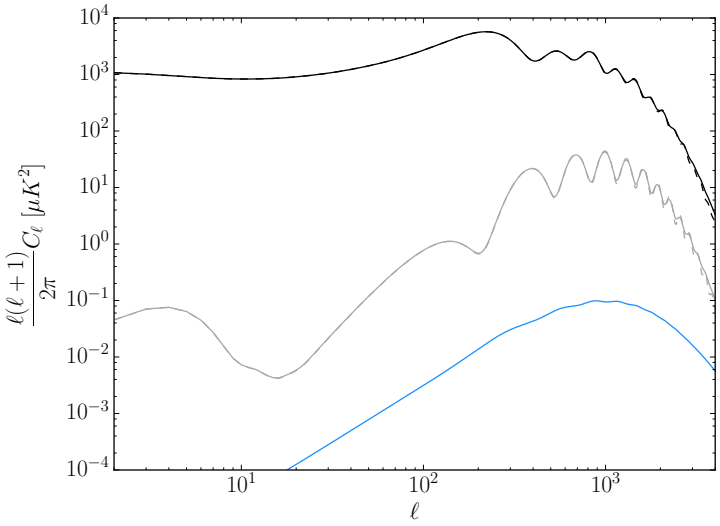

Figure 1.2: Lensed CMB temperature (*black*), E-mode (*grey*) and B-mode (*blue*) auto-spectra in the fiducial cosmology (*solid curves*). The corresponding unlensed spectra for T and E are shown (*dashed curves*); there is no primary B-mode contribution in the fiducial ($r = 0$) cosmology. At the map level, the temperature signal is a factor ~ 10 larger than the E-mode signal, which is itself $\sim 10\times$ larger than the B-mode signal. The *reionisation bump* around $\ell \approx 5$ in EE is clearly visible, and observational constraints of this are relevant for the detection of neutrino mass in cosmology (Sec. 1.3.3; Chapter 4). We discuss the physical origin of the structure in the CMB anisotropies at length in Sec. 1.3.

line-of-sight integrals over the matter overdensity δ, i.e. $F_i(\chi, \mathbf{n}) = \delta(\chi, \mathbf{n})$. Within linear theory, the modes of δ grow in time according to the *growth function* $D(\eta)$, so that $\delta(\mathbf{k}, \eta) = D(\eta)\delta(\mathbf{k})$ where $\delta(\mathbf{k})$ is the linear matter overdensity *today*, and here we normalise such that $D(\eta = \eta_0) = 1$. Then $P_{\delta\delta}(k; \eta_1, \eta_2) = D(\eta_1)D(\eta_2)P(k)$, where $P(k)$ is the linear matter power spectrum today, and Eq. 1.18 reduces to

$$C_\ell^{ij} = \frac{2}{\pi}\int dk\ k^2 P(k)\bar{W}_\ell^i(k)\bar{W}_\ell^j(k), \tag{1.19}$$

where $\bar{W}_\ell^i(k) = \int d\chi D(\eta_0 - \chi)W^i(\chi)j_\ell(k\chi)$. For sufficiently small angular scales and slowly-varying $P(k)$, the *Limber approximation* holds, and Eq. 1.19 simplifies to

$$C_\ell^{ij} = \int d\chi \frac{W^i(\chi)W^j(\chi)}{f_k^2(\chi)} P\left(\frac{\ell}{f_k(\chi)}, \eta_0 - \chi\right), \tag{1.20}$$

which is valid in curved universes (Limber, 1953; Bartelmann & Schneider, 2001). Here we have reabsorbed the (identical) growth functions $D(\eta_0 - \chi)$ back into the linear matter power spectrum. Heuristically, the Limber approximation corresponds to a physical scale k at a distance $f_k(\chi)$ being

picked out by the observed angular scale ℓ.

1.1.4 Gaussian random fields

A random field f is said to be *Gaussian*, or a *Gaussian process*, if for any finite vector of positions $[\mathbf{x}_1 \ ... \ \mathbf{x}_n]^T$ the distribution of $\mathbf{f} = [f(\mathbf{x}_1) \ ... \ f(\mathbf{x}_n)]^T$ is a multivariate Gaussian:

$$\mathbb{P}[\mathbf{f}] = \frac{1}{\sqrt{\det(2\pi\mathbb{C})}} \exp\left[-\frac{1}{2}(\mathbf{f}-\boldsymbol{\mu})^T\mathbb{C}^{-1}(\mathbf{f}-\boldsymbol{\mu})\right], \tag{1.21}$$

where the vector $\boldsymbol{\mu}$ has elements $\mu_i = \mu(\mathbf{x}_i)$ and the matrix $\mathbb{C}$ has elements $C_{ij} = C(\mathbf{x}_i, \mathbf{x}_j)$. Here $\mu = \langle f \rangle$ is the *mean function* and $C = \langle (f-\mu)(f-\mu)^T \rangle$ is the *covariance function* of the field. By Wick's theorem, all moments of the field f can be expressed in terms of the mean and covariance, and hence these functions contain all the information about the statistics of the field (they are *sufficient statistics*).

In cosmology we frequently deal with zero-mean, homogenous and isotropic random fields. In this case, $\mu = 0$, the covariance function is equivalent to the correlation function: $C_{ij} = \xi(|\mathbf{x}_i - \mathbf{x}_j|)$, and is a function only of the distance between two points. Furthermore, since the Fourier-transformed field is linear in the field f itself, the probability distribution for the *modes* $f(\mathbf{k})$ is also Gaussian, with zero mean, and a covariance function determined by the power spectrum $P(k)$ (Eq. 1.9).

1.2 Cosmological structure formation

In this section we first give a qualitative outline of the current understanding of the evolution of the universe to the present day. We follow this with a description of the Standard Model of cosmology, ΛCDM, and a more mathematical treatment of inflation and the growth of perturbations in General Relativity.

1.2.1 Qualitative timeline of the universe

It is well-understood that our universe at early times was in a dense, hot, expanding phase, in which the contents of the universe were approximately homogeneously distributed; this is the so-called *Hot Big Bang* (Peebles, 1980). There were small ripples in the density and velocity of

matter and radiation, which seeded the growth of cosmological structure. One mechanism for the production of the primordial inhomogeneities is an *inflationary* epoch in the very early universe, driven by a scalar field (or fields) which fluctuate quantum-mechanically. This is currently the most popular scenario among theorists, but observationally there remain large uncertainties about the behaviour of the universe before the Hot Big Bang phase.

Within inflation, quantum fluctuations of the scalar field were expanded to scales well-beyond the cosmological horizon at that time. Eventually inflation ended, the scalar field ϕ decayed (in a process called *reheating*), and the products included baryonic matter, dark matter and radiation. Fluctuations in ϕ were thereby imprinted as initial fluctuations in the gravitational potential and the matter and radiation fields.

Because an inflationary period is able to give rise to *super-horizon* perturbations, inflation can explain why, for example, the temperature of the CMB is approximately constant across the sky, solving the so-called *horizon problem* (Sec. 1.3). Without an inflationary period, parts of the sky separated by more than ≈ 1 deg would never have been in causal contact, and therefore could never have achieved thermal equilibrium. Inflation expands a tiny patch of the early universe, small enough to equilibrate to a given temperature, to a size larger than our current cosmological horizon. This same process dilutes any spatial curvature of in the universe, providing an explanation to the *flatness problem* – that the universe is consistent with a spatially flat geometry – which otherwise requires fine-tuning of the initial curvature to solve.

After reheating, the background expansion of the universe was set by the dominant radiation component; the universe entered the *radiation era*. The active neutrino species contributed as radiation to the background expansion at these early times. As the universe expanded and cooled, nuclei of light elements (mostly helium) formed in the dense, high-temperature universe, but these reactions effectively ceased (*froze out*) when the reaction rate became smaller than the expansion rate, and the light-element abundances were fixed.

The energy density of radiation diminishes with the expansion more quickly than matter, and at $z = z_{\rm eq} \approx 3400$ they contributed equally to the total energy density. After this time, the universe entered the *matter era*. From thereon the evolution of the background expansion was set by the matter density and the gravitational potentials were sourced by perturbations in the matter field.

Around redshift $z \approx 1100$, the Compton scattering of photons by free electrons froze out in a process called *decoupling*. The number density of free electrons rapidly decreased as electrons combined with protons to form neutral hydrogen (*recombination*), and the number of high-energy ionising photons was diluted by the expansion. Due to decoupling, the *last-scattering surface* was formed, which is the apparent surface on which photons were last scattered by free electrons. This surface, in reality, has thickness; decoupling is not instantaneous. The radial distribution of the last-scattering surface is quantified by the *visibility function*.

Prior to recombination, the Compton scattering of photons by free electrons, and the electrostatic interaction of electrons and protons, meant that radiation and baryonic matter were held together in a tightly-coupled fluid – the *photo-baryonic fluid*. Perturbations in the baryonic matter and radiation were prevented from growing due to radiation pressure; the balance of pressure and gravity lead to acoustic oscillations in the photo-baryonic fluid. These are visible today as spatial variations in the temperature and polarisation of the CMB and in the distribution of galaxies (in this context, known as the baryonic acoustic oscillations, or BAO). Neutrinos, although acting as radiation at this time, did not interact through the electromagnetic force. After they decoupled as relativistic particles when the universe had a temperature of around $k_B T \approx 1$ MeV, they free-streamed out of overdensities smaller than the horizon size and therefore do not contribute to late-time clustering on these scales.

While the photo-baryonic fluid was oscillating, sub-horizon dark matter perturbations – decoupled from radiation – continued to grow, first logarithmically in the scale factor a during the radiation era, then linearly in a during matter domination. This underlying growth of the dark matter perturbations is essential to explaining the extent of structure formation observed in the universe today. After decoupling, the baryons (free to collapse gravitationally) quickly fell into the potentials set up by the dark matter perturbations, while photons free-streamed, permeating the universe with what is now detected as the primary CMB.

Well after decoupling, the matter perturbations grew linearly with the scale factor a. At the peaks in the density field baryonic matter was sufficiently dense to collapse to form stars and galaxies. These same matter perturbations, which seed the growth of virialised structures, also gravitationally deflect photons from the last-scattering surface and distant galaxies, inducing *gravitational lensing*. Small-scale structures in the late-time universe grew non-linearly, introducing

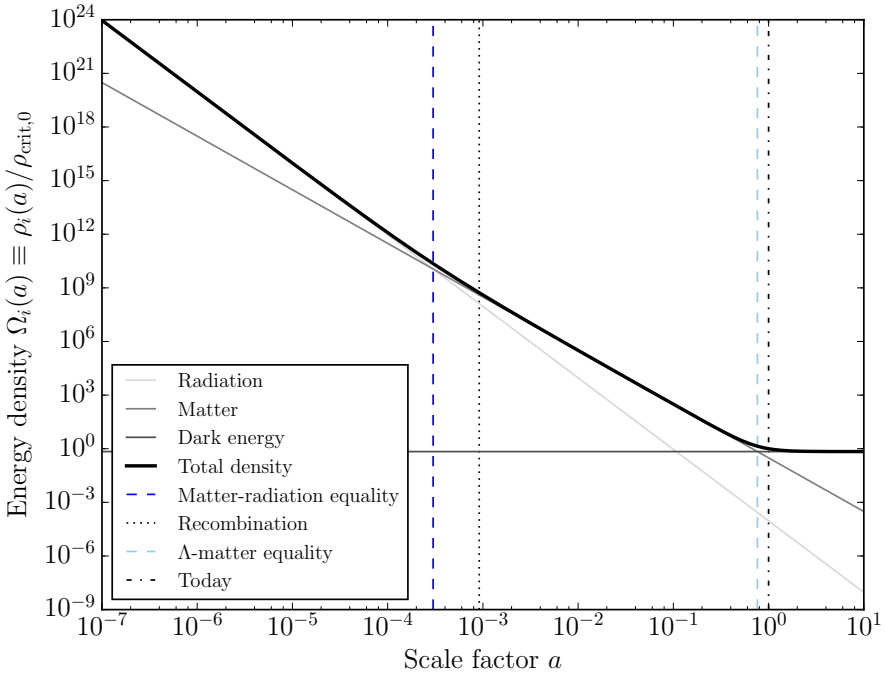

Figure 1.3: Evolution of the energy density for different components in a ΛCDM universe. The total density $\sim a^{-4}$ during radiation domination, $\sim a^{-3}$ during matter domination and $\sim$ constant during Λ domination.

an additional scale dependence in the growth. At late times, around $z \approx 0.7$, the energy density of the universe became dominated by dark energy which caused an accelerating expansion. The universe to transitioned into the *dark energy era*.

1.2.2 ΛCDM

The Λ cold dark matter model (ΛCDM), or *concordance* cosmological model, is the standard theoretical framework in which much of cosmological and astrophysical research is conducted. The model asserts that space-time – on cosmological scales – can be described by a perturbed FRW metric with dynamics governed by Einstein's theory of General Relativity (GR). Within ΛCDM, the universe is spatially flat, expanding, and filled with radiation, neutrinos, cold dark matter (CDM), baryonic matter and a cosmological constant Λ. The initial conditions for the matter and gravitational potential fields in ΛCDM are adiabatic and Gaussian, with a power-law primordial power spectrum. We note that although these initial conditions are consistent with those from slow-roll inflation, they are taken as an ansatz of the model, independent of their

origin.

This model is able to explain many of the observed properties of the universe: for instance, the late-time accelerated expansion (through Λ); the presence of the cosmic microwave background (via expansion, and the corresponding cooling of the universe); the CMB anisotropies (acoustic peaks are set up by Compton scattering of radiation by free electrons in the primordial plasma); and the properties of large-scale structure (the Einstein equations applied to a perturbed FRW metric specify uniquely the growth of perturbations in a ΛCDM universe).

Alternatives to ΛCDM generally come in three flavours: firstly, those in which GR is replaced by a modified theory of gravity (e.g., $f(R)$ gravity); secondly, those in which the *contents* of the universe are different, such as the inclusion of dynamical dark energy ($\Lambda = \Lambda(a)$) or additional scalar fields; thirdly, alternative initial conditions, such as primordial isocurvature perturbations or a running scalar spectral index. There has been no conclusive evidence to suggest ΛCDM is fundamentally unable to explain cosmological observations to date. Ongoing observations, such as measurements of the dark energy equation of state with CMB and large-scale structure experiments, will continue to test the explanatory power of ΛCDM, and theorists will continue to propose alternative models against which ΛCDM can be compared.

In Fig. 1.3 we show the evolution of matter, radiation and dark energy densities in a ΛCDM universe. At early times, radiation dominates the total energy density until matter-radiation equality. The universe is then matter dominated until around 10 billion years after the big bang. At late times, the cosmological constant comes to dominate the total energy density, and will ultimately lead to exponential expansion.

1.2.3 Inflation

As discussed above, inflation is motivated by its ability to solve the horizon problem, the flatness problem and in providing a mechanism for laying down the initial Gaussian fluctuations. GR is described by the Einstein-Hilbert action:

$$S = \int d^4x\sqrt{-g}\left[-\frac{R}{16\pi G} + \mathcal{L}\right], \tag{1.22}$$

where we include the Lagrangian for matter fields $\mathcal{L}$ (e.g., the inflaton). Variation of the action with respect to the metric leads to the Einstein equations:

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = 8\pi G T_{\mu\nu}, \tag{1.23}$$

which link the Ricci tensor $R_{\mu\nu}$ (quantifying space-time curvature) to the stress-energy tensor of the matter $T_{\mu\nu}$ (quantifying the energy-momentum of the cosmological fluid). We can write the Lagrangian for a minimally-coupled single scalar field ϕ as

$$\mathcal{L} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - V(\phi). \tag{1.24}$$

For convenience, we can decompose the scalar field into a homogenous term $\bar{\phi}$ and a perturbation term $\delta\phi$:

$$\phi(\eta, \mathbf{x}) = \bar{\phi}(\eta) + \delta\phi(\eta, \mathbf{x}). \tag{1.25}$$

This partitioning is useful because perturbations about the background are small (as required by observations such as the variance of the CMB temperature). At zeroth order in the perturbations, the Einstein equations give

$$\ddot{\phi} + 2\mathcal{H}\dot{\phi} + a^2 V_{,\phi} = 0 \tag{1.26}$$

for the background evolution of the inflaton field, and a Friedmann equation

$$\mathcal{H}^2 = \frac{8\pi G}{3}\left[\frac{1}{2}\dot{\phi}^2 + a^2 V(\phi)\right] \tag{1.27}$$

determining the expansion rate. For a given potential these coupled equations can be solved (generally numerically) for the background field value $\bar{\phi}$ and scale factor a as a function of time. Here $\mathcal{H} \equiv d\ln a/d\eta = aH$ is the conformal Hubble rate, and a comma denotes differentiation with respect to the proceeding quantity (in this case, ϕ).

In the *slow-roll approximation*, the kinetic energy density of the field is assumed to be much smaller than its potential energy, viz., $|\frac{1}{2}\dot{\phi}^2| \ll a^2|V(\phi)|$, and the field acceleration is also taken to be small: $|\ddot{\phi}| \ll |2\mathcal{H}\dot{\phi} + a^2 V_{,\phi}|$. In this regime, the Hubble rate H is approximately constant, leading to rapid exponential expansion $a \sim \exp(Ht)$, and thereby a solution to the horizon and

flatness problems (Sec. 1.3).

The inflaton field will also fluctuate quantum mechanically about the mean field (the $\delta\phi(\eta, \mathbf{x})$ term of Eq. 1.25). Through a canonical quantisation of the classical field equation of the inflaton fluctuations, one can compute the power spectrum of the super-horizon inflaton fluctuations. This is achieved by expanding the quantised field in creation and annihilation operators and applying the canonical commutation relations, giving

$$P_{\delta\phi}(k) = \left(\frac{H}{2\pi}\right)\Bigg|_{aH=k}, \tag{1.28}$$

where the Hubble rate is evaluated at horizon exit for the particular comoving mode k (Baumann, 2009). Since H remains approximately constant during inflation, the power spectrum of inflaton perturbations is approximately scale-invariant. The fluctuations in the inflaton field can be related to fluctuations in the (gauge-invariant) comoving curvature perturbation $\mathcal{R}$, which remains constant outside the horizon (Weinberg, 1972; Baumann, 2009). Therefore, within single-field slow-roll inflation, after a mode re-enters the horizon during radiation domination, the primordial curvature power spectrum is the same as at horizon exit, and can be written

$$P_{\mathcal{R}}(k) = A_s \left(\frac{k}{k_0}\right)^{n_s - 1}, \tag{1.29}$$

for some pivot scale k_0, and scalar spectral index n_s which quantifies the deviation from scale-invariance. Furthermore, the fluctuations in single-field inflation are Gaussian, fully described by their power spectrum. The deviation from scale-invariance is required within slow-roll inflation to bring the inflationary period to an end. The theoretical prediction for n_s is model dependent, in that it depends on the details of the particular inflationary potential $V(\phi)$ considered (e.g., Martin et al., 2014).

After inflation ends – in some models, this is due to the breakdown of the slow-roll approximation – *reheating* follows. The details of reheating are observationally unconstrained, but it could the thought of as a period in which ϕ oscillates about a potential minimum, slowing the expansion and allowing the inflaton field to couple with a matter field (or fields). This would transfer energy into products which eventually decay into the known species (quarks, electrons, neutrinos, photons and so forth). Fourier modes in which all species have equal energy density on

hypersurfaces of constant total energy density, and equal peculiar velocities, are called *adiabatic* modes. Although the details of reheating are highly uncertain, they are unimportant, in the sense that they do not affect the super-horizon curvature perturbations, which remain constant until re-entry.

The consequence of an inflationary epoch is to seed small, Gaussian fluctuations in a universe energetically dominated by radiation, with additional dark and baryonic matter components. The end of reheating signals the beginning of the *Hot Big Bang* phase. Primordial fluctuations grow under gravity to create the variety of structure seen today in the late-time universe. The growth of structure is the next topic in this chapter.

1.2.4 Structure formation

The distribution and statistics of the large-scale structure are determined by the initial conditions and the physics of the evolution of perturbations in an expanding universe. The initial comoving curvature perturbation can be taken to be a Gaussian random field (Sec. 1.1.4), realised from the power spectrum in Eq. 1.29, with the ratio of the density of each matter component equal everywhere in the comoving gauge (by adiabacity). Within GR, the evolution from the initial conditions is determined uniquely by the Einstein equations for the cosmological fluids in a perturbed FRW metric. Since the fluctuations in the matter and gravitational potentials are small, linear perturbation theory can be used to make this calculation tractable, accurately modelling all but the late-time, non-linear structure.

The perturbed FRW metric can be written, in the conformal Newtonian gauge, as

$$ds^2 = a^2(\eta)\left[(1+2\psi)d\eta^2 - (1-2\phi)\delta_{ij}dx^i dx^j\right] \tag{1.30}$$

(e.g., Ma & Bertschinger, 1995). The Einstein tensor $G_{\mu\nu}$ for this perturbed metric can be calculated (to first-order in the gravitational potentials ϕ, ψ) to give the left hand side of Einstein's field equations (Eq. 1.23). The right hand side of the field equations is proportional to the perturbed energy-momentum tensor $T^{\mu\nu}$. The field equations can be matched, order-by-order in the perturbations, to obtain the background evolution and the dynamics of the perturbations. Energy-momentum conservation gives GR versions of the continuity and Euler equations for the

density and velocity perturbations of each non-interacting fluid, while the remaining field equations determine how the potentials evolve in time (as sourced by the matter perturbations). The curvature perturbation ϕ is related to the total density perturbation in the comoving gauge Δ (corresponding to an observer locally comoving with the fluid) by the Poisson equation:

$$\nabla^2\phi = 4\pi G a^2 \rho \Delta, \tag{1.31}$$

where ρ is the mean density. Furthermore, in a universe with no anisotropic stress – as is assumed from hereon – the potentials are equal ($\phi = \psi$), and evolve according to the following second-order differential equation:

$$\ddot{\phi} + 3\mathcal{H}(1+w)\dot{\phi} + \left[2\dot{\mathcal{H}} + \mathcal{H}^2(1+3w)\right]\phi = 4\pi G a^2 w \rho \Delta, \tag{1.32}$$

where derivatives are with respect to conformal time and w is the (assumed constant) equation of state of the cosmological fluid (i.e., such that pressure $p = w\rho$). These two equations (Eq. 1.31–1.32) form a closed system. For a general mixture of fluids the equation of state is complicated, and time varying, but within a given energy regime (e.g. matter or radiation domination) Eq. 1.32 holds, and this system can be solved analytically in certain limits.

Radiation, baryon and CDM perturbations

In this section we discuss various limits of Eq. 1.31–1.32, identifying the physical effects at play. Many of the concepts discussed below are illustrated in Fig. 1.4, which displays the evolution of perturbations of the photons, baryons and CDM for two different Fourier modes. For later convenience, we define here *matter-radiation equality* as the time at which the energy density of matter and radiation are equal. This occurs when the universe has a scale factor $a_{\rm eq} = \Omega_{\rm r}/\Omega_{\rm m}$, or a redshift $z_{\rm eq}$ given by

$$1 + z_{\rm eq} = 2.4 \times 10^4 (\Omega_{\rm m} h^2) = 3386, \tag{1.33}$$

where the last equality holds for the fiducial model, and we have assumed photons and three neutrino species contributing to the radiation density $\Omega_{\rm r}$.

During the radiation era, the gravitational potential ϕ is set initially by the dominant radiation

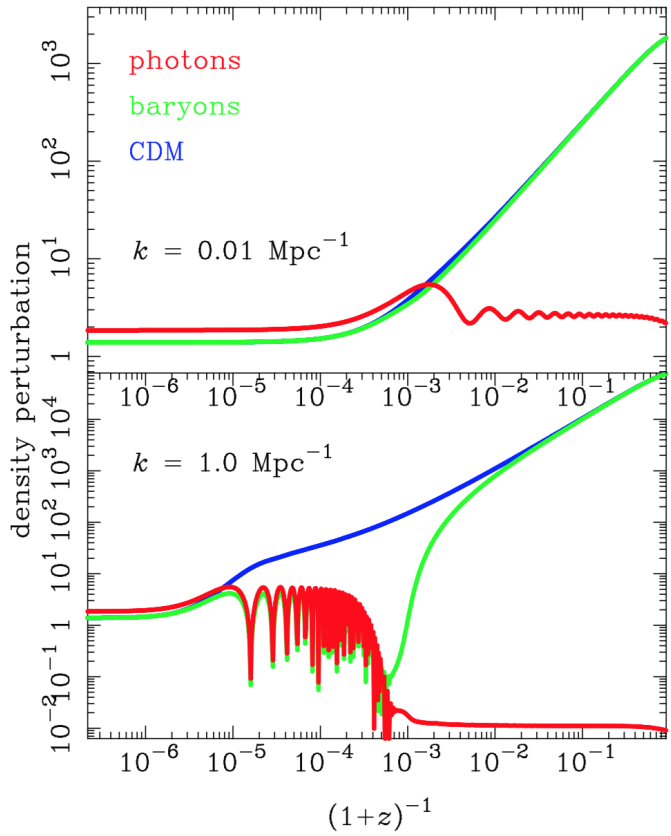

Figure 1.4: *Source: Fig. 4 of Challinor (2009).* Density perturbations in the conformal Newtonian gauge for two different modes. Note the different horizontal scale than used in Fig. 1.5. The smaller-scale mode, in the bottom panel, enters the horizon earlier: sub-horizon CDM perturbations grow logarithmically in the scale factor during radiation domination (the *Mézáros* effect; e.g., Mészáros, 1974; Dodelson, 2003), then linearly during matter domination. After decoupling the baryons fall into the potentials set up by the CDM growth (green curve). Radiation modes oscillate on entering the horizon under the competing influence of gravity and radiation pressure; these oscillations are driven during radiation domination (Hu & Sugiyama, 1995), and are damped by the expansion during matter domination when the potential ϕ is constant (visible in top panel). Additionally, on small scales, photons are able to *diffuse* according to their mean free path; this exponentially suppresses power in high-k modes (*bottom panel*), and is called *Silk damping* (Silk, 1968). Baryon modes which become sub-horizon before decoupling oscillate with the photons due to tight coupling from Compton scattering. Dark energy decays the potentials at late-times, slowing the growth rate; this is visible in both panels from $a \gtrsim 0.5$ and more clearly in Fig. 1.5.

perturbations ($\rho\Delta = \rho_r\Delta_r$ in Eq. 1.31). In this regime, ϕ is constant outside the horizon and undergoes oscillations inside the horizon, with an amplitude which diminishes with time (Challinor, 2009). The radiation perturbations themselves are driven by the decaying potential (i.e. Δ_r increases), then, well inside the horizon, they oscillate with roughly constant amplitude (Hu & Sugiyama, 1995). Baryons are tightly coupled to the photons through Compton scattering, and hence they also oscillate on sub-horizon scales: these are the acoustic oscillations visible in the CMB and late-time galaxy distribution. Diffusion of photons, determined by their mean free path through the photo-baryonic fluid, causes the amplitude of small-scale photo-baryonic perturbations to be exponentially suppressed with scale k; this is called diffusion damping or *Silk damping* (Sec. 1.3.2; Hu & Sugiyama, 1995; Silk, 1968).

During this epoch, while the potential is dominated by radiation perturbations, super-horizon CDM perturbations grow like $\sim a^2$ in the comoving gauge, and sub-horizon CDM perturbations grow only *logarithmically* in the scale factor (Fig. 1.6; Dodelson, 2003; Baumann, 2015). Eventually ϕ decays until the CDM perturbations dominate and begin to determine the potential ($\rho\Delta = \rho_c\Delta_c$ in Eq. 1.31). This occurs even before matter-radiation equality: although $\rho_m < \rho_r$, we still have $\rho_m\Delta_m > \rho_r\Delta_r$, since the radiation perturbations oscillate and cease to grow on sub-horizon scales. In this regime, in which CDM perturbations determine the potential, one can show – by combining Eq. 1.31 and Eq. 1.32 with $w = 0$ for a universe with just matter and radiation – that the growing mode CDM perturbations evolve like

$$\Delta_c \sim 1 + \frac{3}{2}\left(\frac{a}{a_{\mathrm{eq}}}\right). \tag{1.34}$$

That is, the alteration of the expansion rate by radiation *suppresses* sub-horizon CDM growth until after matter-radiation equality, when it grows linearly (Dodelson, 2003). The suppression of growth of CDM perturbations by radiation, as in Eq. 1.34, is called the *Mészáros effect* (Mészáros, 1974). This solution (plus the corresponding decaying mode solution) must be matched onto the logarithmic growth deep in the radiation era when radiation perturbations determine the potential. We now discuss the growth of matter perturbations after decoupling, when curvature and dark energy can become important.

Structure growth after decoupling

After decoupling, the baryons quickly fall into the potentials set up by the dominant dark matter component which have been growing, uninhibited, while pressure support in the photo-baryonic fluid stalled the growth of the baryon perturbations (Fig. 1.4; Challinor, 2009; Baumann, 2015). CDM and baryon perturbations then follow the same distribution, and we can define the *total* matter density contrast Δ_m (in the comoving gauge). Setting $w = 0$ (for matter) and substituting Eq. 1.31 into Eq. 1.32, this simplifies to

$$\ddot{\Delta}_m + \mathcal{H}\dot{\Delta}_m - \frac{3\Omega_{\rm m}H_0^2}{2a}\Delta_m = 0. \tag{1.35}$$

From the Friedmann equation, the Hubble rate $H \equiv d\ln a/dt$ is given by

$$H^2(a) = H_0^2\left(\frac{\Omega_{\rm r}}{a^4} + \frac{\Omega_{\rm m}}{a^3} + \frac{\Omega_k}{a^2} + \Omega_\Lambda(a)\right), \tag{1.36}$$

for a universe with radiation, matter, curvature and dark energy, where $\{\Omega_i\}$ are the density parameters today ($a = 1$) for each component. In general the background energy density of a non-interacting fluid i is given by

$$\rho_i(z) = \Omega_i\rho_{\rm crit}\exp\left(3\int_0^z \frac{1+w(z')}{1+z'}dz'\right) \tag{1.37}$$

where $\rho_{\rm crit} = 3H_0^2/8\pi G = 1.878\times10^{-26}h^2$ kg m^{-3} is the *critical density* today. For this reason, the combination $\Omega_i h^2$ is the *physical* energy density parameter. Using $\mathcal{H} = aH$ and Eq. 1.36, we can rewrite Eq. 1.35 as

$$\Delta_m'' + \frac{a}{2}\left[\frac{\Omega_{\rm m} + 2\Omega_k a + \Omega_\Lambda(a)a^3[1-3w(a)]}{\Omega_{\rm r} + \Omega_{\rm m}a + \Omega_k a^2 + \Omega_\Lambda(a)a^4}\right]\Delta_m' - \frac{a}{2}\left[\frac{3\Omega_{\rm m}}{\Omega_{\rm r} + \Omega_{\rm m}a + \Omega_k a^2 + \Omega_\Lambda(a)a^4}\right]\Delta_m = 0, \tag{1.38}$$

where now derivatives $'$ are with respect to the logarithmic scale factor ($\ln a$). We denote by $w(a)$ the equation of state of the dark energy component. This form is particularly amenable for implementing a numerical integration scheme, since the pre-factor functions in square brackets are well-behaved as $a \to 0$. For the special case of no radiation ($\Omega_{\rm r} = 0$) and a cosmological constant

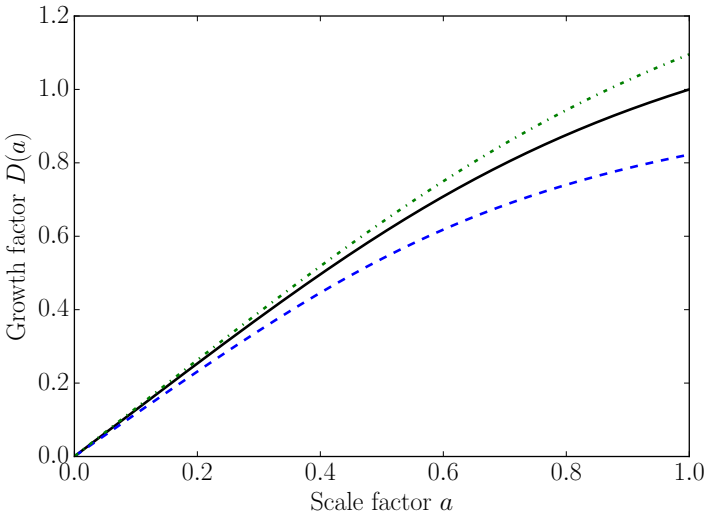

Figure 1.5: Matter growth factor in the fiducial cosmology ($\Omega_\mathrm{m} \approx 0.3$), normalised to unity today (*black solid*). We also show the growth factor for an $\Omega_\mathrm{m} = 0.2$ (*blue dashed*) and $\Omega_\mathrm{m} = 0.4$ (*green dot-dashed*) cosmology. At early times, during matter-domination, $D(a) \sim a$. Structure grows more readily in a universe with a higher matter density (note the relative amplitude of each curve). Suppression of growth by the cosmological constant explains the flattening at late times; this effect becomes more marked as Ω_m decreases (so that Ω_Λ increases).

($w = -1$), there is a closed-form integral solution for the growing-mode amplitude:

$$D(a) = \frac{5}{2}\Omega_\mathrm{m} H_0^2 H(a) \int_0^a \frac{da'}{a'^3 H^3(a')}, \tag{1.39}$$

where $\Delta_m(\mathbf{x}, a) = \frac{D(a)}{D(a_i)}\Delta_m(\mathbf{x}, a_i)$ for some initial amplitude $\Delta_m(\mathbf{x}, a_i)$ (Heath, 1977; Hamilton, 2001). In Eq. 1.39, the growth factor D is normalised such that $D(a) = a$ at early times. We show the growth factor for various cosmologies in Fig. 1.5, computed by numerical integration of Eq. 1.38.

1.2.5 The matter power spectrum and non-linear growth

Throughout linear evolution of the initial Gaussian perturbations, the matter power spectrum fully describes the statistics of the matter field (Sec. 1.1).

The primordial power spectrum of the comoving curvature perturbation is almost scale invariant within ΛCDM (Eq. 1.29). This can be related to fluctuations in the gravitational potential ϕ, giving a scale dependence $P_\phi \sim k^{n_s - 4}$ for the initial power spectrum of ϕ, where n_s is the scalar

		RD		MD	
		Φ	δ_m (Δ_m)	Φ	δ_m (Δ_m)
$k \gg k_{\rm eq}$:	superhorizon	*const.*	*const.* (a^2)	–	–
	subhorizon	a^{-2}	$\ln a$	*const.*	a
$k \ll k_{\rm eq}$:	superhorizon	*const.*	*const.* (a^2)	*const.*	*const.* (a)
	subhorizon	–	–	*const.*	a

Figure 1.6: *Source: Table 5.1 of Baumann (2015).* Scale factor dependence of the potential ϕ and matter density perturbation in the Newtonian gauge $\delta_{\rm m}$ and comoving gauge $\Delta_{\rm m}$. This summarises much for the discussion of Sec. 1.2.4–1.2.5.

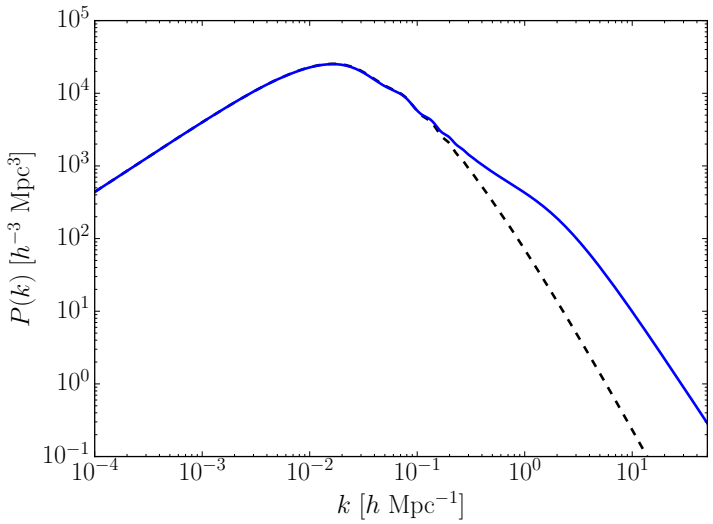

Figure 1.7: Matter power spectrum today $P(k, z = 0)$ in the fiducial cosmology for linear (*black dashed*) and non-linear (*blue solid*) evolution. The acoustic oscillations are visible to the right of the peak at $k = k_{\rm eq}$ (Eq. 1.40). At $z = 0$ the non-linear corrections are observationally important beyond $k \gtrsim 0.1\ h$ Mpc^{-1} and modelling uncertainties can attribute systematic uncertainties when performing clustering or lensing analyses (Chapter 5).

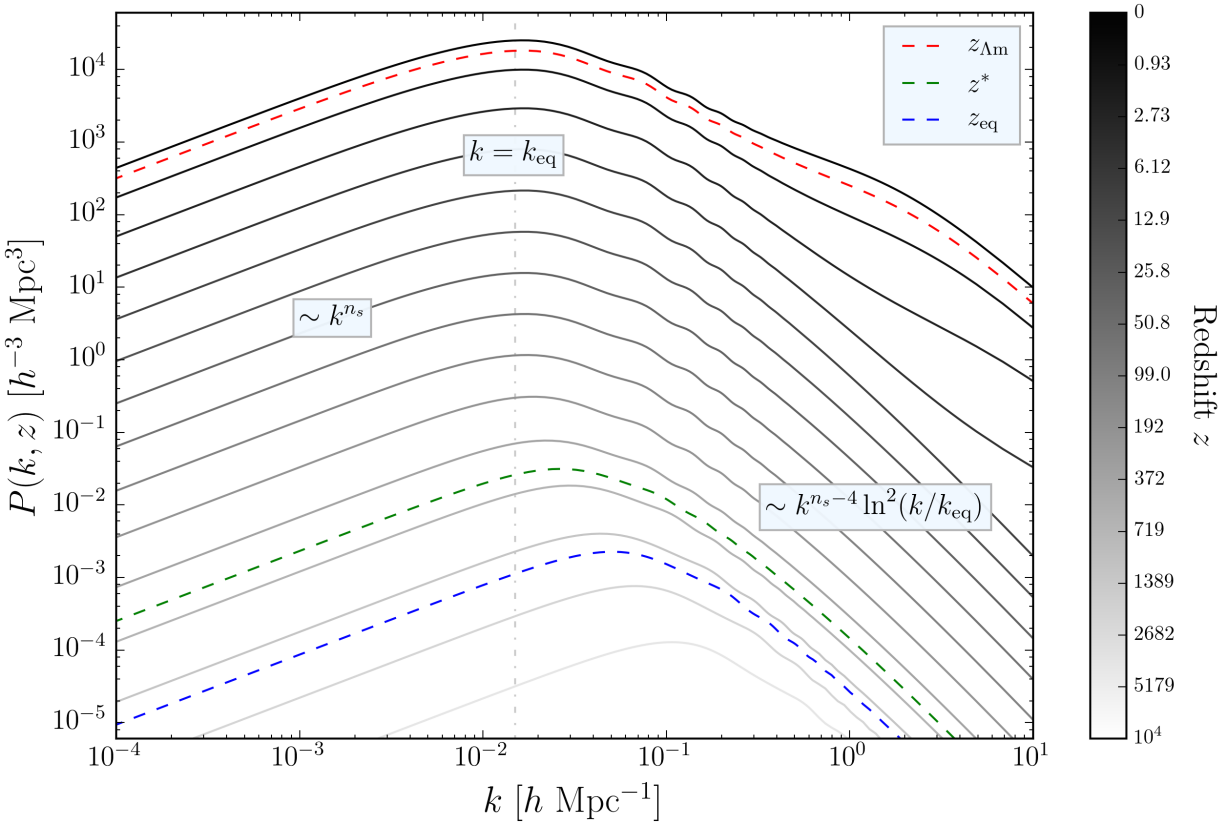

Figure 1.8: Matter power spectrum $P(k, z)$ at 15 epochs – logarithmically spaced in $1+z$ between $z = 10^4$ and today ($z = 0$) – assuming the fiducial cosmology (*the redshift of each grey/black curve correponds to a tick label in the colourbar*). At any given time, modes which are super-horizon, or super-horizon at equality, retain the primordial shape of the power spectrum $P(k) \sim k^{n_s}$ ($k \ll k_{\rm peak}$). At early times the peak scale $k_{\rm peak}$ corresponds directly to the horizon size. However, after equality, modes on all scales grow at the same rate since the universe is matter dominated (Sec. 1.2.5); the peak position then becomes fixed at $k = k_{\rm eq}$ (*dot-dashed vertical line*, Eq. 1.40). The growth of modes entering the horizon before equality is suppressed by the *Mézáros effect*, leading to a small-scale asymptotic shape $P(k) \sim k^{n_s-4} \ln^2(k/k_{\rm eq})$, for $k \gg k_{\rm peak}$. Non-linear corrections become important at late times and on small scales (*top right*). The acoustic oscillations on scales $k > k_{\rm peak}$ are frozen in (non-evolving) after decoupling ($z = z^* = 1090$, *green dashed*). For reference, we also show the matter power spectrum at matter-radiation equality ($z = z_{\rm eq} = 3386$, *blue dashed*) and matter-Λ equality ($z = z_{\Lambda m} = 0.31$, *red dashed*). On this figure, the different growth factors for each regime of radiation ($\sim a^2$), matter ($\sim a$) and dark energy ($\sim$ const.) domination are manifest as equal spacing between curves – within a given regime – at large scales. This spacing decreases monotonically from high-redshift to today.

spectral index (Dodelson, 2003). The primordial matter power spectrum in the comoving gauge P then follows from the Poisson equation in Fourier space: $\Delta_m \sim k^2\phi$, giving $P(\eta = 0) \sim k^{n_s}$.

An important scale is the *equality scale* $k_{\rm eq}$, which is the wavenumber corresponding to the horizon size at matter-radiation equality. Since the horizon size is approximately equal to the Hubble radius, we can define $k_{\rm eq} \equiv a_{\rm eq}H(a_{\rm eq})/c$, i.e. the inverse (comoving) Hubble radius[1]; then

$$k_{\rm eq} = \sqrt{\frac{2\Omega_{\rm m}H_0^2}{a_{\rm eq}c^2}} = 0.015\left(\frac{\Omega_{\rm m}h}{0.212}\right) h\ {\rm Mpc}^{-1} \tag{1.40}$$

Those modes which enter the horizon before equality will grow only logarithmically (while radiation perturbations determine the potential), whereas the super-horizon matter perturbations in the comoving gauge Δ_m grow like $\sim a^2$ (e.g., Peebles, 1980; Baumann, 2015). The growth of small-scale modes, which enter the horizon before equality, is correspondingly suppressed relative to the growth of super-horizon modes. This effect leads to a scale-dependence of the small-scale power spectrum $P(k) \sim k^{n_s-4}\ln^2(k/k_{\rm eq})$ for $k \gg k_{\rm eq}$ (Dodelson, 2003).

This result can be understood more physically as follows. The small-scale comoving mode k enters the horizon ($\approx$ Hubble radius) when $aH = k$, and during radiation domination $H \sim a^{-2}$; the mode therefore enters the horizon when the universe has a scale factor $a = a_k \sim 1/k$. Once within the horizon, the mode growth is modified relative to super-horizon modes, which are growing like $\sim a^2$ during radiation domination; growth until equality is suppressed, therefore, by a factor $(a_k/a_{\rm eq})^2 = (k_{\rm eq}/k)^2$, which in the power spectrum introduces the factor k^{-4} into the transfer function. The correction factor $\ln^2(k/k_{\rm eq})$ is due to the modes growing logarithmically (rather than not at all) on entering the horizon during radiation domination – smaller-scale modes enter the horizon earlier and hence undergo longer logarithmic growth. After equality, modes on all scales (in the comoving gauge) then simply grow like $\sim a$; there is no additional processing as modes enter the horizon and the broadband shape of the matter power spectrum shape is fixed.

Since the perturbations which remain super-horizon until equality enter the horizon during matter domination, when the potential ϕ is constant and the growth factor $\sim a$ on all scales, these modes simply grow in amplitude from $\eta = 0$, with no additional scale dependence induced.

[1]Defining instead $k_{\rm eq} \equiv \chi_{\rm eq}^{-1}$, where $\chi_{\rm eq}$ is the exact cosmological horizon size at equality, gives essentially the same number ($k_{\rm eq} = \frac{1}{2(\sqrt{2}-1)}\sqrt{\frac{\Omega_{\rm m}H_0^2}{a_{\rm eq}c^2}} = 0.062(\Omega_{\rm m}h^2){\rm Mpc}^{-1} = 0.013\ h\ {\rm Mpc}^{-1}$ in the fiducial cosmology), as pointed out in Dodelson (2003).

Summarising,

$$P(k) \sim \begin{cases} k^{n_s} & k \ll k_{\rm eq}, \\ k^{n_s-4} \ln^2(k/k_{\rm eq}) & k \gg k_{\rm eq}. \end{cases} \tag{1.41}$$

Taken together, the considerations above explain the characteristic bump in the matter power spectrum, centred on the equality scale $k_{\rm eq}$, and the asymptotic scale-dependence at small and large k. We show the matter power spectrum today in the fiducial cosmology in Fig. 1.7. The dashed black curve is the linear matter power spectrum, calculated by numerically integrating the Einstein equations for a linearly-perturbed FRW metric with CAMB (Lewis et al., 2000). The solid blue curve shows how non-linear corrections affect the small-scale power spectrum, calculated using the HALOFIT software within CAMB (Takahashi et al., 2012). Matter on small scales evolves non-linearly at late times as the density constrast Δ_m approches unity, and first-order perturbation theory becomes insufficient to model growth, which is more accurately described by the inherently non-linear field equations. This quickly becomes analytically intractable, and N-body simulations must be used to calibrate the non-linear correction across a wide-range of cosmologies. Baryonic feedback effects introduce additional uncertainty into theoretical models of the small-scale power spectrum (e.g., Mead et al., 2015, 2016).

In Fig. 1.8 we draw together many of the results of this section into a summary plot, showing the evolution of the matter power spectrum from the early universe ($z = 10^4$; well before equality), to today.

1.2.6 Background expansion with massive neutrinos

In Chapter 4 we consider in detail the effect of neutrinos on cosmological observables. Here we consider the mass-dependent effect on neutrinos on the background expansion, since it is relevant for accurate calculations of the expansion rate and all distance measures in cosmology (e.g., Eq. 1.36).

In the early universe neutrinos are in kept in equilibrium with the cosmological fluid through the weak interaction. When the universe cools to a temperature of around $k_B T = 1$ MeV, the relativistic neutrinos decouple, and the shape of the Fermi-Dirac distribution function $f(p)$ at

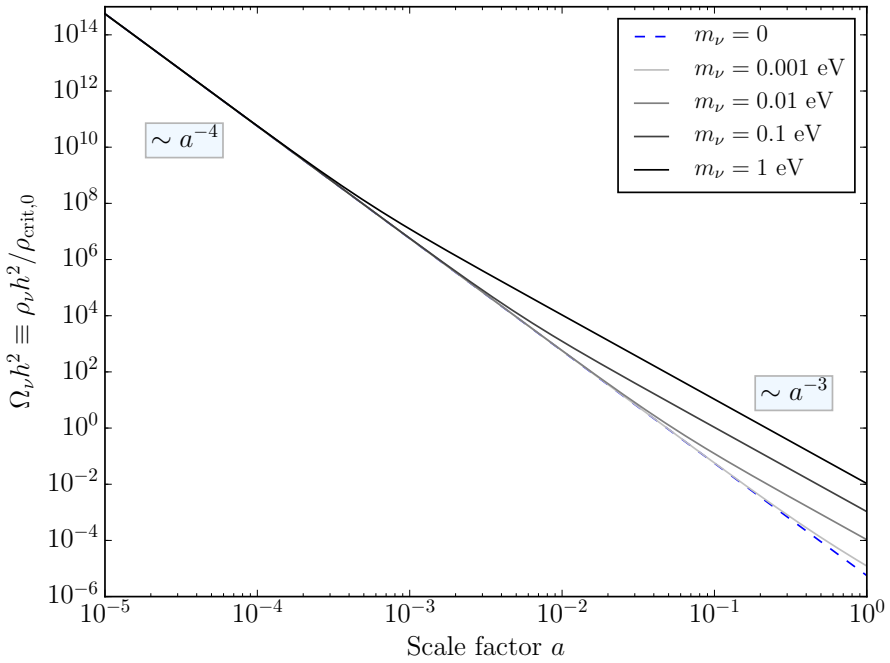

Figure 1.9: Neutrino energy density evolution, normalised to the critical density today, for various values of the neutrino mass m_ν. Note the radiation and matter regimes at early and late times, respectively. Heavier neutrinos become non-relativistic earlier and give a larger contribution to the energy density today ($a = 1$). We plot the exact form given by Eq. 1.44.

decoupling

$$f(p|T) = \frac{1}{1 + \exp\left(\frac{pc}{k_B T}\right)} \tag{1.42}$$

is preserved thereafter, since they become effectively collisionless (Lesgourgues & Pastor, 2012). The neutrino temperature evolves as $T/T_{\nu,0} = a^{-1}$, where $T_{\nu,0}$ is the neutrino temperature today, and the energy density of a single mass eigenstate m_ν when the universe has a scale factor a can then be written

$$\rho_\nu(a) = g_\nu \int_0^\infty \frac{4\pi p^2 dp}{(2\pi\hbar)^3} \frac{\sqrt{p^2c^2 + m_\nu^2 c^4}}{1 + \exp\left(\frac{apc}{k_B T_{\nu,0}}\right)}. \tag{1.43}$$

Defining a dimensionless dummy variable $x \equiv apc/(k_B T_{\nu,0})$, and using $g_\nu = 2$ for the two helicity states, this reduces to

$$\rho_\nu(a) = \frac{1}{\pi^2 c^3 \hbar^3} \left(\frac{k_B T_{\nu,0}}{a}\right)^4 \int_0^\infty dx \frac{x^2\sqrt{x^2 + \alpha^2}}{1 + \exp x}, \tag{1.44}$$

where $\alpha \equiv a/a_{\rm nr}$ and $a_{\rm nr} \equiv k_B T_{\nu,0}/(m_\nu c^2)$ defines the scale factor when the neutrino becomes non-relativistic. The integral expression must be, in general, evaluated numerically, but Eq. 1.44 reduces analytically to the expected radiation and matter regimes in the appropriate limits:

$$\rho_\nu(a) = \begin{cases} \left[\frac{7\pi^2 (k_B T_{\nu,0})^4}{120 c^3 \hbar^3}\right] a^{-4} & a \ll a_{\rm nr}, \\ \left[\frac{3\zeta(3)(k_B T_{\nu,0})^3}{2\pi^2 c^3 \hbar^3} m_\nu c^2\right] a^{-3} & a \gg a_{\rm nr}. \end{cases} \tag{1.45}$$

We note that the neutrino temperature today $T_{\nu,0}$ can be related to the photon temperature today through $T_{\nu,0} = \left(\frac{4}{11}\right)^{1/3} T_0$; this difference arises because neutrinos decouple prior to the energy injection from electron-positron annihilation. For any distance calculations which span both the relativistic and non-relativistic regimes of the neutrino density evolution, such as the angular diameter distance to recombination, it is important for computational accuracy that the full expression in Eq. 1.44 is used. We plot this quantity for various values of the neutrino mass in Fig. 1.9.

Assuming all massive neutrinos are fully non-relativistic today, with a mass sum Σm_ν, we can use Eq. 1.45 to derive the well-known relation between neutrino mass and the physical massive neutrino density parameter (defined today, at $a = 1$):

$$\Omega_\nu h^2 = \frac{\Sigma m_\nu c^2}{94.1 \text{ eV}}. \tag{1.46}$$

1.3 The cosmic microwave background

Around 400,000 years after the Big Bang (at redshift $z = z^* \approx 1100$), the universe underwent a transition from being completely ionised – and opaque due to Compton scattering of photons by free electrons – to being almost exactly neutral, and transparent. The reason for this transition is that, under expansion, the photon temperature redshifts away ($T(a) \sim a^{-1}$) until the universe becomes cool enough, around $T = T^* \approx 3500K$, that electrons and protons can combine to form neutral hydrogen without being re-ionised by energetic photons (the *recombination* epoch). The process of recombination significantly reduces the *ionisation fraction* x_e of the universe, increasing the mean-free path for interactions between electrons and photons, which previously were tightly-coupled through Compton scattering. As $x_e \to 0$, the interaction rate drops below the expansion

rate, effectively *decoupling* photons from the baryons. The photons can then free-stream, and are detected today as the cosmic microwave background (CMB).

The mean temperature of the CMB today, $T = T_0 = 2.7255$ K, reflects the factor of ≈ 1100 expansion of the universe since recombination (Mather et al., 1994). Inhomogeneities in the universe at last scattering – in the local density, velocity and potential – and gravitational effects on the propagating photon cause the temperature and polarisation of the observed CMB to be *anisotropic*. Physically, the observed photon temperature depends on the radiation energy density ρ_r of the region of space from which it last scattered (since this determines the intrinsic temperature of the photons, through $\rho_\mathrm{r} \sim T^4$), the Doppler shift due to the bulk motion of the photo-baryonic fluid, and the local potential (which determines the global gravitational redshift). Secondary effects include the distribution of gravitational potentials along the line-of-sight (which can lens the photon), and the integrated Sachs-Wolfe (ISW) effect, which induces an additional gravitational redshift as large-scale potentials evolve during the passage of the photon.

The fractional size of the CMB fluctuations is small: $\Delta T/T_0 = 1.5 \times 10^{-5}$ (in the fiducial cosmology, via Eq. 1.15), where ΔT is the rms variation in temperature across the sky. This justifies the use of perturbation theory to describe the inhomogeneities in the early universe.

1.3.1 Physics of the CMB anisotropies

At its most basic level, the primary anisotropies of the CMB are a projection of the inhomogeneities in the universe at decoupling. We observe the spatial fluctuations in the photon distribution as they were, 'frozen in', at the last-scattering surface. After decoupling the photons *free-stream*, following null geodesics, and are affected thereafter only by evolving gravitational potentials and lensing effects (Sec. 1.4).

The comoving cosmological horizon size at decoupling is given by

$$\chi^* = \int_0^{t^*} \frac{c}{a(t)} dt = \int_0^{a^*} \frac{c}{a^2 H(a)} da = 2 d_H \sqrt{\frac{a_\mathrm{eq}}{\Omega_\mathrm{m}}} \left[\left(1 + \frac{a^*}{a_\mathrm{eq}} \right)^{1/2} - 1 \right], \tag{1.47}$$

where $d_H \equiv c/H_0$ is the Hubble radius today, a^* is the scale factor at decoupling, $a_\mathrm{eq} = \Omega_\mathrm{r}/\Omega_\mathrm{m}$ is the scale factor at matter-radiation equality, and we have made use of Eq. 1.36. Here Ω_m includes only *cold* matter components (i.e. CDM and baryons); not neutrinos, which are relativistic at

early times and therefore contribute instead to $\Omega_{\rm r}$ (Sec. 1.2.6). Curvature and dark energy are negligible in the early universe. Substituting numbers for the fiducial model into Eq. 1.47, we find $\chi^* = 280$ Mpc. This is the (comoving) distance that light can travel before decoupling, and therefore represents the causal horizon at this time. The proper size of the horizon at decoupling is $a^*\chi^* = 0.26$ Mpc in the fiducial model.

The angular diameter distance to decoupling is given by Eq. 13 of Chapter 4, and in the fiducial model is $d_A(a^*) = 12.7$ Mpc; the angle subtended today by the cosmological horizon at decoupling is hence $\theta = a^*\chi^*/d_A(a^*) = 1.2$ deg. This means that, in the standard Big Bang model, parts of the universe at decoupling which are now separated by more than 1.2 degrees on the sky were never in causal contact. They are observed, however, to have almost exactly the same temperature – suggesting that indeed they *were* in causal contact, in order that they could achieve thermal equilibrium. This apparent contradiction is called the *horizon problem*, which can be explained by invoking inflation (Sec. 1.2.3).

The size of the cosmological horizon at decoupling is important for understanding two important features of the CMB anisotropies. Firstly, it means that any primary fluctuations we observe on largest angular scales are *super-horizon*, and therefore *unprocessed* by sub-horizon physics; they correspond directly to the primordial fluctuations. Secondly, the *sound horizon* at decoupling χ_s^* – the maximum distance of causal effect for the photo-baryonic fluid – is closely related to the cosmological horizon. Modes of the photo-baryonic fluid grow as they enter the horizon and then begin to oscillate when radiation pressure becomes important (Fig. 1.4). These oscillations are forced by the gravitational potentials ψ and ϕ and their time-evolution (Hu & Sugiyama, 1995; Dodelson, 2003). The sound speed of the fluid is $c_s = c/\sqrt{3(1+R)}$, where $R = \rho_b/\rho_\gamma$ is the baryon-to-photon ratio, which remains small ($R < 1$) throughout decoupling for allowed models. The comoving sound horizon at decoupling χ_s^* is around 150 Mpc, approximately a factor $\sqrt{3}$ smaller than the cosmological horizon[2].

As modes enter the horizon the photo-baryonic perturbations are driven by the gravitational potential which decays during radiation domination (Sec. 1.2.4; Hu & Sugiyama, 1995). Those modes which correspond to the sound horizon size at decoupling ($k \approx k^* = 1/\chi_s^*$) grow to a maximum amplitude by decoupling, contributing excess angular power when these modes are

[2] $\chi_s^* = 144.6$ Mpc in the fiducial cosmology

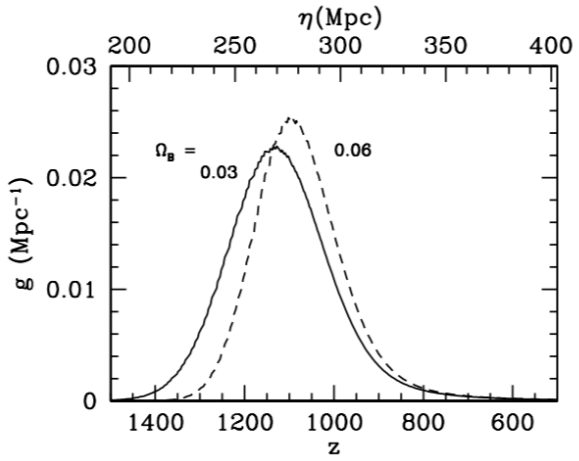

Figure 1.10: *Source: Fig. 7.10 of Dodelson (2003).* The visibility function g describing the last-scattering surface; $g(\eta)$ is the probability an observed photon last scattered at conformal time η, and peaks around $z \approx 1100$.

projected on the sky. The harmonics of these modes ($k = nk^*, n \in \mathbb{Z}$) are also at an extremum of their oscillation at decoupling, while intermediate modes have gone through a non-integer number of oscillations, and so will be observed with a reduced amplitude; this gives rise to the peak structure seen in CMB power spectra (e.g., Fig. 1.12). The physical scale of the sound horizon therefore sets the angular scale of the acoustic peaks in the CMB; around 0.7 deg, or $\ell \approx 200$.

Velocity fluctuations in the photo-baryonic fluid contribute additional anisotropy through the Doppler effect; because the velocity fluctuations are 90 degrees out-of-phase from density fluctuations, they boost power in the troughs of the CMB power spectrum (Hu et al., 1997).

The last-scattering surface is described by the *visibility function* $g(\eta) = -\dot{\tau}(\eta)e^{-\tau(\eta)}$, which specifies the probability that an observed photon was last scattered at conformal time η (Fig. 1.10). Here $\tau(\eta) = \int_\eta^{\eta_0} d\eta' \sigma_T n_e(\eta') a(\eta')$ is the integrated optical depth of the photo-baryonic fluid (discounting reionisation). The visibility function g is small at early times because τ is large since the free electron density n_e is large. At late times, well after decoupling $\tau \to 0$ as the free electron density decreases dramatically due to recombination, but the scattering rate $\dot{\tau}$ is also small as photon and baryons are effectively decoupled; the visibility function therefore has a peak, around $z \approx 1100$, and becomes negligible for $z \lesssim 800$ and $z \gtrsim 1400$.

Drawing together the effects described above, the temperature multipole observed today (at

$\eta = \eta_0$) due to a mode k can be written as

$$\begin{aligned} T_\ell(k,\eta_0) = &\int_0^{\eta_0} d\eta\ g(\eta)[\Theta_0(k,\eta) + \psi(k,\eta)] j_\ell[k(\eta_0-\eta)] \\ &- \int_0^{\eta_0} d\eta\ g(\eta)\frac{iv(k,\eta)}{k}\dot{j}_\ell[k(\eta_0-\eta)] \\ &+ \int_0^{\eta_0} d\eta\ e^{-\tau(\eta)}[\dot{\psi}(k,\eta) + \dot{\phi}(k,\eta)] j_\ell[k(\eta_0-\eta)] \end{aligned} \tag{1.48}$$

(Dodelson, 2003). The *effective temperature* fluctuation is given by $\Theta_0 + \psi$, i.e., the sum of the photon monopole (the intrinsic temperature fluctuation) and the gravitational potential (which attributes a redshift to the photons as they climb out of the local potential well). The first term in Eq. 1.48 is simply a projection of this effective temperature onto the sky, weighting contributions from different times appropriately by the visibility function. The second term is the analogous Doppler contribution from the bulk velocity of photo-baryonic fluid. The final term is the *integrated Sachs-Wolfe* (ISW) contribution due to the evolution of the potentials through which the photons propagate. The *early* ISW effect, due to the residual radiation contribution to the expansion rate around decoupling, adds power on scales just larger than the first acoustic peak. The *late-time* ISW effect comes from the presence of dark energy causing the decay of the late-time potentials. The relative size and scale dependence of these effects are shown in Fig. 1.11.

1.3.2 Silk damping

Prior to decoupling, photons diffuse, travelling along geodesics between scatterings by free electrons. This washes out power on the smallest scales. The physical photon mean free path is given by $\lambda_{\rm phys} = (n_e\sigma_T)^{-1}$, where σ_T is the Thomson scattering cross-section and n_e the physical free electron density. This corresponds to a *comoving* mean free path $\lambda = \lambda_{\rm phys}/a = (an_e\sigma_T)^{-1}$. In a time dt, the photon will undergo $dN = cdt/\lambda_{\rm phys} = d\eta/\lambda$ scatterings, where $d\eta$ is the corresponding increment in conformal time. The mean-square comoving distance travelled is hence $\langle d\chi^2\rangle = \lambda^2 dN = \lambda d\eta = d\eta/(an_e\sigma_T)$. These contributions add in quadrature to set a comoving *damping* wavenumber k_D, which at decoupling is given by

$$k_D^2(\eta^*) = \left[\int_0^{\eta^*} \frac{d\eta'}{a(\eta')n_e(\eta')\sigma_T}\right]^{-1}. \tag{1.49}$$

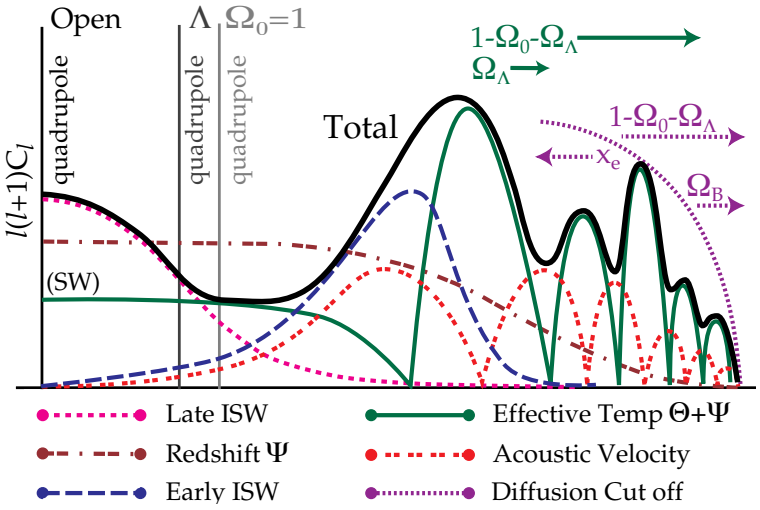

Figure 1.11: *Source: adapted from Fig. 5 of Hu et al. (1997).* Contributions to the primary CMB temperature power spectrum: the effective temperature (*green*), the Doppler contribution from the velocity of the fluid (*red*), and the early and late-time ISW (*pink and blue*). Small-scale fluctuations are suppressed by diffusion damping. The origin of the horizontal axis depends on the geometry of the universe, as indicated by the vertical lines, while the precise peak structure is a sensitive function of the background cosmological parameters (*Planck* Collaboration, 2015c).

This expression, derived heuristically, is a good approximation to a more accurate formula which accounts for the baryon content (Dodelson, 2003). Fluctuations at small scales ($k \gg k_D \approx 0.12h$ Mpc^{-1}), then, are exponentially damped by photon diffusion, while large scale perturbations ($k \ll k_D$) are unaffected. This is called *Silk damping*. The suppression of power in the small-scale photon distribution at decoupling (visible in the lower panel of Fig. 1.4) corresponds directly to the suppression of power in the small-scale angular power spectrum of the CMB for $\ell > \ell_D \approx 1000$ (Fig. 1.12).

1.3.3 CMB polarisation

The CMB is *polarised*. If electrons in the photo-baryonic fluid are, in their rest frame, subject to a quadrupolar pattern of incoming radiation from the surrounding photon bath, the scattered radiation perpendicular to the plane of the quadrupole will be linearly polarised (Fig. 1.13, left panel).

A quadrupole moment in the rest frame of the electrons is generated by the velocity perturbations of the photo-baryonic fluid in the inhomogeneous universe. The polarisation field of the CMB can be decomposed into two physical components, E-modes and B-modes (Kamionkowski et al.,

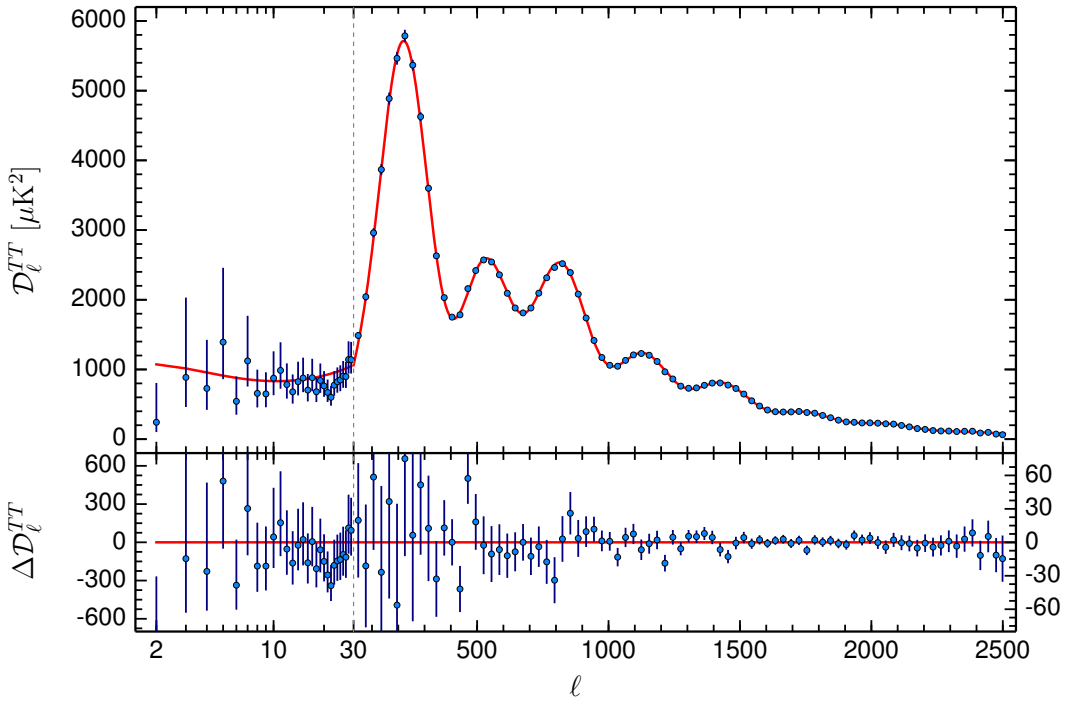

Figure 1.12: *Source: Fig. 48 of Planck Collaboration (2015b).* CMB temperature power spectrum as measured by the *Planck* satellite. The peak structure coming from acoustic oscillations of the photo-baryonic fluid is clearly visible. Small-scale power is suppressed by Silk damping, while the large-scale power reflects primordial fluctuations at super-horizon scales (the *Sachs-Wolfe plateau*, which appears approximately flat in $\mathcal{D}_\ell = \ell(\ell+1)C_\ell/(2\pi)$ at $\ell \lesssim 30$). The peaks are broadened by projection, the finite width of the visibility function g, and the Doppler contribution from the velocity of the fluid.

1997). These quantities are rotationally-invariant measures of the local pattern of polarisation vectors around any given line-of-sight (Fig. 1.13, right panel).

Within linear theory, only E-mode polarisation is generated by the scalar perturbations to the metric considered above. To see heuristically why this should be the case, consider an electron and its surrounding photo-baryonic fluid falling towards an overdensity, as projected in the plane of the sky. In its rest frame, as well as an isotropic bath of photons from the surrounding fluid, it will experience quadrupolar incident radiation; cooler along the radial direction – towards the overdensity – and warmer in the tangential direction. This is essentially a *tidal* effect. The linear polarisation vector generated from scattering along the line of sight by this electron will then be *radially* aligned since the amplitude of the electric field from the warmer region is larger. By circular symmetry, this argument holds at any position around the overdensity, and hence an $E < 0$ pattern is generated. A similar argument holds for underdensities, which generate

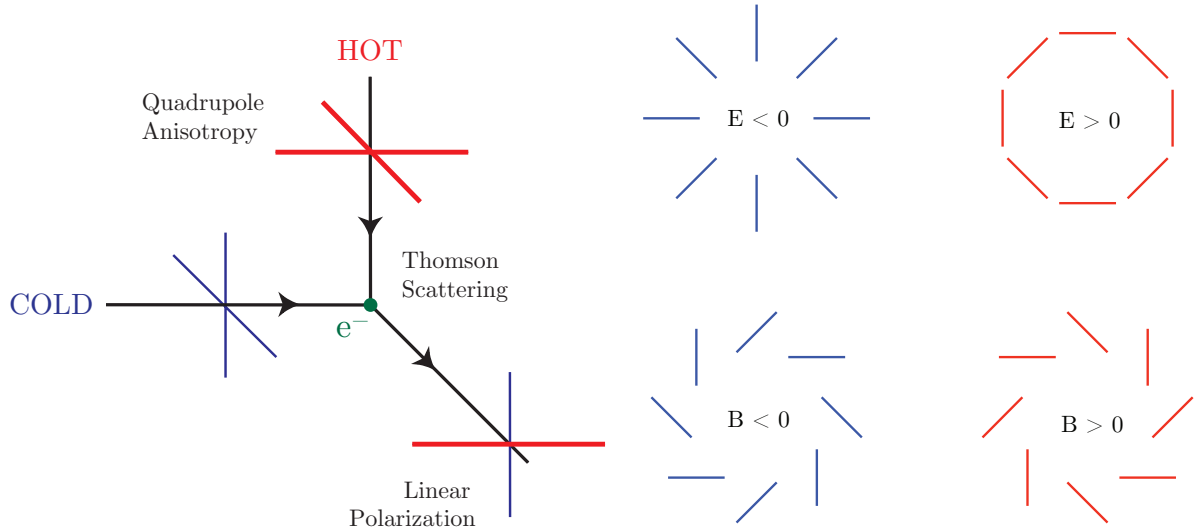

Figure 1.13: *Source: Figs. 4 & 5 of Riotto (2010).* (*Left panel*) Incident quadrupolar radiation gives rise to linearly polarised scattered light. A quadrupole moment in the rest frame of the electrons is generated by the velocity perturbations of the photo-baryonic fluid in the inhomogeneous universe. (*Right panel*) The polarisation field can be decomposed into two physical components, E-modes and B-modes. These quantities are rotationally-invariant measures of the local pattern of polarisation vectors around any given line-of-sight. Gravitational waves from inflation would imprint tensor perturbations on the metric; at late times these would be manifest as a quadrupolar component to the radiation perturbations, generating both E- and B-mode polarisation. Primordial B-modes have not been detected to date.

a local, positive E-mode pattern. In linear theory, with scalar perturbations, the velocity field is *irrotational*, and hence only E-modes are generated (Seljak & Zaldarriaga, 1997; Zaldarriaga & Seljak, 1997; Spergel & Zaldarriaga, 1997). Because the generation of polarisation relies on incident quadrupolar radiation, E-modes are primarily generated late on in the decoupling phase, when the photon mean free path increases; this increases the frequency of interactions with distant photons, and therefore photons with different temperatures.

The large-scale E-mode signal is sensitive to the process of *reionisation*. The universe becomes reionised through astrophysical processes at late times, starting around redshift $z = z_{\rm re} \approx 10$ and ending around $z \approx 6$. This leaves the universe again permeated with free electrons. Just as at recombination, the temperature quadrupole incident on these free-elections will be scattered as linearly polarised radiation. The quadrupole at reionisation is given by the free-streamed monopole from the last-scattering surface:

$$T_2(\chi_{\rm re}) = (T_0 + \psi)(\chi^*) j_2 \left[k(\chi_{\rm re} - \chi^*)\right], \tag{1.50}$$

where $\chi_{\rm re}$ is comoving horizon size at reionisation (Zaldarriaga, 1997). Using the properties of the spherical Bessel function, the peak of the quadrupole is at wavenumber $k \approx 2/(\chi_{\rm re} - \chi^*)$, which appears on the sky at multipole $\ell \approx k(\chi_0 - \chi_{\rm re}) \approx 2(\chi_0 - \chi_{\rm re})/(\chi_{\rm re} - \chi^*) \approx 5$. This contribution is clearly visible in the E-mode power spectrum (Fig. 1.2).

The scattering of CMB photons by free electrons in the reionised universe washes out power in the anisotropies (e.g., Griffiths et al., 1999). We can easily obtain a good estimate of the magnitude of this effect by computing the optical depth to reionisation assuming instantaneous reionisation at $z_{\rm re}$:

$$\begin{aligned}\tau(\chi_{\rm re}) &= \int_{\chi_{\rm re}}^{\chi_0} a n_e(a) \sigma_T d\chi \\ &= \frac{2cn_{e,0}\sigma_T}{3\Omega_{\rm m}H_0}\left(\left[\Omega_{\rm m}(1+z_{\rm re})^3 + \Omega_\Lambda\right]^{1/2} - 1\right)\end{aligned} \tag{1.51}$$

where $n_{e,0} = \Omega_{\rm b}\rho_{\rm crit}(1-3Y/4)/m_p$ is the free-electron density today, Y is the helium fraction and m_p is the proton mass. For the fiducial model, this yields

$$\tau(z_{\rm re}) \approx 0.08\left(\frac{1+z_{\rm re}}{11}\right)^{3/2}, \tag{1.52}$$

which is comparable to the *Planck*-derived value (*Planck* Collaboration, 2015c).

1.3.4 The CMB power spectra

The main observable statistic of interest in CMB analysis is the angular power spectrum C_ℓ. The reasons are various: importantly, because the CMB anisotropies by construction have zero mean, the power spectrum is the first non-vanishing moment of the CMB distribution function (in harmonic space). This makes the power spectrum cosmologically interesting, yet analytically more straightforward than dealing with the full distribution function in pixel space. The construction of estimators from the raw CMB maps is also correspondingly simple. The power spectrum is most useful, however, because the primary CMB can be modelled as a zero-mean *Gaussian* random field; initial Gaussian fluctuations will evolve linearly to decoupling, which will then project to a Gaussian field on the sky. This means that the power spectrum contains all the information relevant to the statistics of the field, fully specifying its distribution function (Sec. 1.1).

In practice, we have access to only one realisation of the CMB fields (one each of T, E and B;

Sec. 1.3.3). We therefore have only a finite number of modes from which to estimate the angular power spectrum: observed over the whole sky, there are $2\ell+1$ independent azimuthal modes with a given multipole ℓ. Observed over a finite sky fraction $f_{\rm sky}$, different azimuthal modes become degenerate, reducing the effective number of independent modes to $(2\ell+1)f_{\rm sky}$ (e.g., Hobson & Magueijo, 1996). Given the model of the CMB as a statistically isotropic random field, the spherical harmonic coefficients $a_{\ell m}$ satisfy Eq. 1.13:

$$\langle a_{\ell m} a^*_{\ell' m'} \rangle = C_\ell \delta_{\ell\ell'} \delta_{mm'}, \tag{1.53}$$

and we can define an unbiased estimator of the underlying power spectrum which generates the observed realisation $\{a_{\ell m}\}$ by

$$\hat{C}_\ell = \frac{1}{2\ell+1} \sum_{m=-\ell}^{m=+\ell} a^*_{\ell m} a_{\ell m}. \tag{1.54}$$

This quantity, sometimes called the *pseudo-C_ℓ's*, is a *sufficient statistic*; i.e., the probability of observing a set of harmonic coefficients $\{a_{\ell m}\}$ depends only on $\hat{C}_\ell$. This is extremely useful, since it allows for lossless compression of the data – the two dimensional CMB map – to the one-dimensional empirical power spectrum, speeding up the computational analysis, e.g., for parameter estimation.

Using Gaussianity of the modes $a_{\ell m}$, the sampling distribution of the power spectrum estimator is a generalised χ^2-distribution with $2\ell+1$ degrees of freedom. This has a probability density

$$p(\hat{C}_\ell | C_\ell) = \frac{1}{\Gamma(\ell + 1/2)} \frac{1}{\hat{C}_\ell} \left(\frac{2\ell+1}{2C_\ell} \hat{C}_\ell \right)^{\ell+1/2} \exp\left[-\frac{2\ell+1}{2C_\ell} \hat{C}_\ell \right], \tag{1.55}$$

and variance $\mathbb{C}_\ell = 2C_l^2/(2\ell+1)$. This generalises to a Wishart distribution when considering the sampling distribution of the empirical power spectra estimated from multiple correlated Gaussian fields. Since the estimator in Eq. 1.54 is a sum of many random variables of finite variance, then for large ℓ this tends to a Gaussian sampling distribution by the central limit theorem:

$$p(\hat{C}_\ell | C_\ell) \approx \frac{1}{\sqrt{2\pi\mathbb{C}_\ell}} \exp\left[-\frac{1}{2} (\hat{C}_\ell - C_\ell)^T \mathbb{C}_\ell^{-1} (\hat{C}_\ell - C_\ell) \right]. \tag{1.56}$$

This Gaussian approximation has been used to date by the small-scale ground-based CMB exper-

iments as the basis for the likelihood: $\mathcal{L}(\theta) \equiv \mathbb{P}(\hat{C}_\ell|\theta) = \mathbb{P}[\hat{C}_\ell|C_\ell(\theta)]$, where the last equality holds because the CMB is a zero-mean Gaussian field, and so is fully described by its power spectrum (Sievers et al., 2013; Naess et al., 2014; Hou et al., 2014).

An important caveat to the discussion above is that, in practice, CMB maps used for power spectral analyses are not well-defined across the full sky. Effects such as Galactic and point source emission must be masked out to prevent contamination of the inferred power spectrum (and therefore inferred parameters). Furthermore, most CMB experiments observe only a fraction of the sky. The corollary is that the observed CMB is modulated by an experiment-dependent window function, which includes masking and the effect of the telescope resolution (the *beam*). Using the naive estimator of Eq. 1.54 in the presence of a window function will give spurious results (Chapter 5). The effects of a window function, masking and the beam can be corrected for by introducing a *mode-coupling matrix* to provide an unbiased estimator (Sec. 5.3; Das et al., 2013). Alternatively, since the window function is localised in real-space, for the largest scales a pixel-based likelihood can be preferable, although this is computationally unfeasible across all scales for current high-resolution CMB maps (*Planck* Collaboration, 2015c).

Atmospheric noise and ground pickup (which dominate at large scales for ground-based telescopes), and detector noise (an almost white-noise contribution which dominates at small scales) will additionally induce a *noise bias* for the estimator given in Eq. 1.54. This bias can be handled by splitting the data by time of observation (e.g., 6 month chunks): a map i for each of the N data splits is constructed such that the atmospheric and detector noise correlations between each split are negligible. Using $\langle a^{i*}_{\ell m} a^j_{\ell m}\rangle = C_\ell + \delta_{ij} N_\ell$, where N_ℓ is the noise power spectrum, a noise-bias-free power spectrum estimator can then be constructed as

$$\hat{C}_\ell = \frac{1}{(2\ell+1)N(N-1)} \sum_{m,i,j} a^{i*}_{\ell m} a^j_{\ell m} (1-\delta_{ij}), \tag{1.57}$$

where the δ-function ensures the split auto-spectra, which contain noise bias, are discounted (Das et al., 2013). This procedure is lossy in that it removes a fraction $1/N$ of the data. In this sense N should be taken to be as large as possible while requiring that the noise correlations between splits remain consistent with null. For instance, $N = 4$ is chosen for the ACT and ACTPol analyses. Note that ground pickup – which is likely to be common between splits – will not be removed by

this procedure.

There are four non-vanishing power spectra given the three physical primary CMB fields: $C_\ell^{TT}, C_\ell^{TE}, C_\ell^{EE}$ and C_ℓ^{BB}. The TB and EB cross-spectra of the CMB are identically zero due to parity considerations (Kamionkowski et al., 1997). We show the temperature and polarisation power spectra in the fiducial model in Fig. 1.2 and the *Planck*-derived temperature power spectrum in Fig. 1.12. Although the primary, unlensed CMB can be well modelled as a Gaussian random field, other components such as foregrounds and the lensed temperature and polarisation are non-Gaussian fields. The non-Gaussian components will still contribute to the observed temperature and polarisation power. These must be carefully modelled, therefore, when comparing observations to theory. We discuss the effects of different cosmological parameters on the CMB power spectra in Chapter 4.

Although B-modes cannot be generated from scalar perturbations, within inflationary theory primordial tensor perturbations generically produce B-modes, with an amplitude determined by the energy scale of inflation (e.g., Liddle, 1994; Baumann, 2009):

$$E_{\rm inf} \approx 10^{16} (r/0.01)^{0.25} \text{ GeV}. \tag{1.58}$$

The free parameter is the *tensor-to-scalar ratio* r, which relates the amplitude of the tensor power spectrum to the primordial scalar power spectrum of the comoving curvature perturbation. In the usual power-law parameterisation of the initial power spectra, $r = A_T/A_s$, where the (implicit) pivot scale is typically taken to be $k = 0.05$ Mpc^{-1} (*Planck* Collaboration, 2015f). Current observations are consistent with no primordial B-mode signal; $r < 0.07$ at 95% confidence (Array et al., 2015). But because of the implications of r as a 'smoking gun' signature of inflation, CMB experiments are observing – and are being developed – to push this limit further, and advances in this area are being made rapidly (e.g., ACTPol, AdvACT, POLARBEAR, SPTpol, SPT-3G, SPIDER, LiteBIRD, BICEP/Keck and the Simons Array; Sec. 1.9).

Gravitational lensing of the CMB is the deflection of photons from the last-scattering surface by the intervening large-scale structure (Sec. 1.4). Lensing can generate B-modes through the coupling of the primary E-mode polarisation with the lensing potential ϕ (Zaldarriaga & Seljak, 1998; Lewis & Challinor, 2006). In terms of the the unlensed E-mode power spectrum $\tilde{C}_\ell^{EE}$ and

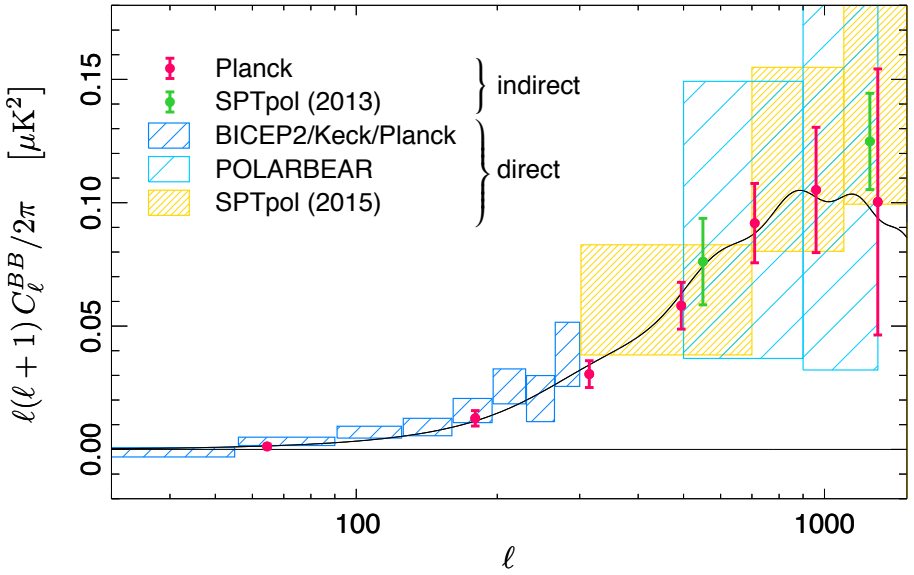

Figure 1.14: *Source: Fig. 9 of Planck Collaboration (2015h).* B-mode power spectrum C_ℓ^{BB} as measured by various experiments (*points and boxes*). The ΛCDM expectation is also shown (*black solid line*). The *Planck* result represents a 12σ detection of lensing-induced B-modes, with an amplitude ($A = 0.97 \pm 0.08$) consistent with the ΛCDM expectation ($A = 1$). The 'direct' results correspond to estimation of C_ℓ^{BB} from the usual quadratic combination of the observed B-mode modes. The 'indirect' measurements are made by constructing a B-mode *template* from the observed E-mode and lensing maps, and cross-correlating this template in Fourier space with the observed B-mode map.

the lensing potential power spectrum $C_\ell^{\phi\phi}$, the B-mode power spectrum due to lensing can be written

$$C_\ell^{BB,\mathrm{lensing}} = \int \frac{d^2\mathbf{l}'}{(2\pi)^2} \left[\mathbf{l}' \cdot (\mathbf{l} - \mathbf{l}')\right]^2 C_{|\mathbf{l}-\mathbf{l}'|}^{\phi\phi} \tilde{C}_{l'}^{EE} \sin^2(2\phi_{\mathbf{l}\mathbf{l}'}), \tag{1.59}$$

where $\phi_{\mathbf{l}\mathbf{l}'}$ is the angle between $\mathbf{l}$ and $\mathbf{l}'$ in the Fourier plane (Lewis & Challinor, 2006). Recent measurements from the *Planck* satellite have provided a state-of-the-art 12σ detection of the lensing-induced B-mode power spectrum, shown in Fig. 1.14 (*Planck* Collaboration, 2015h). When searching to measure the tensor-to-scalar ratio r, it is important to account for the lensing-induced B-mode, which acts essentially as a foreground to the primordial B-mode signal (BICEP2/Keck and *Planck* Collaborations, 2015). Although any primordial signal is expected to appear at larger angular scales than the peak in the lensing B-mode power spectrum, the lensing contribution and its inherent uncertainties must be properly modelled and marginalised for any definitive result.

1.3.5 Secondary anisotropies and parameters

Other physical effects, beyond the primary anisotropies described above, contribute to the observed CMB temperature and polarisation maps. These include the Sunyaev-Zel'dovich effect, gravitational lensing (Sec. 1.4), and additional foregrounds such as galactic emission, radio point sources and the cosmic infrared background (Dunkley et al., 2013). These can be accounted for at various levels. At the map level, multi-frequency coverage can be used to separate different components by using a physical prior on the frequency dependences of each source, or by blind component-separation algorithms. At the power spectrum level, one can include multiple components in the data model and estimate 'CMB-only' band-powers which are marginalised over the nuisance parameters which parameterise the secondary sources (Dunkley et al., 2013). At the level of parameters, one can – more directly – sample the cosmological parameters and nuisance parameters simultaneously. In Chapter 4 we discuss in detail many of the physics effects of particular cosmological parameters on the CMB, and we defer further discussion of parameters until then.

1.4 Gravitational lensing

Free-streaming photons propagate along null geodesics. They are therefore deflected by the presence of matter, as this matter induces curvature of the local space-time; this effect is called *gravitational lensing*. Cosmological lensing by the large-scale structure of the universe has been measured through observations of both galaxies and the CMB; the observed images of these sources are distorted by the intervening matter (e.g., Heymans et al., 2013; Becker et al., 2015; Kuijken et al., 2015; *Planck* Collaboration, 2015e). We defer discussion of galaxy lensing to Chapter 5, instead focussing here on the theory and practical extraction of the CMB lensing signal.

1.4.1 CMB lensing

The presence of gravitating structure causes CMB photons to be deflected as they propagate through the universe (see Lewis & Challinor, 2006, for a review). The effect on the observed CMB is an apparent shift in the angular position of a given photon, as projected on the sky, relative to

its unlensed source direction:

$$T(\mathbf{n}) = \tilde{T}(\mathbf{n} + \mathbf{d}(\mathbf{n})), \tag{1.60}$$

where T is the *lensed* temperature field, $\tilde{T}$ is the *unlensed* temperature field, and $\mathbf{d}$ is the deflection field induced by matter fluctuations transverse to the line of sight $\mathbf{n}$. The polarisation scalar fields Q and U satisfy an analogous relation between the lensed and unlensed field.

Considering the path of a cosmological photon through a linearly perturbed FRW metric, one can show the deflection angle is given by

$$\mathbf{d}(\mathbf{n}) = -2 \int_0^{\chi^*} d\chi \frac{f_k(\chi^* - \chi)}{f_k(\chi^*) f_k(\chi)} \nabla_{\mathbf{n}} \Psi(\mathbf{n}, \chi; \eta_0 - \chi), \tag{1.61}$$

where η_0 is the conformal time today and $\nabla_{\mathbf{n}}$ is the covariant derivative on the sphere, written as ∇ hereafter for conciseness (Lewis & Challinor, 2006). The Born approximation has been adopted (i.e., integrals are performed along the unperturbed path of the photon), and the result holds to first order in the Weyl potential Ψ. The lensing potential ϕ, a two-dimensional field on the sphere, is then defined such that $\mathbf{d} \equiv \nabla\phi$, giving

$$\phi(\mathbf{n}) = -2 \int_0^{\chi^*} d\chi \frac{f_k(\chi^* - \chi)}{f_k(\chi^*) f_k(\chi)} \Psi(\mathbf{n}, \chi; \eta_0 - \chi). \tag{1.62}$$

In this section we reserve ϕ to denote the lensing potential rather than, as elsewhere, the curvature perturbation in the conformal Newtonian gauge. Neglecting curvature and anisotropic stress, the Weyl potential Ψ is equal to the gravitational potential, and in GR can be related directly to the total comoving density perturbation $\rho\Delta$:

$$\nabla^2 \Psi = 4\pi G a^2 \rho \Delta, \tag{1.63}$$

where ∇^2 is the 3D Laplacian (Lewis & Challinor, 2006). The above expression is a restatement of Eq. 1.31. A realisation of the density field therefore maps directly onto a realisation of the lensing potential, and, through Eq. 1.60, onto a realisation of the lensed CMB (given the unlensed CMB). This connection suggests that observations of the lensed CMB can be used to learn about the large-scale structure distribution and its statistics. We now demonstrate how this inference is made in practice.

1.4.2 CMB lensing reconstruction

In this section we work in the flat-sky approximation, allowing us to replace the spherical-harmonic transform with the Fourier transform. Due to statistical isotropy, the unlensed CMB temperature and polarisation fields $\tilde{x}, \tilde{x}' \in \{\tilde{T}, \tilde{E}, \tilde{B}\}$ are diagonal in Fourier space:

$$\langle \tilde{x}(\boldsymbol{\ell})\tilde{x}'(\boldsymbol{\ell}')\rangle = \tilde{C}_\ell^{xx'}\delta(\boldsymbol{\ell}+\boldsymbol{\ell}), \tag{1.64}$$

where the average $\langle\cdot\rangle$ is over realisations of the CMB (Eq. 1.13). However, the *lensed* temperature and polarisation fields are not diagonal in Fourier space:

$$\langle x(\boldsymbol{\ell})x'(\boldsymbol{\ell}')\rangle = f_{xx'}(\boldsymbol{\ell},\boldsymbol{\ell}')\phi(\boldsymbol{L}), \tag{1.65}$$

to first order in ϕ, where $\boldsymbol{L} = \boldsymbol{\ell} + \boldsymbol{\ell}' \neq \boldsymbol{0}$ (e.g., Hu & Okamoto, 2002; Kesden et al., 2003). Here $x, x' \in \{T, E, B\}$ are the *lensed* fields. The pair-response functions $f_{xx'}(\boldsymbol{\ell},\boldsymbol{\ell}')$ quantify the correlations between CMB modes induced by a given lensing mode $\phi(\boldsymbol{L})$; lensing *couples* modes in temperature and polarisation. As an expository example, we now derive the TT pair response function f_{TT}; the pair-response functions for all other quadratic pairs can be derived analogously, which for completeness we list in Table 1.1. Expanding Eq. 1.60 to first order in the lensing potential and taking the Fourier transform yields

$$T(\boldsymbol{\ell}) = \tilde{T}(\boldsymbol{\ell}) - \int \frac{d^2\boldsymbol{\ell}'}{2\pi}\boldsymbol{\ell}'\cdot(\boldsymbol{\ell}-\boldsymbol{\ell}')\phi(\boldsymbol{\ell}-\boldsymbol{\ell}')\tilde{T}(\boldsymbol{\ell}') + \ldots, \tag{1.66}$$

and hence,

$$\begin{aligned} T(\boldsymbol{\ell})T(\boldsymbol{\ell}') &= \tilde{T}(\boldsymbol{\ell})\tilde{T}(\boldsymbol{\ell}') \\ &- \int \frac{d^2\boldsymbol{\ell}''}{2\pi}\left[\boldsymbol{\ell}''\cdot(\boldsymbol{\ell}-\boldsymbol{\ell}'')\phi(\boldsymbol{\ell}-\boldsymbol{\ell}'')\tilde{T}(\boldsymbol{\ell}')\tilde{T}(\boldsymbol{\ell}'') + \boldsymbol{\ell}''\cdot(\boldsymbol{\ell}'-\boldsymbol{\ell}'')\phi(\boldsymbol{\ell}'-\boldsymbol{\ell}'')\tilde{T}(\boldsymbol{\ell})\tilde{T}(\boldsymbol{\ell}'')\right] + \ldots \end{aligned} \tag{1.67}$$

Field pair	$2\pi f_{xx'}(\boldsymbol{\ell},\boldsymbol{\ell}')$
TT	$\tilde{C}_\ell^{TT}(\boldsymbol{L}\cdot\boldsymbol{\ell})+\tilde{C}_{\ell'}^{TT}(\boldsymbol{L}\cdot\boldsymbol{\ell}')$
TE	$\tilde{C}_\ell^{TE}\cos 2\varphi_{\boldsymbol{\ell}\boldsymbol{\ell}'}(\boldsymbol{L}\cdot\boldsymbol{\ell})+\tilde{C}_{\ell'}^{TE}(\boldsymbol{L}\cdot\boldsymbol{\ell}')$
TB	$\tilde{C}_\ell^{TE}\sin 2\varphi_{\boldsymbol{\ell}\boldsymbol{\ell}'}(\boldsymbol{L}\cdot\boldsymbol{\ell})$
EE	$[\tilde{C}_\ell^{EE}(\boldsymbol{L}\cdot\boldsymbol{\ell})+\tilde{C}_{\ell'}^{EE}(\boldsymbol{L}\cdot\boldsymbol{\ell}')]\cos 2\varphi_{\boldsymbol{\ell}\boldsymbol{\ell}'}$
EB	$[\tilde{C}_\ell^{EE}(\boldsymbol{L}\cdot\boldsymbol{\ell})-\tilde{C}_{\ell'}^{BB}(\boldsymbol{L}\cdot\boldsymbol{\ell}')]\sin 2\varphi_{\boldsymbol{\ell}\boldsymbol{\ell}'}$
BB	$[\tilde{C}_\ell^{BB}(\boldsymbol{L}\cdot\boldsymbol{\ell})+\tilde{C}_{\ell'}^{BB}(\boldsymbol{L}\cdot\boldsymbol{\ell}')]\cos 2\varphi_{\boldsymbol{\ell}\boldsymbol{\ell}'}$

Table 1.1: Pair-response functions $f_{xx'}$ for CMB temperature and polarisation fields $x, x' \in \{T, Q, U\}$, defined by Eq. 1.65, where $\boldsymbol{L} = \boldsymbol{\ell} + \boldsymbol{\ell}'$. A tilde denotes an *unlensed* quantity.

Taking the expectation over realisations of the CMB gives

$$\begin{aligned}\langle T(\boldsymbol{\ell})T(\boldsymbol{\ell}')\rangle &= \delta(\boldsymbol{\ell}+\boldsymbol{\ell}')\tilde{C}_\ell \\ &- \int \frac{d^2\boldsymbol{\ell}''}{2\pi}\left[\boldsymbol{\ell}''\cdot(\boldsymbol{\ell}-\boldsymbol{\ell}'')\phi(\boldsymbol{\ell}-\boldsymbol{\ell}'')\delta(\boldsymbol{\ell}'+\boldsymbol{\ell}'')\tilde{C}_{\ell'}^{TT} + \boldsymbol{\ell}''\cdot(\boldsymbol{\ell}'-\boldsymbol{\ell}'')\phi(\boldsymbol{\ell}'-\boldsymbol{\ell}'')\delta(\boldsymbol{\ell}+\boldsymbol{\ell}'')\tilde{C}_\ell^{TT}\right] + \ldots,\end{aligned} \tag{1.68}$$

which, on performing the integral, provides the required result:

$$\langle T(\boldsymbol{\ell})T(\boldsymbol{\ell}')\rangle = \frac{1}{2\pi}\left[\tilde{C}_\ell^{TT}(\boldsymbol{L}\cdot\boldsymbol{\ell}) + \tilde{C}_{\ell'}^{TT}(\boldsymbol{L}\cdot\boldsymbol{\ell}')\right]\phi(\boldsymbol{L}) + O(\phi^2), \tag{1.69}$$

for $\boldsymbol{L} = \boldsymbol{\ell} + \boldsymbol{\ell}' \neq \boldsymbol{0}$. Lensing from a fixed realisation of the large-scale structure therefore breaks statistical isotropy, in particular causing a local distortion of the observed CMB angular power spectra. Physically, this can be understood as the effect of spatially-varying magnification in dilating or compressing apparent angular scales. Over a region of sky with magnification μ, an intrinsic angular multipole ℓ is shifted to an observed angular multipole ℓ/μ. We note that further averaging over realisations of large-scale structure restores isotropy for the lensed CMB fields (e.g., Eq. 1.65).

That CMB modes are coupled in the presence of a lensing potential leads naturally to defining a lensing estimator which is *quadratic* in the observed CMB fields:

$$\hat{\phi}_{xx'}(\boldsymbol{L}) = A_{xx'}(\boldsymbol{L})\int d^2\boldsymbol{\ell}\; F_{xx'}(\boldsymbol{\ell}, \boldsymbol{L}-\boldsymbol{\ell})x(\boldsymbol{\ell})x'(\boldsymbol{L}-\boldsymbol{\ell}), \tag{1.70}$$

where $A_{xx'}(\boldsymbol{L})$ is an (as yet undetermined) normalisation term and $F_{xx'}(\boldsymbol{\ell}, \boldsymbol{\ell}')$ is an arbitrary filter

which specifies the weight given to each pair of CMB modes in the estimator. Requiring that the estimator is unbiased at first order in the lensing potential, i.e., $\langle\hat{\phi}_{xx'}(\boldsymbol{L})\rangle = \phi(\boldsymbol{L})$, then specifies

$$A_{xx'}(\boldsymbol{L}) = \left[\int d^2\boldsymbol{\ell}\; F_{xx'}(\boldsymbol{\ell}, \boldsymbol{L}-\boldsymbol{\ell}) f_{xx'}(\boldsymbol{\ell}, \boldsymbol{L}-\boldsymbol{\ell})\right]^{-1}. \tag{1.71}$$

Further requiring that $\hat{\phi}$ is a *minimum variance* estimator, i.e., such that $\langle\hat{\phi}_{xx'}(\boldsymbol{L})\hat{\phi}_{xx'}(\boldsymbol{L})\rangle$ is minimised with respect to the filter F, specifies the precise form of the filters F (see, e.g., Hu & Okamoto, 2002, for the full expressions). For instance, the minimum variance TT filter is

$$F_{TT}(\boldsymbol{\ell}, \boldsymbol{L}-\boldsymbol{\ell}) = \frac{f_{TT}(\boldsymbol{\ell}, \boldsymbol{L}-\boldsymbol{\ell})}{2C_\ell^{TT} C_{|\boldsymbol{L}-\boldsymbol{\ell}|}^{TT}}, \tag{1.72}$$

where the denominator involves the *total* observed power spectra i.e., the lensed signal plus the detector and atmospheric noise power spectra. This result makes physical sense; noisy modes are down-weighted, while mode pairs which are sensitive to a given lensing mode (i.e. f is large) are up-weighted. Eq. 1.72 is essentially a signal-to-noise weighting for each pair of modes.

We now explicitly derive Eq. 1.72 to outline the general approach for all quadratic pairs, suppressing the TT label here for conciseness. Absorbing the normalisation term $A(\boldsymbol{L})$ into the filter G, we have for a general quadratic estimator $\hat{\phi}(\boldsymbol{L}) = \int d^2\boldsymbol{\ell}\; G(\boldsymbol{\ell}, \boldsymbol{L}-\boldsymbol{\ell}) T(\boldsymbol{\ell}) T(\boldsymbol{L}-\boldsymbol{\ell})$. The condition that $\hat{\phi}$ is unbiased can then be written as $I[G] = 1$ using Eq. 1.65, where we define the functional

$$I[G] \equiv \int d^2\boldsymbol{\ell}\; G(\boldsymbol{\ell}, \boldsymbol{L}-\boldsymbol{\ell}) f(\boldsymbol{\ell}, \boldsymbol{L}-\boldsymbol{\ell}). \tag{1.73}$$

We then compute

$$\langle\hat{\phi}(\boldsymbol{L})\hat{\phi}^*(\boldsymbol{L}')\rangle = \int d^2\boldsymbol{\ell} d^2\boldsymbol{\ell}'\; G(\boldsymbol{\ell}, \boldsymbol{L}-\boldsymbol{\ell}) G(\boldsymbol{\ell}', \boldsymbol{L}'-\boldsymbol{\ell}') \langle T(\boldsymbol{\ell}) T(\boldsymbol{L}-\boldsymbol{\ell}) T(-\boldsymbol{\ell}') T(\boldsymbol{\ell}'-\boldsymbol{L}')\rangle, \tag{1.74}$$

using Wick's theorem to evaluate the four-point function, viz.,

$$\langle T(\boldsymbol{\ell}) T(\boldsymbol{L}-\boldsymbol{\ell}) T(-\boldsymbol{\ell}') T(\boldsymbol{\ell}'-\boldsymbol{L}')\rangle = \delta(\boldsymbol{L})\delta(\boldsymbol{L}') C_\ell C_{\ell'} + \delta(\boldsymbol{L}-\boldsymbol{L}') C_\ell C_{|\boldsymbol{L}-\boldsymbol{\ell}|} \left[\delta(\boldsymbol{\ell}-\boldsymbol{\ell}') + \delta(\boldsymbol{\ell}+\boldsymbol{\ell}'-\boldsymbol{L})\right] \tag{1.75}$$

to obtain

$$\langle\hat{\phi}(\boldsymbol{L})\hat{\phi}^*(\boldsymbol{L}')\rangle = 2\delta(\boldsymbol{L}-\boldsymbol{L}') \int d^2\boldsymbol{\ell}\; G^2(\boldsymbol{\ell}, \boldsymbol{L}-\boldsymbol{\ell}) C_\ell C_{|\boldsymbol{L}-\boldsymbol{\ell}|}, \tag{1.76}$$

for $\boldsymbol{L}, \boldsymbol{L}' \neq \mathbf{0}$ and using the symmetry of G in its arguments. For the minimum variance estimator, we then wish to minimise the functional

$$V[G] \equiv \int d^2\boldsymbol{\ell}\; G^2(\boldsymbol{\ell}, \boldsymbol{L}-\boldsymbol{\ell}) C_\ell C_{|\boldsymbol{L}-\boldsymbol{\ell}|} \tag{1.77}$$

subject to the constraint $I[G] = 1$, which is easily peformed using functional calculus. Defining

$$\mathcal{L}[G] \equiv V[G] - \lambda(I[G]-1) \tag{1.78}$$

and setting the derivative

$$\frac{\partial \mathcal{L}}{\partial G} = \int d^2\boldsymbol{\ell}\;\left[2G(\boldsymbol{\ell}, \boldsymbol{L}-\boldsymbol{\ell}) C_\ell C_{|\boldsymbol{L}-\boldsymbol{\ell}|} - \lambda f(\boldsymbol{\ell}, \boldsymbol{L}-\boldsymbol{\ell})\right] \tag{1.79}$$

to zero, we see that $G(\boldsymbol{\ell}, \boldsymbol{L}-\boldsymbol{\ell}) = \lambda(\boldsymbol{L}) f(\boldsymbol{\ell}, \boldsymbol{L}-\boldsymbol{\ell}) / \left[2C_\ell C_{|\boldsymbol{L}-\boldsymbol{\ell}|}\right]$ is a solution. The Lagrange multiplier $\lambda(\boldsymbol{L})$ is determined by substitution of this solution into $I[G] = 1$. This yields the final result given in Eq. 1.71–1.72. The minimum variance filters for all other quadratic pairs can be obtained in this manner.

Substituting Eq. 1.72 into Eq. 1.70 and simplifying gives

$$\frac{\hat{\phi}_{TT}(\boldsymbol{L})}{A_{TT}(\boldsymbol{L})} = \int d^2\boldsymbol{\ell} \left[\frac{\tilde{C}_\ell^{TT}(\boldsymbol{L}\cdot\boldsymbol{\ell})}{C_\ell^{TT}} T(\boldsymbol{\ell})\right] \left[\frac{1}{C_{|\boldsymbol{L}-\boldsymbol{\ell}|}^{TT}} T(\boldsymbol{L}-\boldsymbol{\ell})\right], \tag{1.80}$$

which is, manifestly, a convolution integral. Using the convolution theorem, Eq. 1.80 can be efficiently evaluated by computing the two-dimensional inverse Fourier transform of the square bracketed terms (giving filtered versions of the data $T(\mathbf{x})$), multiplying these in real space, and then Fourier transforming back to obtain $\hat{\phi}(\boldsymbol{L})/A(\boldsymbol{L})$.

This procedure applies analogously to all the lensing quadratic estimators (i.e., each pair of fields x, x'). We have implemented this *real-space reconstruction* approach in the ACTPol lensing pipeline to achieve a factor ≈ 100 speed up in the lensing reconstruction compared to evaluating the convolution (Eq. 1.80) directly. This is particularly valuable given that lensing power spectrum estimation requires $O(10^3)$ simulations for pipeline validation, covariance estimation and bias corrections. Rapid lensing reconstruction on the high-resolution, wide-field ACTPol data has

therefore become a computational necessity.

A similar trick can be performed for efficient evaluation of the normalisation terms $A_{xx'}$. This term is data independent and can therefore be pre-computed, prior to lensing reconstruction. Furthermore, the variance of each estimator, averaged over large-scale structure, is

$$\mathrm{Var}(\hat{\phi}_{xx'}(\boldsymbol{L}), \hat{\phi}_{xx'}(\boldsymbol{L}')) = \langle \hat{\phi}_{xx'}(\boldsymbol{L}) \hat{\phi}_{xx'}(\boldsymbol{L}') \rangle - \langle \phi(\boldsymbol{L}) \phi(\boldsymbol{L}') \rangle = \delta(\boldsymbol{L} + \boldsymbol{L}')(N_L^{(0)} + N_L^{(1)}), \quad (1.81)$$

where $N_L^{(0)} = A_{xx'}(L)$ is the dominant term, zeroth-order in the lensing power spectrum. The $N_L^{(1)}$ term, which acts as a bias in power spectrum estimation and also adds variance, is a sub-dominant contribution – first-order in the lensing power spectrum – arising from higher-order terms in the expansion of Eq. 1.60. The correlation between estimators for current experiments is less than $O(10\%)$, and they may therefore be combined as a *minimum-variance estimator* to improve the signal-to-noise on the reconstructed lensing modes (Hu & Okamoto, 2002). As polarisation measurements improve, the most powerful estimator will transition from TT to EB; since the unlensed B-mode signal is small ($r \ll 1$), the cosmic-variance contribution to the variance of the lensed B-modes is small.

$A_{xx'}(L)$ therefore sets the precision with which a particular mode $\phi(\boldsymbol{L})$ can be reconstructed, and in a related fashion is the dominant contribution to the noise variance of the estimated power spectrum (Kesden et al., 2003):

$$\mathrm{Var}(\hat{C}_L^{\phi\phi}) = \frac{1}{(2l+1) f_{\mathrm{sky}}} \left(C_L^{\phi\phi} + N_L^{(0)} + N_L^{(1)} \right)^2 . \quad (1.82)$$

Efficient computation of $A_{xx'}(L)$ is therefore desirable for real data analysis, but is also useful for computing the expected signal-to-noise of the lensing power spectrum for a given experimental specification, and thereby forecasting parameter constraints (Chapter 4). Inserting the expressions for the minimum-variance filters F into Eq. 1.71, the integral can again be written as a convolution. For example, for the TT estimator, Eq. 1.71 becomes

$$\begin{aligned} A_{TT}^{-1}(\boldsymbol{L}) &= \int d^2\boldsymbol{\ell} \, \frac{[\tilde{C}_\ell^{TT}(\boldsymbol{L}\cdot\boldsymbol{\ell}) + \tilde{C}_{|\boldsymbol{L}-\boldsymbol{\ell}|}^{TT}(\boldsymbol{L}\cdot[\boldsymbol{L}-\boldsymbol{\ell}])]^2}{2 C_\ell^{TT} C_{|\boldsymbol{L}-\boldsymbol{\ell}|}^{TT}} \\ &= \int d^2\boldsymbol{\ell} \, \frac{(\tilde{C}_\ell^{TT})^2 (\boldsymbol{L}\cdot\boldsymbol{\ell})^2}{C_\ell^{TT}} \frac{1}{C_{|\boldsymbol{L}-\boldsymbol{\ell}|}^{TT}} + \int d^2\boldsymbol{\ell} \, \frac{\tilde{C}_\ell^{TT}(\boldsymbol{L}\cdot\boldsymbol{\ell})}{C_\ell^{TT}} \frac{\tilde{C}_{|\boldsymbol{L}-\boldsymbol{\ell}|}^{TT}(\boldsymbol{L}\cdot[\boldsymbol{L}-\boldsymbol{\ell}])}{C_{|\boldsymbol{L}-\boldsymbol{\ell}|}^{TT}}, \end{aligned} \quad (1.83)$$

which can be evaluated efficiently by applying the convolution theorem separately to each term. The TE estimator is not separable for the minimum variance filter, but choosing a slightly sub-optimal filter leads to negligible loss of signal-to-noise and allows the same formalism to be used (Okamoto & Hu, 2003). In Chapter 5 we show simulated Advanced ACTPol convergence maps, given the expected noise properties, and also projected lensing reconstruction noise power spectra.

A weakness of lensing quadratic estimators is that they are sensitive to *any* effect which gives rise to correlations between modes. For example, windowing of observations (e.g., masking or finite-patch observations) causes mode coupling and gives rise to a stochastic *mean-field* contribution, even across an unlensed patch of sky. This must be computed (e.g., through representative Monte Carlo simulations), subtracted from the estimated lensing field, and accounted for in the bandpower uncertainties (e.g., *Planck* Collaboration, 2015e). Other effects include the non-Gaussian thermal Sunyaev-Zel'dovich effect and radio point-sources which are present, at some level, in CMB temperature and polarisation maps.

As such, lensing reconstruction methods beyond the naive quadratic estimator have been developed. These include maximum-likelihood approaches (e.g., Hirata & Seljak, 2003) and *bias-hardened* approaches which mitigate estimator bias from masking, inhomogenous noise or the use of an incorrect fiducial cosmology in the filters (Namikawa & Takahashi, 2014). More recently, the methodology for a fully Bayesian approach – allowing for the simultaneous sampling of the lensing field and the lensed temperature signal, given the observed, noisy T field – was introduced in Anderes et al. (2015), but has yet to be applied to real data.

On the observational side, CMB lensing measurements have advanced rapidly since the first detection in 2007, at 3.4σ significance, using WMAP data in cross-correlation with the clustering of radio galaxies from NVSS (Smith et al., 2007). Because CMB lensing traces the underlying matter field, it correlates with large-scale structure probes such as galaxy clustering, galaxy weak-lensing and the cosmic infrared background (Sec. 1.6; Hirata et al., 2008; Feng et al., 2012; Bleem et al., 2012; *Planck* Collaboration, 2014b; van Engelen et al., 2014; Fornengo et al., 2014; Bianchini et al., 2014; Hand et al., 2015; Allison et al., 2015). This connection allows CMB lensing to be used, not only for cosmological inferences, but also to provide information about astrophysical effects such as tracer bias and halo masses.

Gravitational lensing is sensitive to the amplitude and growth of structure, and is thereby

important for the determination of cosmological parameters such as σ_8, Ω_m and Σm_ν. Lensing measurements help to break the geometric degeneracy between the CMB peak positions and the distance to the last-scattering surface. ACT measurements of the lensing power spectrum have been used to tighten cosmological parameters beyond CMB temperature constraints, a process which is now routine in current analyses of CMB data (Sievers et al., 2013; Das et al., 2014; *Planck* Collaboration, 2015c).

Das et al. (2011) used ACT measurements of CMB temperature (across 324 deg^2 on the celestial equator) to make the first detection of the CMB lensing power spectrum, at 4σ significance. Furthermore, Sherwin et al. (2011) constructed a lensing likelihood from these data which, when combined with primary CMB measurements from *WMAP* and ACT demonstrated, for the first time, evidence for $\Omega_\Lambda > 0$ from CMB measurements alone (at 3.2σ significance).

In Sherwin et al. (2012), the projected positions of quasars were cross-correlated with ACT lensing maps and found to be consistent with ΛCDM expectations, with a non-zero amplitude at 3.8σ significance. This measurement was used to learn about the quasar bias and host halo mass. Hand et al. (2015) were the first to measure the cross-correlation between CMB and galaxy lensing, using ACT data and lensed galaxies from the CS82 survey (Erben et al., 2013).

The cosmic infrared background, produced by thermal emission from unresolved dusty galaxies, is substantially correlated with the CMB (up to 80%; *Planck* Collaboration, 2014b). Data from ACTPol were used to construct lensing maps which were cross-correlated with the *Planck* 545 GHz HFI map, an all-sky dust tracer map. The principal result was a 9.1σ detection of CMB lensing using temperature and polarisation measurements (4.5σ from polarisation alone), for the first time including scales $\ell > 2000$ for this measurement (van Engelen et al., 2014). Allison et al. (2015) cross-correlated lensing maps from ACT and ACTPol with the clustering of active galactic nuclei (AGN) in the FIRST radio survey, finding a non-null signal at 4.4σ (Chapter 3; Helfand et al., 2015).

In other work, recent ACTPol lensing measurements were stacked at the locations of optically-selected galaxies from the BOSS survey (Eisenstein et al., 2011; Madhavacheril et al., 2014). This represented the first evidence, at 3.2σ significance, for lensing of the CMB by individual dark matter halos (the '1-halo' term). This analysis has since been repeated with *Planck* and SPT data (Baxter et al., 2015; *Planck* Collaboration, 2015g).

We note that measurements by the South Pole Telescope (SPT) and the *Planck* satellite have also been used extensively for lensing science (van Engelen et al., 2012; Bleem et al., 2012; Geach et al., 2013; *Planck* Collaboration, 2015c,e; Baxter et al., 2015; Giannantonio et al., 2015). For instance, the lensing auto-spectrum is detected at 14σ significance (5.8σ from polarisation alone) using 100 deg^2 of SPTpol data in Story et al. (2015). These measurements have evolved in tandem with those from ACT, developing the field as a whole. The current state-of-the-art lensing map comes from the *Planck* 2015 data release (*Planck* Collaboration, 2015e), representing a 40σ detection of the gravitational lensing of the CMB across the full sky.

1.5 Galaxy bias

Luminous matter in the universe (e.g., galaxies and galaxy clusters) broadly traces the underlying total matter field; at early times the dark matter perturbations set up gravitationally-dominant potentials into which the baryonic matter falls after decoupling (Sec. 1.2). Over time the matter continues to collapse under gravity until – in some regions of the universe where the density is sufficiently high – non-linear growth and *virialisation* proceeds, allowing stars and galaxies to form. Crucially, this process will only occur in those parts of the universe which are initially dense enough, i.e., at the peaks in the density field.

As such, luminous matter does not *exactly* trace the underlying matter field: there are more luminous objects in a high-density region than in a low-density region than one would expect from simply taking the ratio of the matter density of those two regions. Generically, the density of luminous matter (baryons b) is some function of the underlying matter field: $\delta_b = f(\delta_m)$. Kaiser (1984) showed that, due to this effect, galaxy clusters are expected to be linearly *biased* tracers of the underlying matter field. Bardeen et al. (1986) showed more generally that for Gaussian random fields the peaks are biased with respect to the underlying field, while Mo & White (1996) used extended Press-Schechter theory to derive an analytical expectation for the bias function $b(M, R)$ which relates the overdensity of dark matter haloes δ_h of a given mass M, smoothed over a given radius R, to the underlying matter overdensity:

$$\delta_h(M, R) = b(M, R)\delta_m. \tag{1.84}$$

Since the luminous matter is seeded in dark matter haloes, it is correspondingly biased. As is shown in Chapter 3, radio jets from active galactic nuclei are highly-biased tracers of the dark matter, and therefore exist in extremely massive dark matter haloes of 10^{13}–$10^{14} M_{\odot}$ (Allison et al., 2015).

1.6 Cross-correlations

Many cosmological observables can be related to the underlying matter field; they are said to *trace* the matter fluctuations. This is the case for biased tracers such as galaxies and clusters (Sec. 1.5), and also for unbiased tracers such as galaxy and CMB lensing. We actually observe fields as they are projected onto the sky, and the exact form of the projection kernel $W_i(\chi)$ will vary between tracers (Eq. 1.17). These different cosmological probes are therefore expected to be *correlated*, to a varying extent, depending on their redshift dependence, bias evolution and geometrical properties. Cross-correlations between CMB lensing and other tracers of large-scale structure have been reported extensively in recent years (e.g., Smith et al., 2007; Hirata et al., 2008; Feng et al., 2012; Bleem et al., 2012; *Planck* Collaboration, 2014b; van Engelen et al., 2014; Fornengo et al., 2014; Bianchini et al., 2014).

The scale-dependent correlation between two matter tracers i, j is captured in their *cross-spectrum* C_ℓ^{ij}, which is given by Eq. 1.20, Sec. 1.1. The cross-spectrum is in essence a weighted integral of the matter power spectrum over redshift. For any given cosmological model the theoretical cross-correlation between two probes can – in principle – be calculated, and then compared against observations to test for consistency or to determine the parameters of that theory. When estimating from data, an advantage of cross-correlations over auto-correlations is that they are robust to systematic biases which may be particular to each dataset; a spurious instrumental or methodological signal in one probe will not correlate with the cosmological signal in the other. Furthermore, cross-correlations carry cosmological information over and above the auto-spectra alone: correlated Gaussian fields are *fully* described by both their auto- and cross-spectra (e.g., Das et al., 2013). CMB lensing cross-correlations are discussed further, and calculated in detail, in Chapters 3 and 5.

1.7 Bayesian statistics

Within cosmology we are interested in drawing conclusions about the models, laws and quantities which describe the universe. These conclusions are based on empirical data and, ultimately, our prior state of knowledge and the assumptions we are willing to assert at the beginning of the analysis. Cox (1946) showed that consistent reasoning (e.g., demanding *transitivity* between degrees of belief) requires the use of probability theory to describe the degree of belief that we assign to a given proposition. Hence, the application of probability theory is the correct way to reason scientifically in the presence of epistemological uncertainties (Hennig et al., 2015). For instance, we are able to say that the probability the age of the universe is between 13.8 and 13.9 billion years old is around 65%, given current data and a ΛCDM cosmological model (*Planck* Collaboration, 2015c).

Cox proved that consistent reasoning requires that the probability we assign the joint proposition A *and* B be the same as the product of the probability of A *given* B and the probability of B alone, viz., $\mathbb{P}(A, B) = \mathbb{P}(A|B)\mathbb{P}(B)$ (Sivia, 1996). A corollary of this result is *Bayes' theorem*:

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(B|A)\mathbb{P}(A)}{\mathbb{P}(B)}. \tag{1.85}$$

In Chapter 2 we discuss the application of Bayes' theorem to parameter inference and model selection in cosmology. An additional important result is the *marginalisation* property:

$$\mathbb{P}(A) = \int \mathbb{P}(A|B)\mathbb{P}(B)dB, \tag{1.86}$$

where we abuse notation slightly, now considering each $\mathbb{P}$ as a probability density. This property allows us to deal with unknown *nuisance* parameters: we can integrate them out, conditioning our inference all on plausible values, weighting by their prior probability. For example, we may wish to infer cosmological parameters θ from a CMB map, but the precise calibration (i.e. overall normalisation) y of the map is not known exactly. However, for any given y, the empirical power spectrum can be calculated and parameter inference performed by comparing this to theoretical power spectra to give the conditional posterior density $\mathbb{P}(\theta|y)$. In principle, then, we simply perform parameter estimation for all values of y, and then combine them (integrate), weighting

by the prior probability $\mathbb{P}(y)$ to give the sought-after, marginalised cosmological parameter density $\mathbb{P}(\theta)$. In practice, it may not be computationally feasible to condition on all values of the nuisance parameters: intelligent sampling techniques can be employed to perform this marginalisation automatically, rapidly and accurately (Chapter 2).

As a quantitative example of the Bayesian approach, we outline how parameter estimation is performed in the context of the *general linear model* M, in which the data $\mathbf{d}$ are modelled as a linear combination of parameters $\boldsymbol{\theta}$ with additive (and, in general, correlated) Gaussian noise. Then

$$\mathbf{d} = \mathbf{X}\boldsymbol{\theta} + \mathbf{n}, \tag{1.87}$$

where $\mathbf{n} \sim N(0, \mathbb{N})$ for noise covariance $\mathbb{N}$ and the matrix $\mathbf{X}$ is often called the *design matrix* in the statistics literature. Given the assumptions of the model, the sampling distribution for the data is given by

$$p(\mathbf{d}|\boldsymbol{\theta}, M) = \frac{1}{\sqrt{\det(2\pi\mathbb{N})}} \exp\left[-\frac{1}{2}(\mathbf{d} - \mathbf{X}\boldsymbol{\theta})^T\mathbb{N}^{-1}(\mathbf{d} - \mathbf{X}\boldsymbol{\theta})\right]. \tag{1.88}$$

From Bayes' theorem (Eq. 1.85) and some matrix algebra the posterior distribution for the parameters may then be written

$$p(\boldsymbol{\theta}|\mathbf{d}, M) = \frac{1}{\sqrt{\det(2\pi\mathbb{C})}} \exp\left[-\frac{1}{2}(\boldsymbol{\theta} - \bar{\boldsymbol{\theta}})^T\mathbb{C}^{-1}(\boldsymbol{\theta} - \bar{\boldsymbol{\theta}})\right], \tag{1.89}$$

where the posterior covariance $\mathbb{C} = (\mathbf{X}^T\mathbb{N}^{-1}\mathbf{X})^{-1}$ and mean $\bar{\boldsymbol{\theta}} = (\mathbf{X}^T\mathbb{N}^{-1}\mathbf{X})^{-1}\mathbf{X}^T\mathbb{N}^{-1}\mathbf{d}$, and we have assumed uninformative priors on the parameters $\boldsymbol{\theta}$. Because the generative data model is linear in parameters, and the noise is Gaussian distributed, the posterior is also Gaussian.

Specialising to the scenario where the data signal is some scaled version of a template $\mathbf{m}$, with amplitude A (i.e., $\boldsymbol{\theta} = A$ and $\mathbf{X} = \mathbf{m}$) we can summarise the Gaussian posterior for A by

$$A = \bar{A} \pm \sigma(A) = \frac{\mathbf{m}^T\mathbb{N}^{-1}\mathbf{d}}{\mathbf{m}^T\mathbb{N}^{-1}\mathbf{m}} \pm \frac{1}{\sqrt{\mathbf{m}^T\mathbb{N}^{-1}\mathbf{m}}}. \tag{1.90}$$

The signal-to-noise ratio SNR of the detection of a non-zero amplitude is then

$$SNR \equiv \frac{\bar{A}}{\sigma(A)} = \frac{\mathbf{m}^T\mathbb{N}^{-1}\mathbf{d}}{\sqrt{\mathbf{m}^T\mathbb{N}^{-1}\mathbf{m}}}. \tag{1.91}$$

Finally, we can straightforwardly forecast the expected SNR by setting $\mathbf{d} = \mathbf{m}$ ($A = 1$); further assuming uncorrelated uncertainties, so that $\mathbb{N}_{ij} = \sigma_i^2 \delta_{ij}$, Eq. 1.91 reduces to

$$SNR = \sqrt{\sum_i \frac{m_i^2}{\sigma_i^2}}. \tag{1.92}$$

This expression is applied in Chapter 5 for forecasting the detection significance of future weak-lensing cross-correlations.

1.8 Forecasting

A crucial aspect of deciding the observational strategy of a telescope is forecasting the expected science output of any given choice. For instance, one might want to forecast how well the data will constrain a certain parameter, how well the data will be able to discriminate between theory A or theory B, or how much more information will be learnt about parameter θ with observational strategy X or Y.

Forecasting can answer questions such as 'deep or wide?'; whether to put integration time on a small region of the sky, probing to higher redshift, or whether observe a wide sky area, increasing the number of angular modes over which to average to beat down sample variance[3]. Different measures of the *utility* of an experiment will generically lead to a different hierarchy of strategies. For example, inflationary gravitational waves, if present, would give an extremely faint signal in the large-scale B-mode power spectrum. Any observations designed to detect the corresponding tensor-to-scalar ratio r must produce high-sensitivity polarisation maps, and with a finite total integration time on a given telescope, this requirement would typically favour deep observations over a small region. There is commonly a trade-off between reducing sample variance by observing a greater number of independent angular modes (by going wider) and reducing the noise power in a given direction (by going deeper).

Forecasting, by modelling the data generation process, is therefore integral to experimental design. A decision theory approach can be adopted to select between various options through the

[3]For example, for a CMB experiment, the detector noise gives a map sensitivity in temperature $\Delta T = NET\sqrt{\frac{4\pi f_{\rm sky}}{N_{\rm det} Y \tau}}$, where NET is the noise-equivalent temperature of a single detector, $N_{\rm det}$ is the number of detectors, Y is their yield and τ the integration time (Wu et al., 2014). Hence, for a fixed instrument and integration time, $\Delta T \sim \sqrt{f_{\rm sky}}$: a wider patch is noisier, but this must be balanced against the increase in the number of angular modes accessed ($N_{\rm modes} \sim f_{\rm sky}$).

definition of a utility function $U(D|i)$ – a function of the (*a priori* unknown) data D generated by experiment i – which might account for science output, costs and so forth (Loredo, 2004). The expected utility $E(i)$, averaging over plausible data realisations, is then a natural scoring criterion:

$$E(i) \equiv \mathbb{E}_{p(D|i)}[U(D|i)] = \int dD\ p(D|i)U(D|i). \tag{1.93}$$

Note that the chosen theoretical model is implicit within the experimental context i. As a concrete example, one might use the *Kullback-Leibler (KL) divergence* between posterior and prior – which quantifies the *information* learned about the parameters on obtaining data D – as a utility function:

$$U(D|i) \leftarrow \mathcal{D}_{\mathrm{KL}}[p(\theta|D,i)\ ||\ p(\theta|i)] = \int d\theta\ p(\theta|D,i) \ln \frac{p(\theta|D,i)}{p(\theta|i)}, \tag{1.94}$$

a choice advocated in Loredo (2004) to optimise the choice of sample location for some toy model problems. Inserting Eq. 1.94 into Eq. 1.93 and simplifying gives the expected utility under the choice of a KL utility function:

$$E_{\mathrm{KL}}(i) = \int d\theta\ p(\theta|i) \left[\int dD\ p(D|\theta,i) \ln p(D|\theta,i) \right] - \int dD\ p(D|i) \ln p(D|i), \tag{1.95}$$

which has a physical interpretation: the expected information gain is equal to the expected (prior averaged) information in the *sampling distribution* $p(D|\theta,i)$, minus the information in the *prior predictive* distribution $p(D|i)$. The task of the scientist is, then, to compute Eq. 1.95 for all plausible strategies i, which necessarily means being able to compute the prior $p(\theta|i)$ and sampling distribution $p(D|\theta,i)$ across the full parameter and data domain. This clearly requires a comprehensive understanding of the data generation process; both the production of the signal – given the assumed theoretical model – and also how this propagates through the experiment (e.g., becoming corrupted by noise) to produce the data. In many real world situations an analytical or numerical evaluation of Eq. 1.95 is impossible, and approximations must be found to make progress.

In a Bayesian context – where we condition on known quantities and marginalise over unknown quantities – forecasting is interesting because not only is the *data* realisation unknown but often various model parameters are also not known with infinite precision *a priori*. Different underlying

parameters will generically give rise to different datasets, even within a given experiment i. In general the inferred posterior variance depends on the particular data realisation. The forecasting approach described above naturally encapsulates prior uncertainty on both data and parameters.

Typically we are interested in the uncertainty with which an experiment, or experimental strategy, is going to constrain a given parameter θ, for example, the sum of the neutrino masses Σm_ν (Chapter 5). In this scenario the posterior variance, given the data D, would be the utility function and we might then choose the strategy which is expected, on average, to give the smallest posterior variance on θ.

In many scenarios within cosmology the underlying model parameters are sufficiently well known that any scatter in the plausible future data, as quantified by the width of the prior predictive $p(D|i)$, is primarily due to the experimental uncertainties rather than lack of *a priori* knowledge of the parameters. In this case, because all allowed parameters will give rise to essentially the same datasets, it is justified to pick *fiducial parameters* θ – parameters which are representative of the current state of knowledge – and simply set $p(D|i) = p(D|\theta, i)$. This is the sampling distribution of the data D given the fiducial parameters θ, and is often more straightforward to handle than the full prior predictive $p(D|i)$.

A further approximation – when forecasting the posterior variance – is to assume that the posterior distribution is *Gaussian* (an increasingly good approximation as the data improve). Gaussianity of the posterior $p(\theta|D, i)$ ensures that the we can write the inverse posterior variance (or *precision*), which is in general a function of the data, as

$$\sigma^{-2}(D) = -\frac{\partial^2 \ln p(\theta|D, i)}{\partial \theta^2}, \tag{1.96}$$

where the derivative on the right-hand side can be evaluated at any θ. The expected posterior precision then reduces simply to the *Fisher scalar* F:

$$F(\theta|i) = \mathbb{E}_{p(D|\theta,i)}\left[-\frac{\partial^2 \ln p(\theta|D, i)}{\partial \theta^2}\right]. \tag{1.97}$$

This expression can be generalised to a Fisher *matrix*, which quantifies the expected posterior precision matrix for a multivariate Gaussian posterior; this generalisation is given, and used extensively, in Chapter 5. The inverse Fisher matrix is then a measure of the projected parameter

covariance for the future experiment i.

We emphasise the limitation of this forecasting approach lies in the assumption of Gaussianity of the posterior and in not properly marginalising over prior uncertainties on the parameters. A Bayesian version of the *Cramér-Rao bound*, attributed to Van Trees (1968), implies that the expected posterior variance is bounded from below by the inverse of the prior-averaged Fisher matrix (Rao, 1945; Gill & Levit, 1995), viz.,

$$\int dDp(D|i) \int d\boldsymbol{\theta} p(\boldsymbol{\theta}|D,i) \left[\boldsymbol{\theta} - \int d\boldsymbol{\theta}' p(\boldsymbol{\theta}'|D,i)\boldsymbol{\theta}'\right] \left[\boldsymbol{\theta} - \int d\boldsymbol{\theta}' p(\boldsymbol{\theta}'|D,i)\boldsymbol{\theta}'\right]^T \geq \left[\int d\boldsymbol{\theta} p(\boldsymbol{\theta}|i) F(\boldsymbol{\theta})\right]^{-1}. \tag{1.98}$$

If the posterior is Gaussian with data-independent covariance, the bound is *saturated* (i.e., equality is satisfied), and the inverse Fisher matrix – evaluated at any fiducial parameters – is identically equal to the posterior covariance.

1.8.1 Transformation of basis for covariance matrix evaluation

The Fisher matrix, and corresponding covariance matrix, are defined in a given parameter basis. Suppose we have a covariance matrix $\mathbb{C}_{\mathbf{x}}$ for N parameters $\mathbf{x}$ and we want to know the corresponding covariance matrix in a different parameter basis $\mathbf{f} = \mathbf{f}(\mathbf{x})$. Provided the data are sufficiently constraining, such that the variations in $\mathbf{f}$ corresponding to variations in $\mathbf{x}$ allowed by $\mathbb{C}_{\mathbf{x}}$ are effectively linear, then we can write

$$\mathbb{C}_{\mathbf{f}} = \mathbb{J}\mathbb{C}_{\mathbf{x}}\mathbb{J}^T \tag{1.99}$$

as the covariance matrix in the new parameter bias, where $\mathbb{J} \equiv \partial\mathbf{f}/\partial\mathbf{x}$ is the Jacobian of the transformation. This is utilised in Chapter 4 to change basis, eliminating the Hubble parameter h (which is an input for the CAMB code; Lewis et al., 2000) in favour of the angular diameter distance to recombination θ. In this special case,

$$\mathbb{C}_{\mathbf{x}} = \begin{pmatrix} \mathbb{C}_{hh} & \mathbb{C}_{hi} \\ \mathbb{C}_{ih} & \mathbb{C}_{ij} \end{pmatrix}, \tag{1.100}$$

where i, j label parameters in the original basis other than h, e.g., $\Omega_{\rm c}h^2, \Omega_{\rm b}h^2, \ldots$. The sub-matrices, from left to right, top to bottom, have dimensions 1×1, $1 \times (N-1)$, $(N-1) \times 1$ and $(N-1) \times (N-1)$, respectively. Eliminating h in favour of θ using Eq. 1.99 yields the transformed covariance matrix

$$\mathbb{C}_{\mathbf{f}} = \begin{pmatrix} \frac{\partial\theta}{\partial\mathbf{x}^T}\mathbb{C}_{\mathbf{x}}\frac{\partial\theta}{\partial\mathbf{x}} & \frac{\partial\theta}{\partial\mathbf{x}^T}\mathbb{C}_{\mathbf{x}i} \\ \mathbb{C}_{i\mathbf{x}}\frac{\partial\theta}{\partial\mathbf{x}} & \mathbb{C}_{ij} \end{pmatrix}, \tag{1.101}$$

which, as expected, leaves the $\mathbb{C}_{ij}$ sub-matrix unchanged. To compute this matrix in practice requires analytical or numerical evaluation of the derivatives $\partial\theta/\partial\mathbf{x}$.

1.9 Experiments

In this section we summarise the principal experiments for current and near-future CMB and large-scale structure studies which are of relevance to this thesis. We refer to these experiments throughout the text. As the focus of this thesis, we begin with the ACTPol and Advanced ACTPol experiments.

- **ACTPol**: ACT is a ground-based 6m telescope in Northern Chile. ACT began science observations in 2008 with temperature-sensitivity only – the *MBAC* camera – and upgraded to ACTPol for temperature and polarisation observations from 2013-16 (Swetz et al., 2008; Niemack et al., 2010). The currently operating frequency channels are centred at 90GHz and 150GHz, with some of the detector arrays equipped with a rotating half-wave plate to improve polarisation sensitivity (Henderson et al., 2015). Its large dish and extensive detector arrays ($O(10^3)$ detectors) mean ACTPol, alongside SPTpol, is a leading high-resolution, small-scale CMB experiment. Cosmology results from ACTPol to date include power spectra and parameters from the first season of data (Naess et al., 2014), lensing from dark matter haloes (Madhavacheril et al., 2014), and lensing cross-correlations with the cosmic infrared background and radio-loud AGN (van Engelen et al., 2014; Allison et al., 2015). Parameters and a full lensing power spectrum analysis from the first two seasons of data are upcoming (Louis et al. *in prep*, Sherwin et al. *in prep*). In early 2016 ACTPol will be upgraded to Advanced ACTPol.

- **Advanced ACTPol**: This camera upgrade from the current ACTPol instrument will observe from 2016-19 across five frequency bands (28GHz–230GHz), with roughly double the number of detectors per array compared to ACTPol (Henderson et al., 2015). Advanced ACTPol (AdvACT) will be fitted with rotating half-wave plates which will improve access to the large-angular-scale polarisation signal by suppressing temperature-to-polarisation leakage from atmospheric variations, benefitting primordial B-mode searches. The wide frequency coverage and small beams (*cf.* e.g., *Planck*) will give exceptional foreground removal power, mitigating the effects of dust and synchrotron contamination, and helping in the search for r. AdvACT is scheduled to cover half the sky down to 8 μK-arcmin sensitivity in temperature ($\approx 11\mu$K-arcmin in polarisation); much wider – but shallower – than SPT-3G. This larger area will allow for more angular modes to be observed, which will help for lensing reconstruction and constraining the sum of the neutrino masses (Chapter 5).

1.9.1 Other cosmic microwave background experiments

- **SPTpol**: The other principal small-scale CMB experiment, this 'Stage 2' telescope will be upgraded to SPT-3G in 2016 (Benson et al., 2014). With observing frequencies of 90GHz and 150GHz and a 10m dish, the full SPTpol survey – covering 500 deg^2 – is expected to have a temperature sensitivity of 6 μK-arcmin, with observations taken from 2012-15 and the analysis ongoing. Cosmology highlights to date include a direct measurement of the lensing B-mode power spectrum from 100 deg^2 of data (Keisler et al., 2015), and a 14σ detection of the CMB lensing auto-spectrum (Story et al., 2015).

- ***Planck***: This European Space Agency (*ESA*) satellite is the premier CMB telescope, providing the most sensitive full-sky CMB maps and strongest cosmological parameter constraints to date (*Planck* Collaboration, 2015a,c). *Planck* observed the full sky repeatedly from 2009–2013 across a wide range of frequency channels (30GHz–857GHz), giving (presently) unrivalled foreground handling ability. We show the all-sky temperature map as measured by *Planck* in Fig. 1.15, and the corresponding TT power spectrum in Fig. 1.12. The *Planck* satellite is the scientific successor to the NASA Wilkinson Microwave Anisotropy Probe (*WMAP*).

- **POLARBEAR and the Simons Array**: observing since 2012, the POLARBEAR instrument is a CMB experiment on the 3.5m Huan Tran single-dish telescope in the Atacama Desert (The POLARBEAR Collaboration, 2010). Science results include the direct and indirect measurement of CMB polarisation lensing, and lensing B-modes, from the first season 150GHz data (The POLARBEAR Collaboration, 2013, 2014a,b). Data from seasons 2 and 3 are being analysed. Two additional dishes and more frequency coverage will create the Simons Array; fitted with POLARBEAR-2 receivers, it will be fully deployed in 2017 (Tomaru et al., 2012)[4].

- **SPT-3G**: A 'Stage 3' telescope, SPT-3G is the South Pole equivalent of AdvACT and its full 2500 deg^2 survey will attain a map sensitivity of around 3.5μK-arcmin in polarisation (Benson et al., 2014). In this noise regime the EB lensing estimator will begin to carry more information about the convergence field than the TT estimator, providing scope for extensive lensing science and cross-correlations with large-scale structure surveys.

- **SPIDER**: Balloon-borne from the South Pole, this polarimeter is projected to place an upper limit on the tensor-to-scalar ratio of $r < 0.03$ (at a 99% CL) – although this was forecast before the *Planck* polarised dust results were released, so is likely to be optimistic (Fraisse et al., 2013). Its first flight took place in January 2015, with the data now being analysed[5].

- There are many other current and planned CMB experiments, such as **CLASS**[6], **BICEP-Keck**[7] and **QUIJOTE**[8], but we focus in this thesis on those presented above (Essinger-Hileman et al., 2014; BICEP2 Collaboration et al., 2014; López-Caniego et al., 2014).

1.9.2 Large-scale structure experiments

- **Dark Energy Survey**: DES began in August 2013 and will take five years of science observations. The Dark Energy Camera (DECam), the receiver taking data for the survey,

[4] http://ltd16.grenoble.cnrs.fr/IMG/UserFiles/Images/11_2015_07_23_LTD16_PBSA_Suzuki_Final.compressed.pdf
[5] http://spider.princeton.edu
[6] http://sites.krieger.jhu.edu/class/
[7] http://bicepkeck.org
[8] http://www.iac.es/proyecto/cmb/pages/en/quijote-cmb-experiment.php

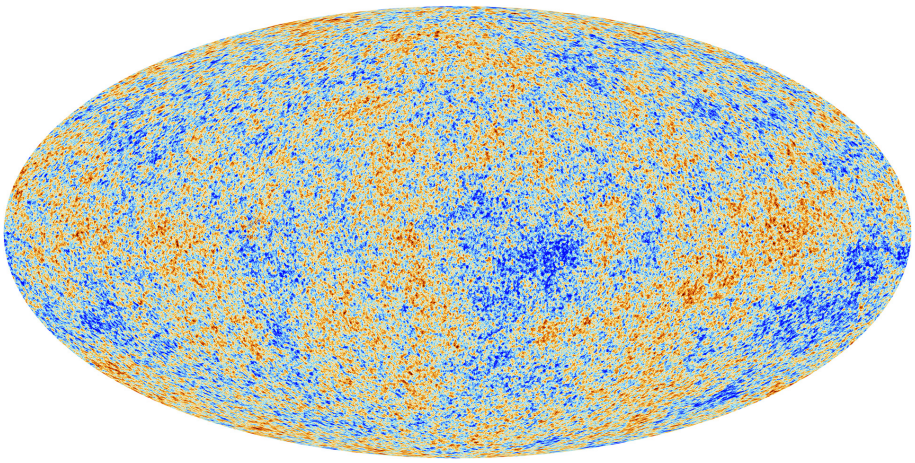

Figure 1.15: *Source: European Space Agency. Planck*-measured CMB temperature all-sky map.

is hosted on the Blanco telescope in the Chilean Andes. The major observables from DES will be clustering, galaxy weak-lensing and supernovae measurements, with full-survey coverage planned across 4000 deg^2. The main science goal is to constrain the properties of dark energy (Diehl et al., 2014). Cosmology results from their 'science verification' data (pre-Year 1) include parameter constraints from cosmic shear tomography and cross-correlations of lensing and clustering measurements with *Planck* and SPT (The Dark Energy Survey Collaboration, 2015; Kirk et al., 2015a; Giannantonio et al., 2016).

- **Kilo-Degree Survey**: KiDS is an optical imaging survey covering 1500 deg^2 with ongoing observations. The coverage – in four filters (u, g, r, i) – is split evenly between KiDS-N on the celestial equator and KiDS-S in the Southern sky, and uses the VLT Survey Telescope in the Atacama desert. There is extensive overlap between the KiDS-N field and the ACTPol 'BOSS-N' field. This is due to be exploited for future lensing cross-correlation science.

- **Hyper-Suprime Cam**: HSC is located on the Subaru telescope and is an ongoing galaxy lensing survey. Due to the overlapping fields of ACTPol, Advanced ACTPol and HSC, there will be a strong case for cross-correlation science, particularly given the depth of HSC; the mean galaxy density is projected to be 20 galaxies arcmin^{-2}, compared to ≈ 10 arcmin^{-2} for full-depth KiDS and DES (Liu & Hill, 2015).

- **Extended Baryon Oscillation Spectroscopic Survey**: eBOSS is the successor to the SDSS BOSS survey which will measure the positions and spectroscopy of $\approx 6 \times 10^5$ quasars

and $\approx 4 \times 10^5$ emission line and luminous red galaxies. These will be used to reconstruct the BAO distance ratio over the redshift range $0.6 < z < 2.2$ (Zhao et al., 2015). The survey began in 2014 and will continue until 2020. It is the precursor to DESI and, in combination with CMB experiments such as *Planck* and AdvACT, will tightly constrain the sum of the neutrino masses.

- **Dark Energy Spectroscopic Instrument**: DESI is the successor to eBOSS and will measure the BAO distance ratio between $0.15 < z < 1.85$ to unprecedented precision: $\lesssim 1\%$ in bins of $\Delta z = 0.1$ (Levi et al., 2013). The DESI survey is due to begin in 2018. This will allow for tight constraints on dark energy and the sum of the neutrino masses Σm_ν. For example, combining AdvACT with DESI is expected to give a 1σ constraint of $\sigma(\Sigma m_\nu) = 22$ meV in the normal hierarchy (Allison et al., 2015).

- **Large Synoptic Survey Telescope**: LSST is under construction in Chile as an 8m optical telescope, with science observations due to begin in 2021. LSST will cover the entire available sky every few nights. With deep and wide weak-lensing and photometric catalogues, LSST will have extensive synergies with AdvACT. Both experiments will cover the same $\approx 50\%$ sky footprint (LSST Dark Energy Science Collaboration, 2012). The LSST weak-lensing 'gold-sample' will contain around 26 galaxies arcmin^{-2} and the AdvACT lensing – LSST lensing correlation will be detected with a signal-to-noise of $\gtrsim 100$, allowing for precision science with this cosmological observable (Chapter 5).

- **Euclid**: Euclid is an *ESA* space-based optical weak-lensing survey telescope, due for launch in 2020. Like LSST, it will also cover approximately half the sky and is designed to test gravity, and to constrain dark energy, neutrino mass and primordial non-Gaussianity through precision weak-lensing measurements (Laureijs et al., 2011).

- **SKA**: The Square Kilometre Array will be built in Australia and South Africa and is due for first light in 2020. Among many other science cases, the sensitivity to the neutral hydrogen 21cm line will allow it to detect $O(10^9)$ galaxies and to do all-sky intensity mapping, providing unprecedented scope for large-scale structure studies at radio wavelengths (Maartens et al., 2015). The SKA will also be a precision probe of reionisation.

1.10 Key questions in cosmology

Here we discuss several unsolved questions which are presently active in cosmology research and relevant to the work presented in this thesis.

Masses of the active neutrinos. There are three known active neutrino species; these are subject to the weak nuclear force, interacting as the three weak eigenstates of the electron, muon and tau neutrinos. Neutrinos propagate as mass eigenstates which are a linear combination of the weak eigenstates: this phenomenon gives rise to neutrino *oscillations*. Because these oscillations are observed, we know the neutrinos must have mass. However, oscillation experiments are only informative about the squared-mass *differences* of the neutrinos; the absolute mass scale of the neutrinos remains unknown. In Chapter 5 we discuss current mass limits and how cosmological observations are expected to measure the neutrino mass sum over the next decade.

The nature of dark energy and dark matter. Dark energy and dark matter represent the majority of the energy density budget of the universe today; around 70% and 25%, respectively. Their presence is strongly favoured by observations; for instance, without dark matter there would not be enough matter density to explain the extent of structure formation we actually measure in the universe. A particle candidate for dark matter has not been discovered, although direct and indirect searches are ongoing. Dark energy is a deeper puzzle; a possible contribution to the energy density from quantum mechanical vacuum fluctuations of a scalar field would be some 120 orders-of-magnitude too high to be the cosmological constant. This is the well-known *cosmological constant problem* (e.g., Burgess, 2013). Contemporary experiments will tightly constrain the equation of state of dark energy $w(z)$ through clustering and weak lensing measurements (Laureijs et al., 2011; LSST Dark Energy Science Collaboration, 2012; Levi et al., 2013). Any statistically significant deviation from $w = -1$ would revolutionise the field and the exact form of this 'non-null' signal would focus present theoretical studies of dark energy. The current constraint on a constant dark energy equation of state w from a combination of *Planck*, BAO, SNe and H_0 data is $w = -1.023^{+0.091}_{-0.096}$ at 95%, entirely consistent with a cosmological constant (*Planck* Collaboration, 2015c).

Theory of gravity. The known issue of the incompatibility of GR with quantum field theory in itself means there must be more to learn; progress on this problem may shed light on the

cosmological constant problem and the accelerating expansion of the universe. String theory provides a way to reconcile quantum theory with gravity, but its predictions are as yet untestable, making it difficult to adopt as a scientific theory on philosophical grounds. Some theorists search to explain the observed cosmic acceleration not through the introduction of a cosmological constant within ΛCDM, but instead through the introduction of an alternative theory of gravity. Solar system constraints and cosmological constraints from the CMB and large-scale structure tightly constrain the allowed parameter space of alternative theories, with current observations generally favouring GR (Clifton et al., 2012; Khoury, 2013). Alternative theories often introduce additional degrees of freedom. They must, correspondingly, explain the observational data much better than GR in order to overcome the increased Occam factor from the expansion of the plausible prior volume (Chapter 2).

Tension in growth parameters. The ΛCDM parameters relating to the growth of structure inferred from *Planck* observations of the CMB are in some tension with those derived from cosmic shear data from CFHTLenS, with the *Planck*-analysis generally favouring larger values of $\Omega_{\rm m}$ and a different σ_8 (*Planck* Collaboration, 2015c; MacCrann et al., 2015; Kitching et al., 2016). Unfortunately, this discrepancy cannot be explained through a large neutrino mass; this would push down H_0, exacerbating the inconsistency of the two datasets. Relating to the point above, this tension could be hinting at the inability of the ΛCDM model to capture the full growth history of the universe, i.e., the growth predicted by early-time measurements (the CMB) cannot be extrapolated to the late-time growth probed by lensing. Perhaps a modified theory of gravity is required. On the other hand, given the statistical precision of current observations, systematic uncertainties are non-negligible, and it is possible that residual systematics, such as in the shear estimation, could explain this discrepancy without the need to invoke modified gravity. This example relates to the wider question of how robust ΛCDM is, in its explanatory power, over a range of cosmological observations. Progress in answering this question will be made with upcoming observations, e.g., through neutrino mass detection or a measurement of $w \neq -1$.

Inflation and gravitational waves. Despite its many explanatory features – a solution to the horizon and flatness problems, and a seed of initial Gaussian fluctuations – inflation is not a consensus theory (see, e.g., Brandenberger, 2011, for some alternatives). Indeed, the model of inflation is really a collection of different theories and physical mechanisms, even within the slow-

roll approximation (Martin et al., 2014). Some authors have highlighted the lack of predictive power as an issue for the inflationary paradigm (Barrow & Liddle, 1997; Gubitosi et al., 2015). However, there is no compelling alternative to inflation which cosmologists currently agree upon, and therefore – from a Bayesian perspective – we should continue to use inflationary theory until observations rule otherwise. Inflationary theories generically produce gravitational waves which may be detectable through observations of the early universe (e.g., large-scale CMB B-mode observations); this is typically quantified through a measurement of the tensor-to-scalar ratio r which, as noted in Sec. 1.3.3, is presently consistent with zero given current data. With the recent first direct detection of gravitational waves by the LIGO experiment – demonstrating that gravitational waves can propagate over cosmological distances – the interest in the detection of primordial gravitational waves from inflation will intensify (Abbott et al., 2016).

Chapter 2

Comparison of sampling methods for Bayesian parameter estimation

In this chapter we present a comparison of sampling algorithms for Bayesian parameter estimation. These provide the main tools needed to extract cosmological information from observed statistics such as angular power spectra. The posterior probability distribution for a set of model parameters encodes all that the data have to tell us in the context of a given model; it is the fundamental quantity for Bayesian parameter estimation. In order to infer the posterior probability distribution we have to decide how to explore parameter space. Here we compare three prescriptions for how parameter space is navigated, discussing their relative merits. We consider Metropolis-Hasting sampling, nested sampling and affine-invariant ensemble MCMC sampling. We focus on their performance on toy-model Gaussian likelihoods and on a real-world cosmological dataset. We outline the sampling algorithms themselves and elaborate on performance diagnostics such as convergence time, scope for parallelisation, dimensionality scaling, requisite tunings and suitability for non-Gaussian distributions. We find that nested sampling delivers high-fidelity estimates for posterior statistics at low computational cost, and should be adopted in favour of Metropolis-Hastings in many cases. Affine-invariant MCMC is competitive when computer clusters can be utilised for massive parallelisation. Affine-invariant MCMC and existing extensions to nested sampling naturally probe multi-modal and curving distributions.

This chapter is organised as follows. We begin in Sec. 2.1 by summarising the fundamental results of Bayesian inference, introducing much of the subsequent notation. We then detail the

three sampling techniques which are the focus of this paper. Single-particle Metropolis-Hastings sampling is discussed in Sec. 2.1.2. The formalism and implementation of nested sampling is outlined in Sec. 2.1.3, and a summary of the Goodman & Weare affine-invariant ensemble MCMC sampler is given in Sec. 2.1.4. Their relative performance on a toy-model Gaussian likelihood and a realistic cosmological dataset are tested in Sec. 2.2 and Sec. 2.3, respectively. We discuss these results in Sec. 2.4, and conclude in Sec. 2.5.

2.1 Exploring parameter space

In the framework of Bayesian data analysis we can rigorously discriminate between competing models and determine the region of parameter space that is favoured by the data. We must be able to traverse parameter space in an efficient manner in order to evaluate both the posterior probability distribution for the parameters and the Bayesian evidence, which is the pertinent quantity for model selection. The interpretation and analysis of empirical data using Bayesian methods has widespread scientific value (e.g., Gilks et al., 1995; Qian et al., 2003; Ruiz de Austri et al., 2006; von Toussaint, 2011; Parkinson & Liddle, 2013; Mana et al., 2013).

The scientific literature is filled with examples of particular sampling techniques prescribing how parameter space is navigated (e.g., Metropolis et al., 1953; Mackay, 2003; Skilling, 2004; Feroz et al., 2013). Many distinct branches of science make use of sampling methods to intelligently traverse the parameter space of their models. Different methods are, of course, tailored to particular problems. In the cosmology community Markov chain Monte Carlo (MCMC) methods are widely used (e.g., Lewis & Bridle, 2002). Because MCMC methods are applied ubiquitously, we discuss them in detail here (both single-particle and ensemble methods), comparing them to nested sampling, which was introduced by Skilling (2004). Each method has its own advantages such as speed, simplicity, range of applicability and scope for parallelisation. Indeed, it is the aim of this chapter to assess the relative merits of each sampling algorithm considered. Throughout this discussion we will highlight the requirement to minimise computations of the likelihood, subject to returning an accurate approximation to the joint posterior distribution. This is an important diagnostic; in cosmology, likelihood evaluations often involve running a simulation, or solving a set of coupled differential equations. As such, the calls to the likelihood function are the limiting

factor computationally. It is in this sense that we require a sampler to be efficient.

2.1.1 Bayesian inference

Suppose we have a model M for an observed phenomenon. We would like to determine the posterior probability distribution $\mathbb{P}(\boldsymbol{\theta}|\mathbf{d}, M)$ for the parameters $\boldsymbol{\theta}$ which describe the model, given data $\mathbf{d}$ (Sec. 1.7; Jaynes, 1957; Jaynes & Bretthorst, 2003). From this quantity we can derive all the usual statistics we are interested in, such as parameter means, uncertainties and correlations. Bayes' theorem allows us to express the posterior as a function of simpler, tractable terms (Bayes, 1958):

$$\mathbb{P}(\boldsymbol{\theta}|\mathbf{d}, M) = \frac{\mathbb{P}(\mathbf{d}|\boldsymbol{\theta}, M)\mathbb{P}(\boldsymbol{\theta}|M)}{\mathbb{P}(\mathbf{d}|M)}. \tag{2.1}$$

Here $\mathcal{L}(\boldsymbol{\theta}) \equiv \mathbb{P}(\mathbf{d}|\boldsymbol{\theta}, M)$ is the likelihood, which expresses the explanatory power of a given set of parameters. The prior $\pi(\boldsymbol{\theta}) \equiv \mathbb{P}(\boldsymbol{\theta}|M)$ represents our degree of belief about the parameters before we have knowledge of the data. We denote the Bayesian evidence $\mathcal{Z} \equiv \mathbb{P}(\mathbf{d}|M)$, which for any given model can be written down in integral form,

$$\mathcal{Z} = \int \mathcal{L}(\boldsymbol{\theta})\pi(\boldsymbol{\theta})d\boldsymbol{\theta}, \tag{2.2}$$

since it is the normalisation factor in Eq. 2.1. This formalism allows us to infer parameters from data. Now consider the case where we have two competing models M_1 and M_2, both of which purport to describe the same phenomenon. We can again make use of Bayes' theorem to find

$$\frac{\mathbb{P}(M_1|\mathbf{d})}{\mathbb{P}(M_2|\mathbf{d})} = \frac{\mathcal{Z}_1}{\mathcal{Z}_2}\frac{\mathbb{P}(M_1)}{\mathbb{P}(M_2)}. \tag{2.3}$$

Assuming that our *a priori* belief in the two models is equal, then the evidence ratio $\mathcal{Z}_1/\mathcal{Z}_2$, also called the *Bayes factor*, completely specifies the relative probability of the two models (e.g., Sivia, 1996; Hobson et al., 2002).

The posterior encodes all that the data have to tell us in the context of a given model. Therefore the salient task, for parameter estimation, is to construct an accurate approximation to the posterior. To do this we must sample the posterior sufficiently densely that numerical uncertainties on any quantity we evaluate are negligible for our purposes. Likelihood evaluations

can be computationally expensive; for example, calculating the cosmic microwave background (CMB) power spectrum likelihood at one set of cosmological parameters requires us to evolve many coupled differential equations to the present day, which can take on the order of seconds to evaluate (e.g., Lewis et al., 2000). Simple, brute force grid sampling is therefore not feasible since the number of likelihood evaluations would scale exponentially with the number of parameters. There are alternatives, however, many of which are particular manifestations of the Markov chain Monte Carlo (MCMC) class of sampling algorithms. MCMC samplers work by constructing a Markov chain in parameter space whose equilibrium distribution is the posterior itself. We summarise the implementation of two such algorithms below. We also outline the method of nested sampling, a non-MCMC prescription for how parameter space is explored.

2.1.2 Metropolis-Hastings sampling

The Metropolis-Hastings algorithm is one of the simplest MCMC sampling methods and has been applied to a huge variety of parameter estimation problems (Metropolis et al., 1953). We summarise the essentials of the method for a one-dimensional problem with posterior $p(x)$ as follows:

1. Choose initial point in parameter space x_0
2. At each step i propose a trial step $x_{\rm trial}$ drawn from a symmetric trial distribution $q(x_{\rm trial}, x_i)$; this is often taken to be a Gaussian centred on x_i.
3. Define $P = \min\{1, p(x_{\rm trial})/p(x_i)\}$. Set $x_{i+1} = x_{\rm trial}$ with probability P, otherwise set $x_{i+1} = x_i$.
4. Iterate from step 2 to obtain the chain of points $\{x_i\}$.

For more details see e.g., Lewis & Bridle (2002) and Dunkley et al. (2005). The algorithm is constructed such that steps towards regions of higher posterior probability ('uphill steps') are always taken, while 'downhill steps' are taken only occasionally, and are less probable for larger descents. The symmetry of the trial distribution q guarantees the stationarity of $p(x)$ under the Markov process and thus that the asymptotic distribution of the chain is $p(x)$.

For a D-dimensional parameter space in which we have N samples of the posterior $\{\mathbf{x}_i\}$, the mean and covariance of the posterior can be estimated using

$$\bar{\mathbf{x}} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i, \quad \mathbf{C} = \frac{1}{N-1}\sum_{i=1}^{N}(\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^{\mathbf{T}}. \tag{2.4}$$

Although the Markov chain is guaranteed to converge asymptotically to a perfect sampling of the posterior, we require, in practice, a high-quality sampling in as few chain steps as possible to mitigate computational overheads. This requires a judicious choice of trial distribution, the shape of which can drastically affect the acceptance rate and convergence time (e.g., Gelman et al., 1996). For a Gaussian likelihood with covariance $\mathbf{C}_0$, and uniform priors, Dunkley et al. (2005) showed the optimal choice for the trial distribution is a Gaussian with covariance

$$\mathbf{C} = (2.4^2/D)\mathbf{C}_0. \tag{2.5}$$

In this case the number of likelihood evaluations for convergence to a given level of accuracy scales linearly with dimension D:

$$N_{\rm like} \approx 330D. \tag{2.6}$$

Here convergence is defined by the requirement that the variance of the chain means, $\sigma^2_{\bar{x}}$, is much less than the variance of the underlying distribution, σ_0^2, for each parameter x (σ_0^2 is estimated from the within-chain variance). Dunkley et al. (2005) define the convergence statistic

$$r = \frac{\sigma^2_{\bar{x}}}{\sigma_0^2}. \tag{2.7}$$

One can estimate r by spectral methods, or by running several parallel chains and evaluating it directly; we stop the chains when $r < 0.01$. The widely used Gelman-Rubin parameter R is related to this statistic by $R \approx 1 + r$ (Gelman & Rubin, 1992).

The cosmology community has historically favoured Metropolis-Hastings as a parameter space sampler (e.g., Spergel et al., 2003; Anderson et al., 2012). This is due its ease of implementation and straightforward interpretation; the posterior is given by the number density of chain samples in the long-chain limit. The principal analysis of temperature data from the Planck satellite

used a modified Metropolis-Hastings algorithm, following the methods of Lewis (2013), which incorporates a decomposition of the full parameter space into *fast* and *slow* subspaces for enhanced speed (*Planck* Collaboration, 2013a).

Metropolis-Hastings sampling has some drawbacks. Choosing the trial distribution relies on our *a priori* knowledge of the target distribution and its degeneracies. In many cases this is precisely what we are trying to find by sampling the posterior. Therefore, in practice, we may be unable to choose an optimal trial distribution, and reaching convergence can take many orders of magnitude longer than indicated by Eq. 2.6. For a D-dimensional parameter space there are $D(D+1)/2$ independent components of the covariance matrix; in high-dimensional parameter spaces a poorly-estimated covariance matrix will result in an impractically long convergence time. Furthermore, Metropolis-Hastings suffers from long *burn-in* times if one makes an initial choice of parameters which lie far from the bulk of the posterior. Burn-in refers to the initial section of the chain which is not a representative sampling of the posterior. These samples must be discarded in the calculation of the posterior and thus, in some sense, require redundant likelihood calls. Moreover, this sampling technique does not naturally return an estimate of the Bayesian evidence.

2.1.3 Nested sampling

Nested sampling was first presented by Skilling (2004) as a method for Bayesian evidence calculation. This method was developed to sidestep computationally-expensive thermodynamic integration techniques. Nested sampling begins with a sampling of the entire prior volume. Samples are then drawn from successively more likely regions of parameter space until the posterior bulk, which is oversampled with respect to the full prior, is located and explored. The samples are weighted appropriately and the posterior and evidence may be estimated.

Nested sampling was first used in a cosmological application by Bassett et al. (2004) for discriminating between dark energy models. Mukherjee et al. (2006) introduced a practical method for sampling from restricted regions of prior space based on using an ellipsoidal approximation to the likelihood contours. This was developed further by Shaw et al. (2007), who demonstrated the improved efficiency of nested sampling in evaluating the Bayesian evidence compared to thermodynamic integration methods. Feroz & Hobson (2008) and Feroz et al. (2009) extended sin-

gle ellipsoidal nested sampling into MULTINEST, which deals with multi-modal likelihoods via a cluster-detection algorithm. Feroz et al. (2013) introduce importance nested sampling which can offer further computational gains. For the sake of performance comparison and analysis we implement our own python nested sampling routine designed to explore the uni-modal likelihoods encountered in common cosmological models. We outline our practical implementation of the scheme below – see Skilling, 2004, for a detailed discussion of the formalism.

Nested sampling formalism and implementation

We introduce the *prior mass* $X(L)$ associated with likelihoods $\mathcal{L}(\boldsymbol{\theta})$ greater than L:

$$X(L) = \int_{\mathcal{L}(\boldsymbol{\theta})>L} \pi(\boldsymbol{\theta}) d\boldsymbol{\theta}. \tag{2.8}$$

Clearly $X_0 \equiv X(0) = 1$ and $X(L \geq \mathcal{L}_{\max}) = 0$. Defining the inverse function $L(X)$ as the likelihood which bounds a prior mass X, one can simplify Eq. 2.2 for the Bayesian evidence from a multi-dimensional integral over the prior volume to a one-dimensional integral over the prior mass:

$$\mathcal{Z} = \int_0^1 L(X) dX. \tag{2.9}$$

Here dX is the prior mass associated with likelihoods in the interval $[L, L + dL]$. The integral is well-defined for continuous $\mathcal{L}$ and π with connected support (Chopin & Robert, 2010).

Following Mukherjee et al. (2006), our nested sampling scheme works as follows:

1. Draw N points (the *active set* A) from the prior.

2. At iteration i the least likely point in the active set, which we denote $\mathbf{x}_i$ and is such that $\mathcal{L}(\mathbf{x}_i) = L_i$, becomes the ith sample point.

3. Sample the prior, subject to $\mathcal{L} > L_i$. To do this:

   (a) Approximate the likelihood contour $\mathcal{L} = L_i$ by an ellipsoid which bounds the active set (see Sec. 2.1.3).

   (b) Draw uniformly from the ellipsoid until one has a point $\mathbf{y}$ such that $\mathcal{L}(\mathbf{y}) > L_i$.

4. This new point $\mathbf{y}$ replaces $\mathbf{x}_i$ in the active set.

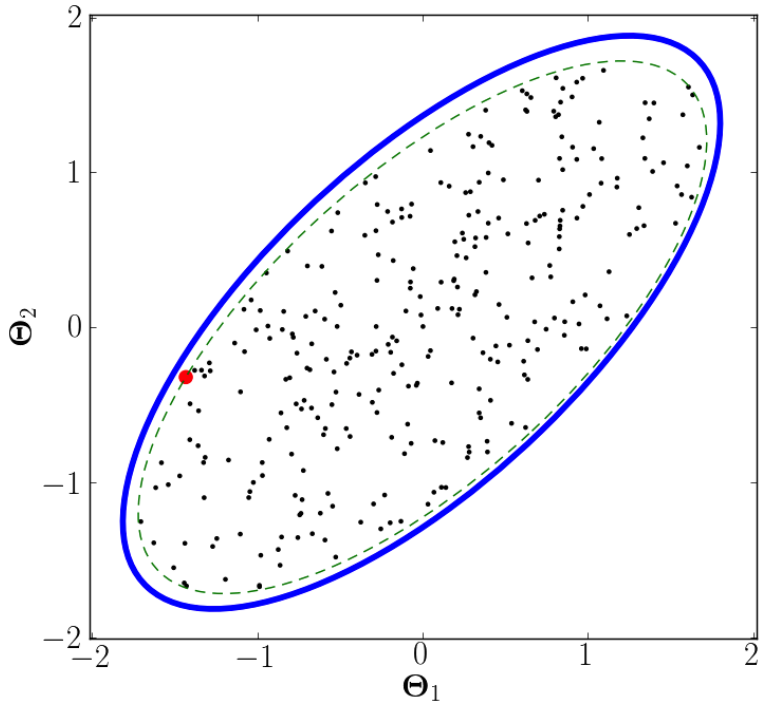

Figure 2.1: *Nested sampling*: the active set ($N = 300$, *black points*) and the inferred bounding ellipse (*blue solid curve*, $f = 1.06$) for a two-parameter Gaussian likelihood ($\sigma_1 = \sigma_2 = 1, \rho = 0.7$). The input mean is $\boldsymbol{\mu} = (0, 0)$. The ellipse is inferred from the mean and covariance of the active set (Eq. 2.17). The true likelihood contour (*green dashed curve*) corresponding to the lowest likelihood point (*red point*) is everywhere encompassed by the ellipse, as necessary for unbiased results.

5. Iterate from step 2 until the stopping criterion is satisfied.

The initial active set represents N points drawn from the prior; this is equivalent, by construction of the prior mass X, to drawing N values for X uniformly over $[0, 1]$. Choosing the lowest likelihood point is equivalent to choosing the largest of these N standard uniform deviates, which has the density

$$p(t) = Nt^{N-1}, \tag{2.10}$$

for $t \in [0, 1]$; we denote the distribution described by this density as T. Thus the prior mass enclosing all points in the active set shrinks to $X_1 = tX_0$, where $t \sim T$. For each subsequent iteration i we choose the lowest likelihood point from which to restrict the prior, and therefore each iteration the prior mass X_i shrinks by a factor t, given probabilistically by Eq. 2.10. We

obtain a sequence for the remaining prior mass at each iteration:

$$X_0 = 1, X_1 = t_1 X_0, X_2 = t_2 X_1, ... \tag{2.11}$$

where the $t_i \sim T$. After many iterations the logarithmic prior mass $\ln(X_i) = \ln(t_1 ... t_i X_0)$ is the sum of many independent and identically distributed random variables, and so the expectation and standard deviation fully characterise the distribution of $\ln(X_i)$:

$$\ln(X_i) = -\frac{i \pm \sqrt{i}}{N}. \tag{2.12}$$

Thus, the prior mass shrinks exponentially with each iteration; to estimate the evidence and posterior we set

$$X_i = \exp\left(-\frac{i}{N}\right), \tag{2.13}$$

and we may straightforwardly propagate the uncertainty on the X_i into an uncertainty on posterior statistics and the final evidence estimate. The evidence is estimated, by the trapezium rule, as

$$\mathcal{Z} = \sum_{i=1}^{M} L_i w_i + \bar{L} X_M, \tag{2.14}$$

where the L_i are the points of lowest likelihood at each iteration, $w_i = \frac{1}{2}(X_{i+1} - X_{i-1})$ is the prior mass over which $\mathcal{L}(\boldsymbol{\theta}) \approx L_i$ and M is the total number of iterations. The additional term $\bar{L} X_M$ is added to account for the contribution from the active set; we assume each point in the active set occupies an equal fraction of remaining prior volume X_M, and $\bar{L}$ is their average likelihood. Including the active set ensures that the peak of the distribution is fully mapped out. The joint posterior distribution may be inferred by binning up the sample points $\mathbf{x}_i$ with weights

$$p_i = \begin{cases} \dfrac{L_i w_i}{\mathcal{Z}}, & i \in \{1, ..., M\}, \\ \dfrac{L_i X_M}{N \mathcal{Z}}, & i \in \{M+1, ..., M+N\}. \end{cases} \tag{2.15}$$

The mean and covariance of the posterior can then be estimated by

$$\bar{\mathbf{x}} = \sum_{i=1}^{n} p_i \mathbf{x}_i, \quad \mathbf{C} = \sum_{i=1}^{n} p_i (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^{\mathbf{T}}, \tag{2.16}$$

where $n = M + N$.

Drawing from a restricted prior

At each iteration i we need to sample the prior restricted to the region of parameter space such that $\mathcal{L}(\boldsymbol{\theta}) > L_i$. Because the prior mass shrinks exponentially with each iteration (Eq. 2.13), it is important, computationally, to avoid sampling the whole prior throughout the course of the algorithm, since this would lead to exponentially-worsening acceptance rates. Mukherjee et al. (2006) and Shaw et al. (2007) suggest using an ellipsoidal approximation to the likelihood contour which is defined by the active set, sampling the prior restricted to this ellipsoid. We summarise this method below, while Fig. 2.1 shows the process pictorially.

We define the ellipsoidal approximation to the bounding likelihood contour by

$$(\mathbf{x} - \boldsymbol{\mu})^T \mathbf{C}^{-\mathbf{1}} (\mathbf{x} - \boldsymbol{\mu}) = k, \tag{2.17}$$

where $\boldsymbol{\mu}$ is the centroid, and $\mathbf{C}$ the covariance matrix, of the active set A. Here

$$k = \max \left\{ (\mathbf{x}_i - \boldsymbol{\mu})^T \mathbf{C}^{-\mathbf{1}} (\mathbf{x}_i - \boldsymbol{\mu}) : \mathbf{x}_i \in A \right\}, \tag{2.18}$$

is defined such that the ellipsoid is scaled to encompass the entire active set. We further expand the ellipsoid along each principal axis by a factor f to ensure the entire prior mass $X(L_i)$ is encompassed; this is required for unbiased results. We adopt $f = 1.06$ following Shaw et al. (2007). This expansion is effected by redefining $k \to kf^2$. Note that correspondingly the volume of the ellipsoid $V \to Vf^D$.

To sample uniformly from the D-dimensional ellipsoid we sample uniformly from the unit D-ball and then construct a linear map $\mathbf{T}$ from the latter to the former. We draw a D-dimensional

vector of standard Gaussian random numbers $\mathbf{w}$ and define the unit-vector

$$\mathbf{z} = \frac{\mathbf{w}}{|\mathbf{w}|}, \tag{2.19}$$

which lies on the surface of the unit D-ball. A uniform deviate $u \in [0, 1]$ then maps $\mathbf{z}$ to a point *within* the unit D-ball:

$$\mathbf{z} \to u^{1/D}\mathbf{z}. \tag{2.20}$$

To map this to the ellipsoid the co-ordinates of $\mathbf{z}$ are firstly rotated into the frame of the principal axes of the ellipsoid, then scaled according, and then rotated back. The overall transformation matrix is

$$\mathbf{T} = \sqrt{k}\mathbf{R}^{\mathbf{T}}\mathbf{D}\mathbf{R}, \tag{2.21}$$

where $\mathbf{D}$ is the square root of the diagonalised covariance matrix and $\mathbf{R}$ is the matrix of eigenvectors of $\mathbf{C}$. Finally, we shift the origin to lie at the centroid $\boldsymbol{\mu}$ of the active set. Thus we obtain a deviate $\mathbf{y}$ drawn uniformly from the bounding ellipsoid:

$$\mathbf{y} = \mathbf{T}\mathbf{z} + \boldsymbol{\mu}. \tag{2.22}$$

Nested sampling involves just two tunable hyper-parameters: the number of points in the active set N and the expansion factor f. Shaw et al. (2007) investigated the effect of varying f on the numerical evidence for toy-model cases (in which the evidence can be computed analytically). They found $f = 1.06$ sufficient to give unbiased results, and we adopt this as our fiducial value in all cases. As shown in Sec. 2.2, the size of the active set directly determines the density of sampling in parameter space and, correspondingly, the number of iterations until convergence. Thus N should be chosen large enough so as to provide an accurate approximation to the posterior, while being not so large as to invoke impractical computational overheads. Typically $N \sim O(10^2)$ offers this compromise, and we investigate this choice in Sec. 2.2.

Stopping criterion

When do we stop sampling the parameter space? For Metropolis-Hastings we use the Gelman-Rubin type statistic r, which ensures the variance of the chain mean is much less than the posterior

variance for each parameter. Unfortunately this stopping criterion is not appropriate for nested sampling; as a consequence of the sampler traversing the full parameter space, the variance of the sample means can become very small long before we have properly explored the bulk of the posterior, particularly when we use a large active set. Instead we consider the evidence accumulated throughout the course of the algorithm following Shaw et al. (2007). The typical trajectory for the values of the posterior weights p_i (or evidence increments $L_i w_i$) can be seen in Fig. 2.2. The peak of this curve corresponds to the exploration of the bulk of the posterior. We use this observation as a guide for when to stop the algorithm: when the points in the active set would, in sum, increment the evidence by only some small fraction of the total evidence accumulated, we can safely assume the posterior bulk has been explored sufficiently. The maximum possible contribution to the evidence from the active set at the ith iteration is approximately $\Delta\mathcal{Z}_i = L_{\max} X_i$, and so this condition can be expressed quantitatively as: *stop if*

$$\log(\mathcal{Z}_i + \Delta\mathcal{Z}_i) - \log(\mathcal{Z}_i) < \kappa. \tag{2.23}$$

Here $L_{\max}$ is the maximum likelihood point in A and we take $\kappa = 0.1$, which is small enough that the error on the accumulated evidence due to truncation is negligible compared to the uncertainty resulting from the deterministic approximation for X_i (Eq. 2.13; Chopin & Robert, 2010).

2.1.4 Affine-invariant ensemble MCMC

Goodman & Weare (2010) introduced an ensemble (many-particle) MCMC sampling algorithm which has the property of affine-invariance; that is, the performance of the algorithm is invariant under linear transformations of the parameter space. Thus, in particular, the sampler works just as well on a highly-degenerate Gaussian distribution as an uncorrelated and isotropic Gaussian distribution. The principle is that many particles, or *walkers*, move through parameter space; at each iteration each walker undergoes a trial move, with the step being accepted with some probability (described below). The trial move is based on the positions of each of the other walkers (the complementary set), since these provide information about the underlying distribution.

We describe the algorithmic scheme below for a one-dimensional density $p(x)$ (see also Foreman-Mackey et al., 2013):

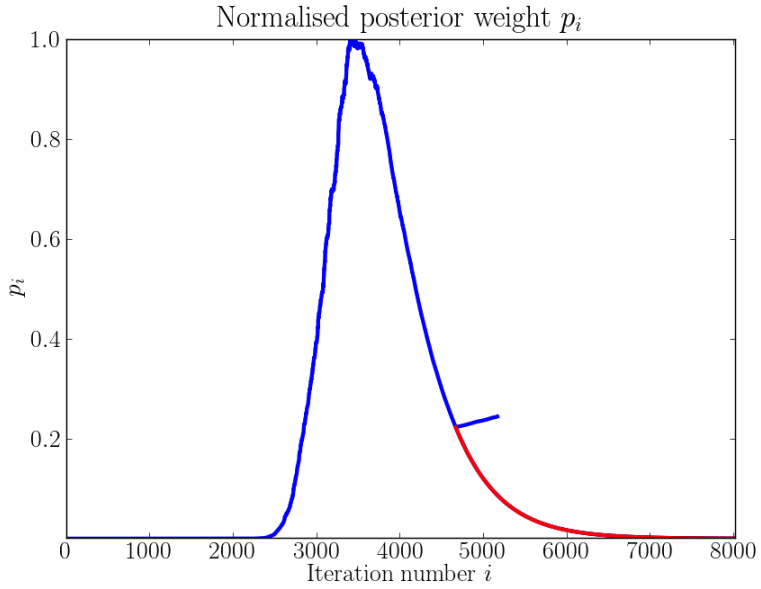

Figure 2.2: *Nested sampling*: typical trajectory for posterior weights p_i (or equivalently evidence increments $L_i w_i$) as a function of iteration number i (*blue*). Note how, initially, the curve is increasing, as samples are drawn from successively higher likelihood regions. As the active set enters the posterior bulk, around $i \sim 3000$, the likelihood L_i flattens off, the exponentially decreasing prior volume weighting factor w_i becomes dominant and the weights p_i decrease. When the stopping criterion is satisfied (Eq. 2.23) the points in the active set (sorted arbitrarily from low to high likelihood) are inserted into the chain resulting in the small correction from $i \sim 4700$. The red curve shows the weights trajectory if we do not use the stopping criterion. Here the active set size $N = 500$, and truncating the algorithm results in less than a 0.02% discrepancy on the computed evidence – a negligible effect.

1. The positions of the n walkers are initialised ($t = 0$). Suppose at iteration t the positions of all the walkers are described by $\mathbf{x}(t)$

2. For each of the walkers $x_j(t), j = 1, ..., n$, successively:

   (a) Propose a trial step $x_{\rm trial}$ (the *stretch move*):

$$x_{\rm trial} = x_k + z(x_j(t) - x_k), \tag{2.24}$$

   where x_k is a random walker from $\mathbf{x}_{[j]}(t)$, the set of positions of the other walkers, and we draw z from

$$g(z) \propto z^{-\frac{1}{2}}, \quad z \in \left[\frac{1}{a}, a\right]. \tag{2.25}$$

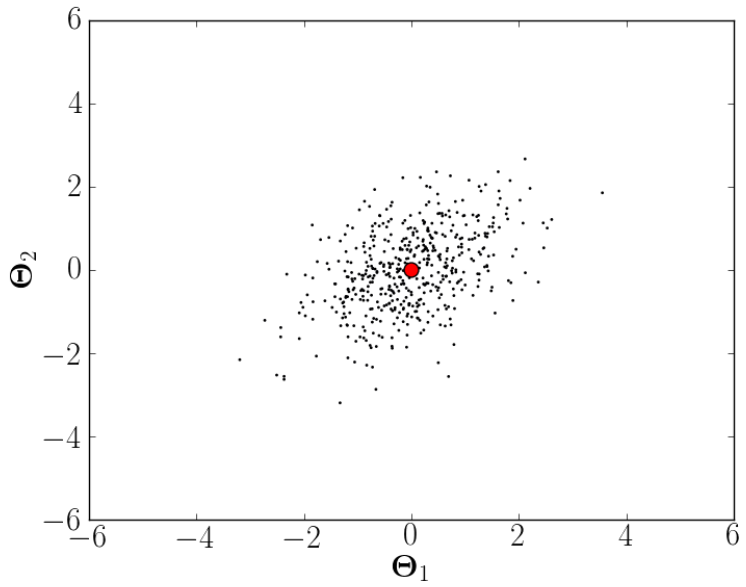

Figure 2.3: *Affine-invariant MCMC*: snapshot of the positions of the $n = 500$ walkers (black points) after $t = 100$ steps (per walker) for a two-parameter Gaussian likelihood ($\sigma_1 = \sigma_2 = 1, \rho = 0.4$). The input mean $\boldsymbol{\mu} = (0, 0)$ has been highlighted (red dot).

(b) Define $s = \min\left\{1, z^{D-1} p(x_{\rm trial})/p(x_j(t))\right\}$. Assign

$$x_j(t+1) \leftarrow \begin{cases} x_{\rm trial} & \text{with prob. } s \\ x_j(t) & \text{with prob. } 1-s \end{cases} \tag{2.26}$$

3. Iterate over t from step 2 to obtain $\{\mathbf{x}(t)\}$.

Note that we set the step-size parameter $a = 2$ in all cases, but in principle a may be varied if the acceptance fraction is too low or too high (see Goodman & Weare, 2010; Foreman-Mackey et al., 2013, for discussion). The above form of $g(z)$ makes the trial step symmetric and, in conjunction with the modified Metropolis-Hasting acceptance probability s, ensures detailed balance and therefore that the asymptotic distribution is p. See Fig. 2.3 for an example of walker positions at one particular time-step, long after burn-in, for a degenerate Gaussian likelihood.

For affine-invariant MCMC we use the autocorrelation time to define the stopping criterion, following Goodman & Weare (2010) and Foreman-Mackey et al. (2013). One must be careful to distinguish burn-in (where the mean walker position can vary considerably) from when the walkers are properly sampling the posterior (where the mean position makes only small oscillations

around the mean parameters). The autocorrelation time measures the longest wavelength of these oscillations, and corresponds to the number of steps between independent samples of the posterior. Burn-in is estimated from the *exponential* autocorrelation time $\tau_{\rm exp}$ for the mean position of the walkers at each subsequent iteration, found by a least-squares fit to the autocorrelation function. We cut $b = 5\tau_{\rm exp}$ steps as burn-in to ensure the initialisation is forgotten, and then estimate in the same way the autocorrelation time τ for the post burn-in samples from the remaining steps (see Sec. 2.3.2 for more details). The *integrated* autocorrelation time $\tau_{\rm int}$ measures the variance on the mean for each parameter x; if after burn-in we have N samples in total:

$$\sigma_{\bar{x}}^2 = \frac{2\tau_{\rm int}}{N}\sigma_0^2. \tag{2.27}$$

We find, as in Akeret et al. (2013), that the autocorrelation function has an exponential form and so we may set $\tau = \tau_{\rm exp} = \tau_{\rm int}$. Thus, using Eq. 2.7 in Eq. 2.27, the total number of likelihood evaluations for convergence is

$$N_{\rm like} = \frac{2\tau}{r} + bn. \tag{2.28}$$

As for Metropolis-Hastings we demand $r = 0.01$ for convergence, so we may make a direct comparison between the two methods.

The mean and covariance of a chain of walkers are given, as for Metropolis-Hastings, by Eq. 2.4, where only the samples after burn in are used. Affine-invariant ensemble MCMC uses the positions of the walkers at each step to provide information about where to sample next. A curving ('banana-shape') distribution would result in a low efficiency for Metropolis-Hastings because the trial distribution cannot be tuned throughout the parameter space (the optimal trial covariance varies as function of position). However for this multiple-walker technique, the positions of the walkers ensure trial steps throughout parameter space are restricted to the posterior bulk, and therefore that the acceptance probability is sufficiently high for practical applications. The only tunable parameters for this sampling technique are the step-size parameter a and the number of walkers n. The number of walkers can be chosen to suit the application, but typically $n \sim O(10)$ or $O(10^2)$.

2.2 Toy-model (Gaussian) likelihoods

In this section we focus on the number of likelihood evaluations required for convergence, $N_{\rm like}$, for each of the three sampling methods, for the case of simple Gaussian likelihoods. In models where likelihood calls are expensive, $N_{\rm like}$ is a diagnostic quantity for the computational effort required for each of these sampling methods. Example python code – for nested and affine-invariant sampling – has been made publicly available; this allows the user to explore the toy-model Gaussian likelihoods studied in this section[1].

2.2.1 An analytical result for nested sampling

We firstly derive a new analytical estimate for the expected number of likelihood evaluations required for convergence in nested sampling, in the case of a multivariate Gaussian likelihood and uniform priors. The result, although strictly holding only for this special case, is nevertheless useful for order-of-magnitude predictions for more complicated distributions. We begin with the stopping criterion in Eq. 2.23. Let the sampler converge after M iterations; then M satisfies $\Delta\mathcal{Z}_M/\mathcal{Z}_M = s$ for constant $s = e^{\kappa} - 1 \approx 0.1$. Using $\Delta\mathcal{Z}_M = L_{\rm max}X_M$ and $\mathcal{Z}_M \approx \mathcal{Z}$, the total number of iterations M will approximately satisfy

$$M = N \ln\left(\frac{L_{\rm max}}{s\mathcal{Z}}\right). \tag{2.29}$$

Consider a Gaussian likelihood

$$\mathcal{L}(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^D \det \mathbf{C}}} \exp\left[-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T\mathbf{C}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right], \tag{2.30}$$

and a uniform prior of volume $V_{\rm p}$ which comfortably contains the bulk of the posterior, so that $\mathcal{Z} = 1/V_{\rm p}$. Then near convergence $L_{\rm max} \approx \mathcal{L}(\boldsymbol{\mu})$, so Eq. 2.29 becomes

$$M = N \ln\left(\frac{V_{\rm p}}{V_{\rm t}s}\right), \tag{2.31}$$

where $V_{\rm t} = \sqrt{(2\pi)^D \det \boldsymbol{C}}$ is a measure of the volume of the target (posterior) distribution.

At each iteration i the ellipsoidal bound will not necessarily be congruous with the likelihood

[1] https://github.com/rupert-allison/sampling

contour $\mathcal{L} = L_i$; indeed, for unbiased results we require only that the locus of the ellipsoid satisfies $\mathcal{L} \leq L_i$ everywhere. Thus some samples, drawn uniformly from the interior of the ellipsoid, will also be drawn from $\mathcal{L} < L_i$, and these must be rejected. In order to calculate the number of likelihood evaluations required for convergence we need to understand this acceptance rate. In the case of uniform priors, the acceptance rate is essentially the ratio of the volume of the restricted prior ($\mathcal{L} > L_i$) which we denote by $V_{\rm like}$, to the volume of the bounding ellipsoid $V_{\rm ell}$. Since we always scale the ellipsoid until it encompasses the active set, and because we expand each principal axis by a factor f, we parameterise the volume ratio – and thus the acceptance rate AR – by

$$\mathrm{AR} = \frac{V_{\rm like}}{V_{\rm ell}} = \left(\frac{\alpha}{f}\right)^D . \tag{2.32}$$

Here α represents the mean ratio between the linear dimension of the likelihood contour and the ellipsoid before we expand by f. We expect this parameter to increase with increasing N and decreasing D as the constraints on the covariance improve. Using this parameterisation and Eq. 2.29 we can finally write down an expression for the number of likelihood evaluations until convergence, $N_{\rm like}$:

$$N_{\rm like} = N\left[\left(\frac{f}{\alpha}\right)^D \ln\left(\frac{V_{\rm p}}{V_{\rm t}s}\right) + 1\right], \tag{2.33}$$

since $N_{\rm like} = M/\mathrm{AR} + N$, where the last term accounts for the active set.

To estimate how α depends on N and D, we performed simple numerical studies based on drawing N points uniformly from an ellipsoid then comparing the volume of the inferred bounding ellipsoid to the actual volume bounded by the lowest likelihood contour. We find that for Gaussian likelihoods, in general, $\alpha \in [0.92, 1]$. The dependance on N and D is weak, justifying the parameterisation used in Eq. 2.32. Setting $\alpha = 0.92$ in Eq. 2.33 gives an upper bound on $N_{\rm like}$ for Gaussian likelihoods.

Since in general $V_{\rm p}/V_{\rm t}$ scales exponentially with dimension D, the logarithmic term gives a contribution linear in D, which is the same scaling as Metropolis-Hastings. The additional exponential term deriving from the imperfect estimation of the covariance from the active set means that convergence for nested sampling scales more severely with dimension.

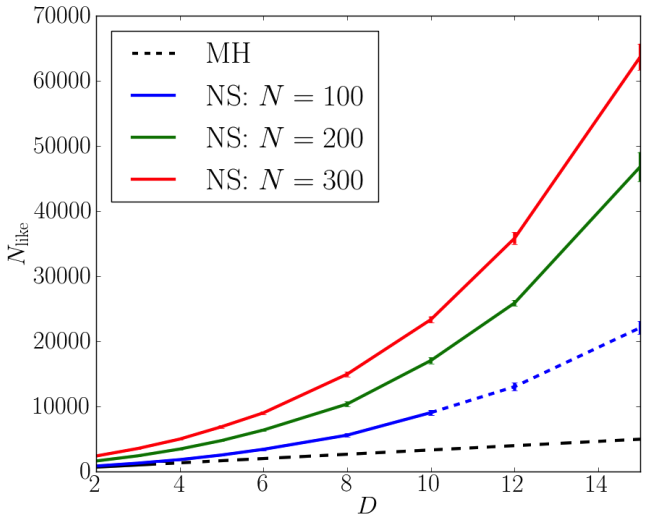

Figure 2.4: Number of likelihood evaluations for convergence, $N_{\rm like}$, as a function of parameter space dimension D for optimal Metropolis-Hastings (MH) and nested sampling (NS). Here we consider an isotropic Gaussian likelihood and a conservative (large) value for the prior to posterior volume ratio of $(20/\sqrt{2\pi})^D$. The error bars for the NS curves show the standard deviation from multiple runs. For MH we show the result of Dunkley et al. (2005): $N_{\rm like} = 330D$, which ensures $r < 0.01$. For NS the time for convergence is a linear function of the size of the active set N, and the steepening gradient with dimension is because probability mass is more concentrated at larger radii in higher dimensions. These curves can be modelled by Eq. 2.33. We dash the $N = 100$ curve for high dimensions because the posterior statistics are likely unreliable in this regime where the active set does not accurately constrain the covariance.

2.2.2 Comparison of sampling methods on Gaussian likelihoods

Here we implement and compare each of the three sampling techniques on Gaussian toy-model likelihoods. We consider Metropolis-Hastings sampling where the trial covariance matrix is optimally tuned (Eq. 2.5). Poorly-tuned schemes can be computationally prohibitive even in the isotropic Gaussian regime, e.g., a factor of 2 under-estimation in the trial step-size in each of 7 dimensions would require a factor of over 10^2 more iterations for convergence than in the optimally-tuned case; see, e.g., Fig. 9 of Dunkley et al. (2005). Additionally, not properly accounting for parameter degeneracies will further lengthen the convergence time.

Fig. 2.4 demonstrates the non-linear scaling of nested sampling with the dimensionality of the parameter space. We find that the form of these curves is accurately modelled by the analytic result given in Eq. 2.33. This effect arises due to the concentration of probability mass at larger radii in higher dimensional spaces, and so more samples are drawn from the region (within the ellipsoid) where $\mathcal{L} < L_i$, reducing the acceptance rate. As seen in Eq. 2.33 the convergence time

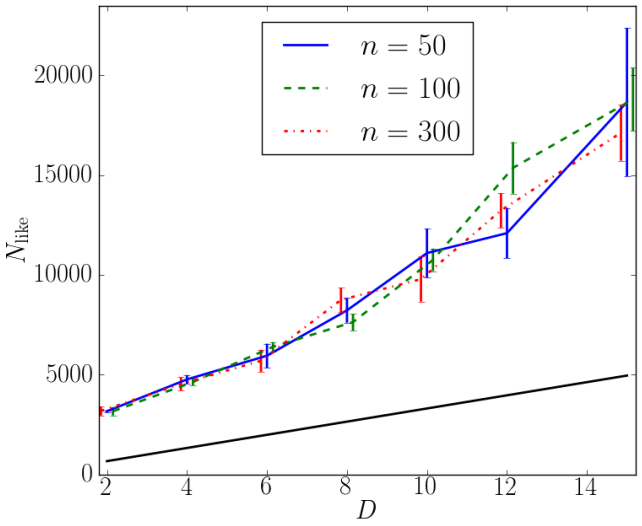

Figure 2.5: Number of likelihood evaluations for convergence $N_{\rm like}$ as a function of parameter space dimension D, for optimal Metropolis-Hastings (MH) and affine-invariant MCMC (Aff-Inv). Here we consider a toy-model isotropic Gaussian likelihood and assume negligible burn-in time. Note that Aff-Inv requires more likelihood evaluations than optimal MH for convergence to the same level of accuracy on the parameter means ($r = 0.01$); this is due to a longer autocorrelation time for the chain. However, for a poorly-tuned trial distribution the convergence time can be an order of magnitude longer for MH than Aff-Inv, which requires essentially no tuning. The convergence time for Aff-Inv is independent of the number of walkers n since we neglect burn-in. We have slightly displaced the $n = 100$ and $n = 300$ data points from integer dimension D for visual clarity. The error bars represent the 1σ sampling uncertainty derived from multiple runs.

only depends logarithmically on the prior to posterior volume ratio, and thus should be insensitive to the particular choice of the extent of the prior. Larger active sets take longer to converge, but provide correspondingly more posterior samples, and hence less statistical uncertainty on the posterior statistics. More quantitatively, the posterior bulk is reached in $N_{\rm tar} \sim N \ln(V_{\rm p}/V_{\rm t})$ iterations, so the effective number of samples of the posterior is $N_{\rm eff} = M - N_{\rm tar} + N \sim N \ln(e/s)$. Thus larger active sets should be chosen for a denser sampling of the posterior (and hence less noisy estimates of the mean and covariance). For quick parameter constraints the smallest possible active set should be chosen, subject to $N \gg D$. This condition ensures that the ellipsoidal bound contains the entire restricted prior volume, and hence provides unbiased results.

Fig. 2.5 shows the convergence time $N_{\rm like}$ (to achieve $r = 0.01$) for optimal Metropolis-Hastings and affine-invariant MCMC on a toy-model isotropic Gaussian likelihood. The number of likelihood evaluations required is higher for affine-invariant MCMC than for Metropolis-Hastings. Non-optimal Metropolis-Hastings, however, can give significantly longer convergence times. Because

we neglect burn-in by starting the walkers from a sampling of the posterior, $N_{\rm like}$ is independent of the number of walkers n. The convergence time for affine-invariant MCMC is lower than nested sampling for all high dimensional spaces, but nested sampling can give less statistical noise on the posterior statistics for the same computational effort.

2.3 Cosmological parameters

Here we test the performance of each of the three sampling techniques on real data. We focus on several performance diagnostics: the number of likelihood evaluations required for convergence, the scope for parallelisation, dimensionality scaling, requisite tunings and performance on non-Gaussian distributions.

2.3.1 The data and the model

We consider cosmic microwave background (CMB) data from the Wilkinson Microwave Anisotropy Probe (*WMAP*) 7 year observations (Jarosik et al., 2011). The data we use are measurements of the maps and angular power spectra of the CMB anisotropies in temperature and polarisation (Larson et al., 2011). We adapt the popular CosmoMC code, replacing the default Metropolis-Hastings sampler by python implementations of nested sampling and the Goodman & Weare affine-invariant ensemble sampler. The python samplers are called within the Fortran CosmoMC, giving us the usability of python (for the modified code) whilst retaining the speed and robustness of the well-tested Fortran code in the computationally intensive part of the calculation.

The chosen sampler passes a set of cosmological parameters to camb: this code then integrates the relevant Boltzmann equations – which encode the physics of the production and propagation of the CMB – and computes theory power spectra, given this parameter set (Lewis et al., 2000). The *WMAP* likelihood code then computes how well the theory explain the data, and this number is passed back to the sampler which then chooses a new parameter set (as explained in Sec. 2.1). This process continues until the stopping criterion is satisfied.

For comparison purposes we consider the concordance ΛCDM model, which is completely described by 6 parameters. We take as a basis $\left\{\Omega_{\rm b}h^2, \Omega_{\rm c}h^2, \theta, \tau, n_s, \ln(10^{10}A_s)\right\}$: the physical baryonic density, the physical cold dark matter density, the angular size of the sound horizon at

	MH		NS				Aff-Inv		
	Optimal	Non-optimal	$N = 50$	$N = 100$	$N = 200$	$N = 300$	$n = 30$	$n = 50$	$n = 100$
$N_{\rm like}$	6400	110800	2300	5400	11600	17900	17900	22900	35800
M	6400	110800	900	1800	3600	5300	17900	22900	35800
AR	0.35	0.09	0.38	0.33	0.31	0.30	0.48	0.45	0.46
r	0.0095	0.0099	0.0053	0.0023	0.0009	0.0008	0.01	0.01	0.01

Table 2.1: Number of likelihood evaluations $N_{\rm like}$ for convergence for each of Metropolis-Hastings (MH), nested sampling (NS) and affine-invariant sampling (Aff-Inv), for a 7-parameter ΛCDM model with *WMAP* 7 year CMB data. We list the acceptance rate AR, total number of samples M and the normalised measure r of the variance on the parameter means (Eq. 2.7). Optimal MH uses a well-tuned trial distribution, non-optimal MH uses a trial distribution with optimal parameter widths but no correlation between parameters and results in a much longer convergence time (Sec. 2.3.2). NS with $N = 50, 100$ produces more accurate posterior statistics (as measured by r) in a shorter convergence time $N_{\rm like}$ than MH. Apart from non-optimal MH, Aff-Inv sampling requires the most likelihood evaluations due to a long burn-in period.

recombination, the optical depth to reionisation, the scalar spectral index and the logarithmic amplitude of the initial scalar fluctuations (at pivot scale $k_0 = 0.002$ Mpc^{-1}), respectively. This choice is convenient since the posterior is approximately Gaussian for each parameter. We also include an additional parameter $A_{\rm SZ}$ describing the amplitude of the Sunyaev-Zel'dovich effect, which increases the observed power in CMB fluctuations across a wide range of angular scales (Larson et al., 2011). The most recent all-sky CMB data now come from the *Planck* satellite (*Planck* Collaboration, 2015a). The resolution and sensitivity of the data available from *Planck* mean many more foreground and nuisance parameters are included in the generative models. This increased number of free parameters makes nested sampling less competitive in this case; see Sec. 2.4.2 for further discussion.

2.3.2 Comparison of the sampling methods

We plot the marginalised one-dimensional posterior probability distributions for each of the basis parameters in Fig. 4.7. We also show examples of two-dimensional posterior contours demonstrating parameter degeneracies seen in the data. All three methods show excellent mutual agreement and the derived means and constraints are fully consistent with those published by the *WMAP* collaboration[2]. We do not show $A_{\rm SZ}$, which is uniform across the entire prior range, because the *WMAP* data do not constrain this parameter.

[2]http://lambda.gsfc.nasa.gov/product/map/dr4/parameters.cfm

Table 2.1 lists the convergence time $N_{\rm like}$ for each of the three sampling routines based on the *WMAP* power spectrum data. The trial covariance matrix for Metropolis-Hastings was obtained as the output covariance matrix from a previous analysis of this data, and thus is close to optimal. We neglect the number of likelihood evaluations to obtain the trial distribution in the quoted result, and as such this number should be regarded as a lower bound. The initial sample point for Metropolis-Hastings was drawn from a non-degenerate Gaussian with widths close to the posterior width for each parameter. For affine-invariant MCMC we chose random positions for the walkers in a small 7-dimensional ball around the *WMAP* 7 year best-fit parameters.

We quote the number of likelihood evaluations for convergence in a single chain. However, to derive estimates of r for Metropolis-Hastings and nested sampling, we actually ran several chains. For affine-invariant MCMC we set $r = 0.01$ to determine $N_{\rm like}$ using Eq. 2.28.

Nested sampling on the *WMAP* data performs comparably to the toy-model case in 7 dimensions, since the posterior is approximately Gaussian for most parameters and the sampler does not require a particular tuning. By comparing the input trial covariance matrix for Metropolis-Hastings sampling with the posterior covariance we find the trial distribution is close to optimal; the fractional error on the input marginalised parameter widths is less than 2% for all parameters, and the absolute errors on the input trial correlation coefficients are less than 0.04. Thus the input covariance can be said to be nearly optimal. As such, we infer that the non-Gaussianity of the posterior (particularly for $A_{\rm SZ}$, which has a uniform distribution) is the principal reason for the factor of three increase in convergence time from the optimal toy-model case.

We also ran Metropolis-Hastings using a Gaussian trial distribution with optimal widths for each parameter but assuming no degeneracy between parameters; the number of likelihood evaluations for convergence was a factor of ~ 20 greater than in the optimal case, demonstrating the sensitivity of Metropolis-Hastings to the choice of trial distribution (see Table 1).

The nested sampling runs deliver much lower numerical uncertainty on the means, as measured by r, compared to both Metropolis-Hastings and affine-invariant MCMC. Small active sets (e.g., $N = 50$) estimate the mean well, yet the tails of the distribution are poorly estimated if N is chosen to be too small, which leads to very noisy estimates of the covariance and the evidence. The time for convergence on real data increases with the number of walkers for affine-invariant MCMC. This is because the burn-in becomes increasingly significant; by $n = 100$ burn-in is the

dominant phase of the algorithm. The increase in the convergence time, with respect to the toy-model case in 7 dimensions (Fig. 2.5), is consistent with the inclusion of burn-in, i.e., the autocorrelation times for the sampler are comparable between the two cases.

To determine the autocorrelation time τ for affine-invariant sampling we make a least-squares fit to the autocorrelation function which is estimated using the python script ACOR[3]. By comparison with estimates made using long chains of $O(10^4)$ steps, we find estimates for τ are order-of-magnitude correct within a few thousand steps (the burn-in samples must be removed for an accurate estimate). The estimated autocorrelation time is then inserted into Eq. 2.28 to give the convergence times $N_{\rm like}$ in Table 2.1.

Despite the large number of likelihood calls required for convergence with affine-invariant MCMC, the walker steps can be made in parallel and so if one has access to a computer cluster the effective time for a likelihood call can be lower by up to a factor $n/2$ (see discussion in Sec. 2.4.2). Thus if n is large, although burn-in is longer, affine-invariant MCMC may in practice perform parameter estimation most quickly out of the three samplers.

2.4 Discussion

In this section we discuss the performance of the three sampling techniques on toy-model and cosmological data. We discuss the tunings, convergence time, scope for parallelisation and robustness of each technique for different parameter estimation problems.

2.4.1 Tunings and convergence

The Metropolis-Hasting algorithm convergence time depends strongly upon the choice of trial distribution. For the results derived in Sec. 2.3 the trial covariance matrix was derived from the output of a previously converged chain for the same model, and so was close to optimal. A trial distribution which is not tuned to the particular problem – for example, an isotropic Gaussian trial used on a highly degenerate target distribution – will fail to return a good sampling of the posterior within a practicable time-scale. For a Gaussian likelihood in a D-dimensional parameter space one must specify the $D(D+1)/2$ independent elements of the target covariance matrix

[3]https://github.com/dfm/acor

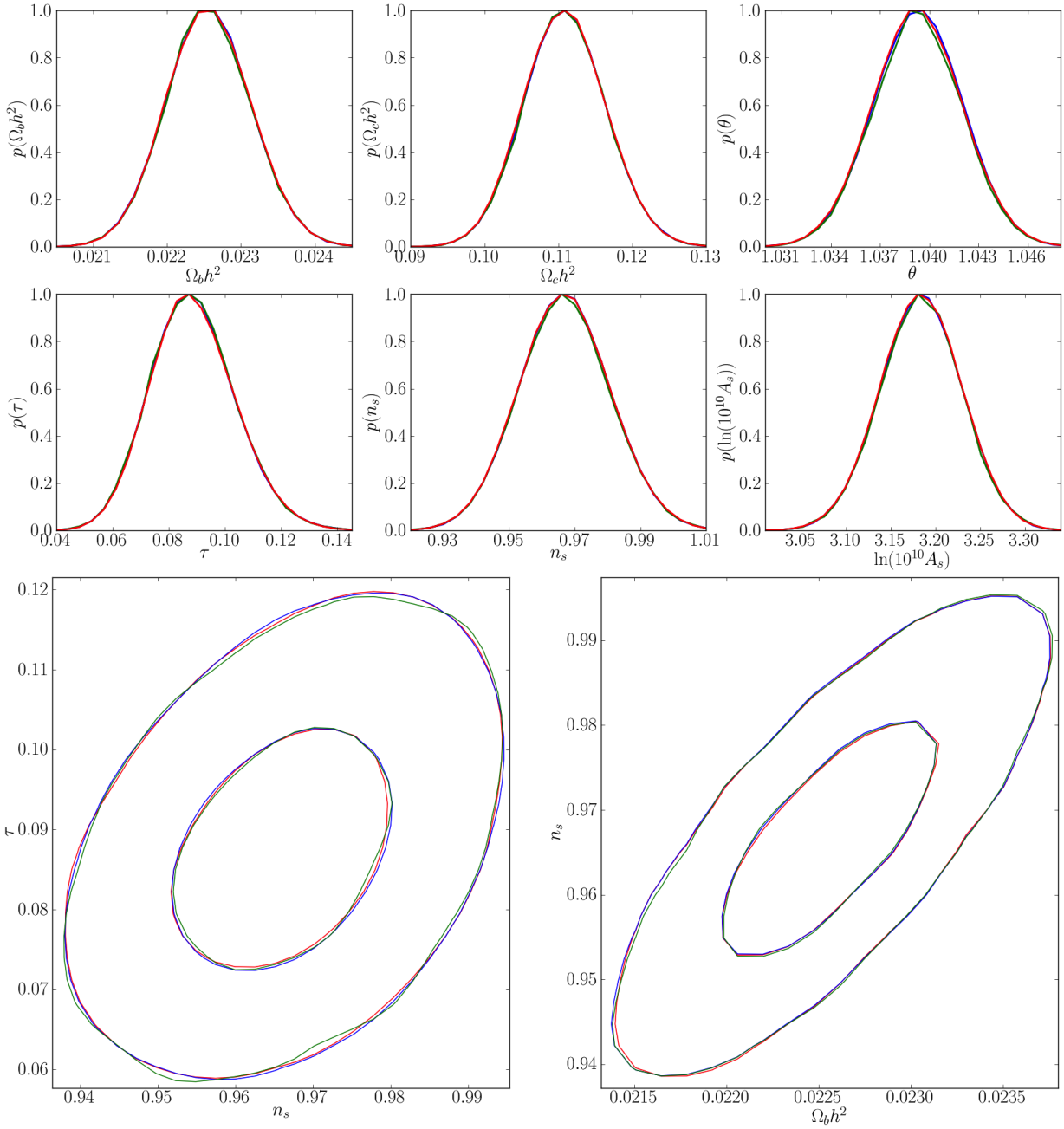

Figure 2.6: *Top panel:* Marginalised posterior probability distributions for each of the basis cosmological parameters (excluding $A_{\rm SZ}$). *Bottom panel:* Examples of posterior contours (68% and 95% CL shown) demonstrating parameter degeneracies. We show the distributions for each of the three sampling methods: Metropolis-Hastings (*blue*), nested sampling ($N = 500$, *green*) and affine-invariant MCMC ($n = 50$, *red*). There is excellent agreement between the three approaches. The plots are produced from chains ran past the time of convergence; for each method noisier posterior plots result from using only samples up to the convergence time.

to high accuracy to ensure an optimal convergence rate. This requires some knowledge of the underlying posterior distribution, which one may not have access to *a priori*. In practice one might run and analyse several short chains to provide a rough estimate of the covariance before the full problem is tackled. Another option is to use an adaptive Metropolis-Hastings procedure which allows the trial covariance to be learned during the initial phase of the exploration (e.g., Lewis, 2013). Non-Gaussian likelihoods clearly present an increased level of difficulty for the Metropolis-Hastings algorithm, and widely spread modes of a posterior may fail to be found at all. However, optimal Metropolis-Hastings on uni-modal Gaussian posteriors of a known shape, due to the simplicity and ease of implementation of the algorithm, is a plausible option when massive parallelisation is unfeasible or unnecessary.

Nested sampling requires no initial guess for the posterior covariance. Instead the algorithm firstly probes the full prior volume, followed by exploration of successively more likely regions of parameter space, adapting – at each step – to the inferred shape of the posterior. All samples taken outside the posterior, i.e., the first $N_{\rm tar} \sim N \ln(V_{\rm p}/V_{\rm t})$ iterations, are effectively burn-in, since they contribute negligible weight to the posterior. The only tunable hyper-parameters are the size of the active set N and the factor f by which the ellipsoidal bound is expanded to ensure unbiased results. We find $f = 1.06$ to be sufficient for all the likelihoods we have considered, while N may tailored to suit the application. A small active set will provide quick (but noisy) parameter constraints, while a large active set densely samples the posterior; Eq. 2.33 allows us to estimate the computational cost for either case. Because of the lack of tuning needed, nested sampling is, in some sense, more objective than Metropolis-Hastings, and perhaps more straightforward to set up on a new problem.

Approximating the likelihood contour as a single ellipsoid, as presented here, would give poor performance on multi-modal posteriors, since the acceptance rate would drop dramatically once the ellipsoid encountered multiple peaks. Cluster detection algorithms have been incorporated into nested sampling implementations to resolve this issue (Feroz et al., 2009, 2013). These extensions also allow nested sampling to be used on distributions containing curving degeneracies, but come with the added cost of a more complex implementation and interpretation.

Nested sampling performs well on the cosmological data set, producing high fidelity estimates of the posterior statistics in a small number of likelihood calls. Nested sampling is the only sampler

presented here that returns a robust estimate for the evidence. This is because MCMC methods do not explore the entire full prior volume in finite time, and the tails of the likelihood may give a large contribution to the evidence integrand. Nested sampling is an invaluable tool for model selection analyses (e.g., Mukherjee & Parkinson, 2008).

Affine-invariant ensemble MCMC also comes with just two tunable hyper-parameters: the number of walkers n and the step-size parameter a, which we set equal to 2 (Sec. 2.1.4). For any given walker W the position of the other walkers provides information on the step size and direction in which W should move to stay within the posterior bulk. A large number of walkers results in a long burn-in time, after which just a few steps, per walker, are enough for robust parameter estimation, since after burn-in each walker is an independent sample from the posterior. As pointed out in Foreman-Mackey et al. (2013), choosing a large number of walkers increases the scope for parallelisation (Sec. 2.4.2).

This sampling prescription, by construction, performs equally well on distributions with strong parameter degeneracies as on isotropic distributions, requiring no extra tuning from the user. Provided the initial walker positions are sufficiently dispersed throughout the prior volume, multi-modal and curving distributions are naturally probed by multi-particle samplers such as affine-invariant ensemble MCMC. Affine-invariant MCMC works well on the toy-model Gaussian likelihoods but requires many more likelihood calls than nested sampling on the real data set considered here. Despite this, parallelisation of the walker moves and its potential for mapping out multi-modal or curving distributions ensures this is a competitive sampling method.

2.4.2 Parallelisation

Metropolis-Hastings sampling can be parallelised by running multiple, independent chains on separate processors. If a single processor requires $N_{\rm like}$ steps for convergence, then $n_{\rm chain}$ processors can achieve the same accuracy r on the posterior means with $N_{\rm like}/n_{\rm chain}$ samples per processor, reducing the actual convergence time significantly. Alternatively, if all processors are ran for $N_{\rm like}$ steps, one obtains an accuracy $r/n_{\rm chain}$ on the posterior mean since there are a factor $n_{\rm chain}$ more samples, which produces less-noisy parameter constraints and smoother distributions.

The number of likelihood evaluations for convergence in nested sampling is set principally by the size of the active set N. One cannot run multiple chains to reduce the number of likelihood

evaluations per processor as in Metropolis-Hastings, since the active set is not independent from step to step. However, multiple processors can be used to produce dense samplings of the posterior by combining the samples from multiple chains with appropriate weights. As pointed out in Feroz et al. (2009), and implemented in MultiNest, an acceptance rate of less than unity means that parallel sampling from the ellipsoid at each iteration can significantly reduce the wall time of nested sampling. This allows it to be used in high-dimensional parameter spaces where otherwise the exponential scaling would rule it out as a practical sampling method (Eq. 2.33). This is utilised in *Planck* Collaboration (2014c), where the full generative model includes more than a dozen foreground and nuisance parameters, to explore spaces with $D > 20$ parameters.

Foreman-Mackey et al. (2013) implement parallelisation of the affine-invariant walker moves in their python code `emcee`. At each step the walkers can be partitioned into two sets containing $n/2$ walkers each; one set is then held fixed while the positions of the other set of walkers are updated, moving them according only to the positions of the fixed set of walkers. Each of the $n/2$ walker moves can be made in parallel because they are independent. The roles of the two sets are then reversed to complete one step. This provides scope for performing very fast parameter estimation, utilising up to $n/2$ processors. This parallelisation is investigated on cosmological data in Akeret et al. (2013). The dependence of trial moves on the other walkers' positions (*cf.* independence of Metropolis-Hastings chains) is precisely the property which allows affine-invariant MCMC to be effective on non-Gaussian and curving distributions.

2.5 Conclusions

We have outlined three sampling methods, presenting the algorithms themselves and discussing their practical implementation and associated stopping criteria. We present a new analytical result for the number of likelihood evaluations required for convergence in nested sampling. A comparison is made between the three sampling methods, focusing on performance on toy-model Gaussian likelihoods and a dataset derived from measurements of the CMB. We make use of and adapt the widely-used cosmological parameter estimation code CosmoMC, developing a general C/Cython bridge between the Fortran CosmoMC and python implementations of the sampling algorithms (Lewis & Bridle, 2002). This bridge allows us to attach essentially any python sampling

code to the well-tested and familiar CosmoMC, with little adaptation of the original Fortran code required.

All three sampling methods return posterior distributions which are mutually consistent and also consistent with those given in the cosmology literature. We find that, although optimised Metropolis-Hastings is in principle the fastest of the three methods for sampling probability distributions, in practice nested sampling and affine-invariant MCMC offer greater flexibility and robustness, requiring the adjustment of only 1 or 2 hyper-parameters. Metropolis-Hastings must be very precisely tuned to the target (posterior) distribution, and this is not possible even in principle when the distribution is non-Gaussian.

Nested sampling and affine-invariant MCMC naturally find and explore parameter degeneracies with no *a priori* input from the user. The number of likelihood evaluations for ellipsoidal nested sampling scales exponentially with the number of free parameters, ruling this out as a practical sampling technique for models with many free parameters unless one uses parallel sampling from the ellipsoid to reduce the wall time of each step. Affine-invariant ensemble MCMC requires the highest number of likelihood evaluations on the cosmological data set, but this sampling prescription is highly parallelisable. Indeed, its potential performance on non-Gaussian, multi-modal and curving distributions means this technique is extremely powerful. Nested sampling, although principally developed as a tool for Bayesian evidence calculation, is shown to deliver low-noise estimates for posterior statistics for low computational cost.

We note that extensions to single ellipsoidal nested sampling extend the range of application of this sampler to multi-modal and curving distributions. We have shown that, for parameter estimation, nested sampling should be adopted over the popular Metropolis-Hastings sampling method in many cases. For future cosmological analyses of multiple datasets with many nuisance parameters, nested sampling would efficiently explore *a priori* unknown degeneracies, although the dimensionality scaling of ellipsoidal nested sampling may be restrictive. New techniques such as PolyChord may offer an alternative, less-expensive procedure to draw from a restricted prior which does not have the strong dimensionality scaling of ellipsoidal nested sampling (Handley et al., 2015).

Chapter 3

The Atacama Cosmology Telescope: measuring radio galaxy bias through cross-correlation with lensing

In this chapter we correlate the positions of radio galaxies in the FIRST survey with the CMB lensing field measured by the Atacama Cosmology Telescope, over 470 deg^2, to determine the bias of these galaxies. This measurement is performed with a view to better understanding cosmology through the combination of different tracers of large-scale structure. We remove optically cross-matched sources below redshift $z = 0.2$ to preferentially select active galactic nuclei (AGN). We measure the angular cross-spectrum $C_l^{\kappa g}$ at 4.4σ significance in the multipole range $100 < l < 3000$, corresponding to physical scales between $\approx$ 2–60 Mpc at an effective redshift $z_{\rm eff} = 1.5$. Modelling the AGN population with a redshift-dependent bias, the cross-spectrum is well fit by the *Planck* best-fit ΛCDM cosmological model. Fixing the cosmology and assumed redshift distribution of sources, we fit for the overall bias model normalization, finding $b(z_{\rm eff}) = 3.5 \pm 0.8$ for the full galaxy sample, and $b(z_{\rm eff}) = 4.0 \pm 1.1$ (3.0 ± 1.1) for sources brighter (fainter) than 2.5 mJy. This measurement characterizes the typical halo mass of radio-loud AGN: we find $\log(M_{\rm halo}/M_\odot) = 13.6^{+0.3}_{-0.4}$.

This chapter is organised as follows. We introduce the measurement in Sec. 3.1. We describe the lensing and radio data, and the cross-correlation analysis methods, in Sec. 3.2. The results

and discussion are presented in Sec. 3.3, with further interpretation of the AGN bias in Sec. 3.3.1. We conclude in Sec. 3.4.

3.1 Introduction

Radio galaxies trace the large-scale structure in the Universe. This has been measured with large-area surveys such as FIRST, WENSS, NVSS, and SUMSS (Becker et al., 1995; Rengelink et al., 1997; Condon et al., 1998; Bock et al., 1999); for an overview see de Zotti et al. (2010). The angular clustering of these galaxies has been measured by Cress et al. (1996); Magliocchetti et al. (1998); Blake & Wall (2002); Overzier et al. (2003); Blake et al. (2004a); Lindsay et al. (2014b). The clustering of radio galaxies will soon be measured over much larger volumes of the universe with the Square Kilometer Array (SKA) and its precursors, allowing cosmological effects such as dark energy, modified gravity and non-Gaussianity to be probed (e.g., Carilli & Rawlings, 2004; Blake et al., 2004b; Raccanelli et al., 2012; Camera et al., 2012; Maartens et al., 2013; Norris et al., 2013; Jarvis et al., 2015; Abdalla et al., 2015; Maartens et al., 2015; Santos et al., 2015).

The bias b of a large-scale structure tracer relates overdensities of that tracer δ to overdensities of the underlying dark matter field $\delta_{\rm DM}$ (Sec. 1.5):

$$\delta = b\delta_{\rm DM}. \tag{3.1}$$

Radio-selected galaxies broadly contain two populations: high-redshift active galactic nuclei (AGN) and low-redshift star-forming galaxies (Condon et al., 2002). AGN dominate the radio emission at high flux ($\gtrsim 1$ mJy) and are highly biased, their hosts being among the most massive galaxies in the early universe (e.g., Jarvis et al., 2001a; Rocca-Volmerange et al., 2004; Seymour et al., 2007; de Zotti et al., 2010; Fernandes et al., 2015). Their bias depends strongly on galaxy mass and redshift (e.g., Seljak & Warren, 2004), and is poorly constrained particularly at high redshift where few optical counterparts are observed. Some progress has been made by identifying redshifts spectroscopically: using Galaxy And Mass Assembly (GAMA) data the bias of FIRST radio galaxies was measured at $z \approx 0.34$ over 200 deg^2 to the 10% level (Lindsay et al., 2014b). On a smaller, square degree region, clustering measurements using data from the Very Large Array (VLA) and VISTA Deep Extragalactic Observations (VIDEO; Jarvis et al., 2013) were used to

show evidence for a strongly increasing bias at $z > 2$ (Lindsay et al., 2014a).

An alternative way to constrain bias is through cross-correlation of the tracer fluctuations with gravitational lensing due to large-scale structure. In particular, the lensing of the cosmic microwave background (CMB) measures the integrated matter fluctuations to $z \approx 1100$. As we will show, the high-redshift radio source distribution overlaps strongly with the broad CMB lensing kernel. Cross-correlations between the CMB and other tracers of large-scale structure have been reported by e.g., Smith et al. (2007); Hirata et al. (2008); Feng et al. (2012); Bleem et al. (2012); *Planck* Collaboration (2014b); van Engelen et al. (2014); Fornengo et al. (2014); Bianchini et al. (2014). The *Planck* Collaboration (2013b) detect the correlation of lensing with radio galaxies from NVSS at 20σ. Sherwin et al. (2012) correlate lensing measurements from the Atacama Cosmology Telescope (ACT) with optically-selected quasars from the Sloan Digital Sky Survey (SDSS), measuring a bias $b = 2.5 \pm 0.6$ at an effective redshift $z \approx 1.4$. Geach et al. (2013) correlate lensing from the South Pole Telescope (SPT) with quasars selected from the Wide-field Infrared Survey Explorer (WISE), measuring a bias $b = 1.61 \pm 0.22$ at $z \approx 1.0$.

In this chapter we measure the angular cross-spectrum $C_l^{\kappa g}$ between the lensing convergence estimated from ACT with the FIRST radio source overdensity. We use lensing maps from the three-year ACT Equatorial survey (Das et al., 2013) together with the first-season ACTPol survey (Naess et al., 2014; Madhavacheril et al., 2014; van Engelen et al., 2014). We consider 36,000 radio sources with flux brighter than 1 mJy, and remove optically cross-matched sources from the Sloan Digital Sky Survey (York et al., 2000) at $z < 0.2$ to preferentially select AGN, discarding the majority of low-redshift star-forming galaxies. We use this measurement to estimate the bias normalization, assuming a fixed cosmological model, and using a redshift distribution and bias-evolution model from the simulated radio catalogue of the SKA Design Study (SKADS, Wilman et al., 2008). We measure $C_l^{\kappa g}$ across a wide range of scales ($100 < l < 3000$) and consider various splits of the radio sources to investigate redshift and flux dependence of the bias.

3.2 Data and analysis

3.2.1 ACT and ACTPol

The Atacama Cosmology Telescope (ACT) is located at an altitude of 5190m in Parque Astronómico Atacama in Northern Chile. The telescope and its current polarisation-sensitive receiver, ACTPol, are described in Sec. 1.9 and Niemack et al. (2010). The two seasons of ACT temperature data and the ACTPol first-season temperature and polarisation data used in this analysis are presented in Das et al. (2013) and Naess et al. (2014). Lensing by large-scale structure induces coupling of otherwise independent temperature and polarisation modes (Sec. 1.4). We construct estimators of the lensing convergence from quadratic combinations of temperature and polarisation maps in Fourier space, following the methodology of Hu & Okamoto (2002). We use the same lensing convergence maps and Monte-Carlo simulations described by Das et al. (2011) and van Engelen et al. (2014).

In this analysis we use two ACT datasets. The first is the ACT Equatorial data which spans a thin strip along the celestial equator with an area of 300 deg^2. This strip is partitioned into six approximately equal area patches over which we compute the cross-spectrum separately then average (weighting by patch area) for the final result. We lose negligible information at the scales of interest and it allows for patch to patch consistency checks. The effective white-noise component of the two-season co-added data is 18 μK-arcmin.

We also fold in the three ACTPol 'deep' fields from the first-season dataset, labelled D1, D5 and D6, with a temperature white-noise component of $16.2, 13.2, 11.2$ μK-arcmin, respectively, over a total area of 206 deg^2 (37 deg^2 of which overlaps with the ACT Equatorial strip). All maps in this analysis use $0.5' \times 0.5'$ pixels, and the patch footprints are shown in Fig. 3.1.

For each ACTPol patch we use the minimum-variance (MV) linear combination of the reconstructed convergence maps estimated from each quadratic pair (TT, TE, EE, EB). Following the The POLARBEAR Collaboration (2013), van Engelen et al. (2014) and Story et al. (2015) we perform the combination in Fourier space, weighting each convergence map by the mode-dependent inverse-variance noise to obtain the MV combination. All lensing convergence maps are mean-field subtracted to remove the lensing-like effect at large scales of mode-coupling from the windowing of the temperature and polarisation fields (Sec. 1.4).

As described in van Engelen et al. (2014), an apodisation window is applied to the ACT and ACTPol temperature and polarisation maps prior to lensing reconstruction. This windowing operation includes a cosine taper at the map edges to remove discontinuous edges and weighting by the pixel hitmap to optimize the signal-to-noise of the reconstruction. The resulting quadratic estimator reconstruction is therefore also windowed, resulting in a scale-dependent suppression of power. Following Bleem et al. (2012), Sherwin et al. (2012), van Engelen et al. (2014) and Hand et al. (2015) we use realistic Monte-Carlo simulations to calculate the transfer function correction by computing the mean cross-spectrum between noiseless lensing realizations and their corresponding reconstructions within the lensing pipeline. This correction ($< 5\%$ for ACTPol, $\approx 10\%$ for ACT) is then applied to the maps when computing the data cross-spectrum to account for the suppression of power due to windowing.

3.2.2 FIRST

The FIRST survey was carried out between 1993 and 2011 at 1.4 GHz with the VLA in B configuration (Becker et al., 1995). The final catalogue contains 946,432 sources covering 10,575 deg^2, with an angular resolution of $5.4''$ (FWHM) and to a completeness of 95% at flux $S_{1.4\mathrm{GHz}} > 2$ mJy (Helfand et al., 2015) . The oblique decision-tree program developed by the FIRST survey team (White et al., 1997) determines the probability that each catalogue entry is the result of a spurious sidelobe response to a nearby bright source. We exclude entries with a sidelobe probability of > 0.1, leaving 720,219 sources above 1 mJy.

To address the issue of extended radio sources resulting in multiple detections for one host galaxy, perhaps none of which corresponds to the core itself (and therefore any associated optical source), we have followed Cress et al. (1996) in applying a collapsing radius of $72''$ (0.02 deg) to the FIRST catalogue. Any FIRST sources within this radius of one another are grouped and combined to form a single entry, positioned at the flux-weighted average coordinates of the group, and attributed with their total flux density. Around 32% of all FIRST sources are collapsed (in groups of average size 2.3 sources per group), forming 17% of the resulting catalogue. These multiple-component sources will come from AGN, which dominate the source population at high flux density and high redshift; at lower flux density and redshift, starbursts and normal star-forming galaxies (SFGs) are increasingly dominant (Condon et al., 2002).

The redshifts of individual sources are not determined by FIRST, but can be found by identifying counterparts in SDSS, which gives redshifts for the brighter, nearby sources in the FIRST sample. The closest sources are most likely star-forming galaxies (e.g., Condon et al., 2002; Wilman et al., 2008); by removing them we simplify the interpretation of the measurement as a constraint on the bias of the dominant astrophysical population (i.e., AGN).

We perform the optical cross-matching by initially taking all sources in the catalogue which lie in the ACT and ACTPol patches described above ($\approx$ 38,000 sources). AGN dominate the radio luminosity function for $L_{1.4\mathrm{GHz}} > 10^{23}$ W Hz^{-1} (Condon et al., 2002; Jarvis & Rawlings, 2004; Mauch & Sadler, 2007). Given the flux limit of FIRST (1 mJy), and assuming a spectral index of $\alpha = 0.8$ ($S_\nu \propto \nu^{-\alpha}$) for the AGN, this luminosity threshold corresponds to sources above redshift $z = 0.2$. We identify optical matches to the radio sources within SDSS, treating as reliable all matches within $2''$ of the radio source, following Lindsay et al. (2014b). Given the density of SDSS sources, the level of spurious optical matches identified with this technique is below 2%. A fraction of 0.27 of the FIRST sources in the lensing fields have an optical match obtained in this manner. We remove all sources with a known redshift below $z = 0.2$, constituting 18% of the sources with a reliable redshift, or 5% of the total number of sources. Given the small fraction of sources removed, this procedure has only a small effect on the results (Sec. 3.3). The final sample comprises $\approx$ 36,000 sources with a mean angular density of 71 sources deg^{-2}.

Within each ACT and ACTPol patch a corresponding map of the overdensity of sources g is produced in a similar way to Sherwin et al. (2012) and Geach et al. (2013). We create a map at the same pixelisation as the lensing map and define the radio-galaxy overdensity map g by

$$g_i = \frac{n_i}{\bar{n}} - 1, \tag{3.2}$$

where n_i is the number of sources in each pixel and $\bar{n}$ is the mean number of galaxies per pixel. This overdensity map is then smoothed with a Gaussian with FWHM of $2'$ to obtain a well-defined pixel window function.

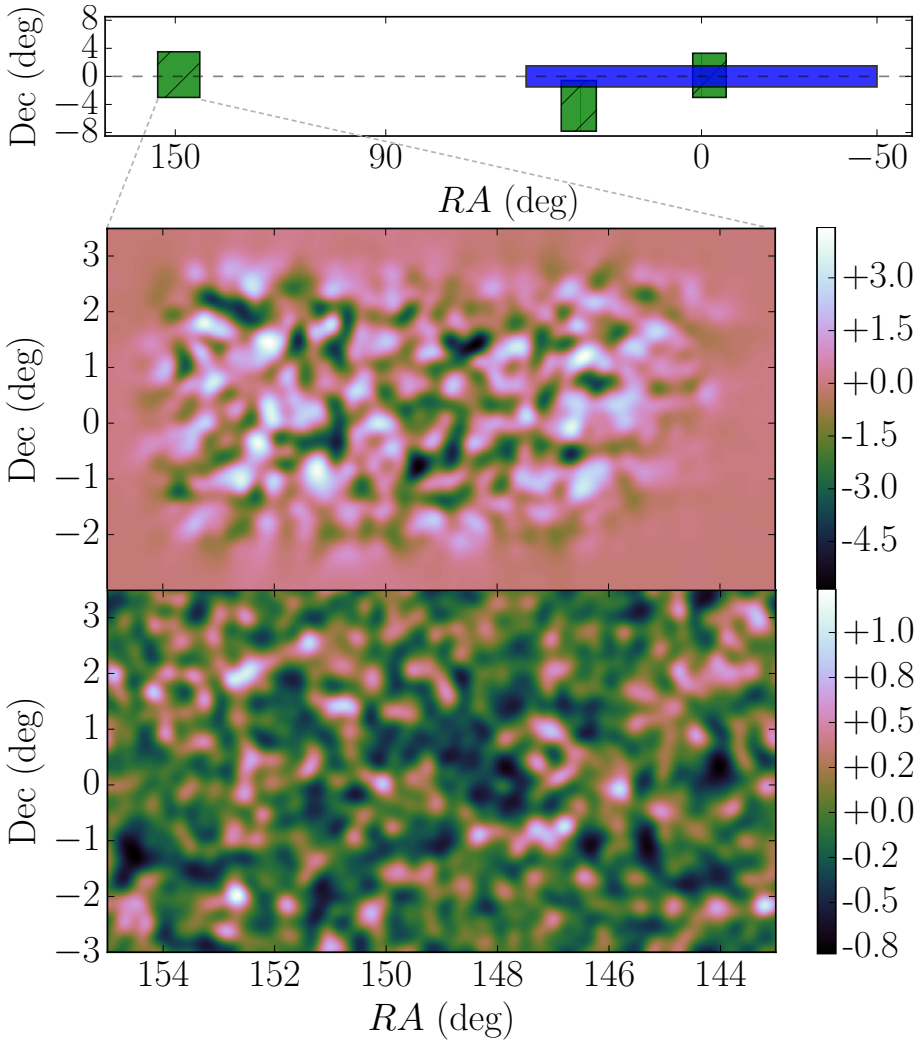

Figure 3.1: *Top panel:* Footprint of the patches used in this analysis: ACT (*blue, filled*) and ACTPol (*green, hatched*, left to right: D1, D6, D5). *Middle panel:* ACTPol D1 lensing convergence map κ smoothed to suppress power below $20'$ scales. The spatial modulation, primarily due to the windowing of the temperature and polarisation maps by the pixel weight map, is evident. The ACTPol lensing convergence is noise dominated for scales $\lesssim 1$ degree. *Bottom panel:* The FIRST overdensity field g over the same patch, smoothed to the same scale, is noise dominated at all scales.

3.2.3 Analysis methods

We compute the cross-spectrum between the lensing convergence from ACT and ACTPol with the FIRST radio galaxy overdensity.

Following the procedures outlined in Das et al. (2011) and Hand et al. (2015), we correct for mode-coupling induced by windowing in real space and from applying annular binning in Fourier space, computing an unbiased estimator of the binned cross-spectrum $C_b^{\kappa g}$ (Chapter 5). The binning we adopt is given in Table 3.1.

To determine the full bandpower covariance matrix we cross-correlate realistic simulations of the reconstructed lensing fields with the radio source maps (which are in principle uncorrelated). Production of these realistic simulations is described in Das et al. (2011) and van Engelen et al. (2014).

This procedure ignores the cosmic variance contribution to the uncertainties in the data coming from the correlated part of the two maps, $C_l^{\kappa g}$. We neglect this as both maps are noise-dominated at the relevant scales for this analysis. Bin-to-bin correlations are $< 10\%$ for all off-diagonal elements of the covariance matrix. We also check that the mean cross-spectrum is consistent with null (Figure 3.2), confirming that our pipeline does not induce spurious cross-power in the absence of correlation.

Approximately 50% of the ACTPol D5 patch and 15% of D6 overlap with the ACT Equatorial strip (Fig. 3.1). There is therefore a correlation between the ACT and ACTPol cross-spectra, as common CMB modes in the primary temperature map have been used to reconstruct the lensing convergence over these regions. Noiseless temperature maps from ACT and ACTPol, and negligible polarisation information from ACTPol, would result in a perfect correlation between the reconstructed convergence maps. However, this overlapping area represents 37 deg^2 of the total 470 deg^2 of this analysis, and hence at most a 4% overestimate of the detection significance, which we neglect given the statistical errors. We thus average the ACT and ACTPol data cross-spectra with inverse-variance weighting.

In order to check for bias in the cross-spectrum estimator, we ran 500 pairs of simple simulated convergence and radio density maps through the cross-correlation pipeline, generating new *correlated* simulations. To obtain these pairs we draw as signal maps aperiodic correlated Gaussian realizations from power spectra obtained assuming *Planck* best-fit cosmological parameters

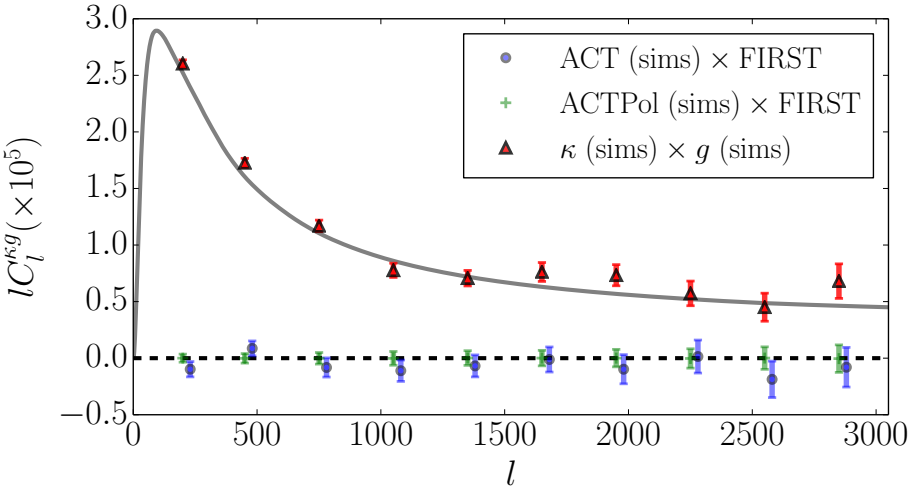

Figure 3.2: Pipeline validation: Mean cross-spectrum $lC_l^{\kappa g}$ of the FIRST radio source map with ACTPol lensing simulations (*green dashes*, $N_{\rm sims} = 2048$) and ACT lensing simulations (*blue circles*, $N_{\rm sims} = 480$) as described in 3.2.3. We displace the ACT points by $\Delta l = 30$ to the right for visual clarity. The measurements are consistent with null, demonstrating that our pipeline does not induce spurious cross-power in the absence of correlation. Error bars shown are the diagonal components of the empirical covariance matrix derived from the same Monte Carlo simulations, scaled appropriately by $\sqrt{N_{\rm sims}}$. We also show the recovered mean cross-spectrum from realistic *correlated* simulations (*red triangles*, Sec. 3.2.3). We cross-correlate input convergence maps, which have added scale-dependent Gaussian noise, with correlated realizations of a galaxy field. This demonstrates that our pipeline is able to recover in an unbiased fashion a known input cross-spectrum (although we note this does not test the lensing reconstruction pipeline, for which we refer to the systematic tests in van Engelen et al., 2014). The generative model for the cross-spectrum is not the fiducial cross-spectrum, but this is unimportant for the purposes of this test.

and a fiducial bias model and source distribution for the radio galaxies (Kamionkowski et al., 1997; *Planck* Collaboration, 2013a). We add Gaussian noise realizations to the convergence maps, appropriate for the temperature sensitivity of ACT (Sec. 3.2.1), using the formalism of Hu & Okamoto (2002) to calculate the reconstruction noise. ACTPol maps are less noisy, but the precise noise level is unimportant for this test. For each pixel i in the radio signal map g we draw a Poisson random variable X_i with mean $\bar{n}(1 + g_i)$, where $\bar{n}$ is the average number of sources per pixel. We set $\bar{n} = 71$ sources deg^{-2} to reflect the source density in the data. We then redefine $g_i \leftarrow X_i/\bar{n} - 1$ and finally smooth the resulting map with a Gaussian beam of FWHM $2'$.

These simulated maps, by construction, have signal, noise and correlation properties which mimic the data, although they do not have the full spatially-anisotropic noise properties. These lensing simulations have not been processed through the lensing reconstruction pipeline, but here we use them simply for checking bias in the cross-correlation pipeline. We refer to the systematic

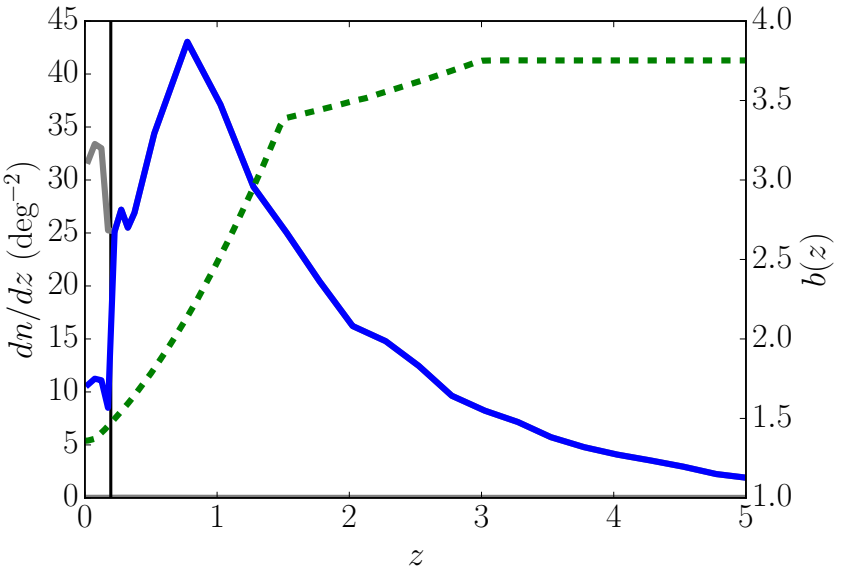

Figure 3.3: *Green dashed:* Fiducial bias model $b(z)$, constructed as a weighted average of the constituent source population bias models (Sec. 3.2.2). *Blue solid:* Source redshift distribution model as derived from SKADS and including the cut of a fraction of the $z < 0.2$ sources. *Grey solid:* As previous but without the redshift cut. *Vertical black line:* $z = 0.2$.

tests in van Engelen et al. (2014) for checks of the lensing pipeline. We find that we do not require apodization of the maps to produce the observed unbiased results; the mean auto- and cross-spectra of these simulations are consistent with the assumed input spectra (Fig. 3.2).

3.2.4 Modelling

The theoretical cross-spectrum can be written under the Limber approximation as

$$C_l^{\kappa g} = \int_0^\infty dz \frac{H(z)}{\chi^2(z)} W_\kappa(z) W_g(z) P\left(\frac{l}{\chi(z)}, z\right), \tag{3.3}$$

where $H(z)$ is the Hubble parameter, $\chi(z)$ is the comoving distance to redshift z, $P(k, z)$ is the non-linear matter power spectrum (wavenumber $k = l/\chi$) and $\{W_i\}$ are the appropriate kernels for the two dark matter probes κ, g which are plotted in Fig. 3.4 (Sec. 1.1). This expression is analogous to Eq. 1.20 for a flat universe. The dominant term in W_g is directly proportional to the tracer bias $b(z)$ multiplied by the normalised source redshift distribution dn/dz:

$$W_g(z) = b(z)\frac{dn}{dz} + M(z), \tag{3.4}$$

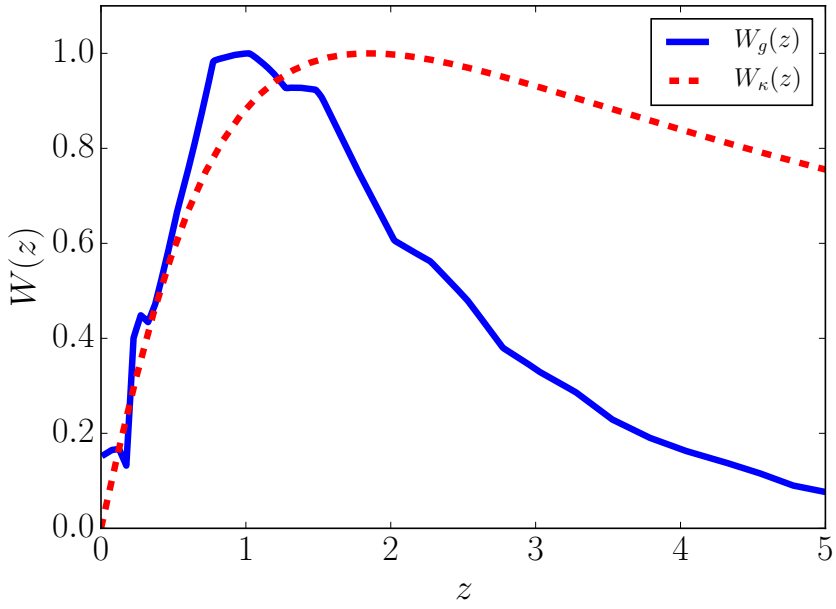

Figure 3.4: Cross-spectrum kernels for the radio source overdensity (g) and CMB lensing convergence (κ), normalised to a unit maximum; see Eq. 4.3. The extensive overlap between the kernels implies a large correlation coefficient between the two fields.

where $M(z)$ is a sub-dominant contribution from the magnification bias. The magnification bias term is independent of the tracer bias, and the full expression (given in e.g., Sherwin et al., 2012) is calculated directly, using the fiducial redshift distribution of FIRST sources in our sample and the inferred slope ($s \approx 0.3$) at the faint end of the luminosity function from the SKA Design Study radio catalogue (see below). The magnification bias contributes a small (<6%), negative contribution to the total cross-spectrum across all relevant scales, and we include it throughout for modelling of the cross-spectrum. A rescaling of the bias amplitude therefore corresponds linearly to a rescaling of the cross-spectrum $C_l^{\kappa g}$. We compute the theory $P(k, z)$ using best-fit *Planck* cosmological parameters, including non-linear corrections using CAMB with HALOFIT (Lewis et al., 2000; Smith et al., 2003; Takahashi et al., 2012).

We use the SKA Design Study (SKADS) simulated radio continuum catalogue to construct a fiducial bias model $b(z)$ and redshift distribution dn/dz for the $S_{1.4\mathrm{GHz}} > 1$ mJy radio sources (see Wilman et al., 2008, 2010, for details). We neglect any scale-dependence in the bias model; see Sec. 3.3 for further discussion. The simulation lacks the mass resolution to directly resolve galaxy- and group-sized haloes for a robust implementation of the galaxy clustering, but the source counts, redshift distribution, and variations in space density are defined by extrapolating

observed luminosity functions, and implementing a bias model, for each of five individual radio populations: AGN (FRI and FRII types, radio-quiet quasars), normal star-forming galaxies and starburst galaxies. These populations are assigned a single halo mass each, used to define $b(z)$ as described by Mo & White (1996), with the bias held fixed above a particular redshift to prevent unphysical clustering where the bias is poorly constrained observationally (see Figure 3 of Raccanelli et al., 2012). The simulated catalogue informs us how the relative numbers of these populations are expected to evolve with redshift, and how the observed bias evolves accordingly for a mixed sample of sources.

Fixing the bias above a particular redshift for each population is clearly a crude model, and is necessitated by the (itself unphysical) assumption of a fixed halo-mass for each population adopted in Wilman et al. (2008); this must break down at high redshifts. Little is known about the high-redshift bias of radio sources, and better-motivated models are lacking (Sec. 3.1). If the data in this study, and studies like this, strongly prefer a much higher bias at high redshift, this will be reflected in the inferred bias normalization or a poor model fit.

By comparing the distribution of known source redshifts (Sec. 3.2.2) with the SKADS simulation, we find an estimated 66% of low-redshift sources are removed by the $z < 0.2$ cut. To construct dn/dz we therefore weight $z < 0.2$ sources by 0.34 relative to higher-redshift sources, similar to the approach of Lindsay et al. (2014a). After the redshift cut $\approx 96\%$ of the sources in our sample are expected to be AGN, with a $\approx 4\%$ contamination fraction of star-forming and starburst galaxies, and we estimate the final sample to have a median redshift $\tilde{z} = 1.3$. These fiducial models are shown in Figure 3.3.

We emphasise that $C_l^{\kappa g}$ is a function of the product $b(z)dn/dz$, and so inference about the tracer bias is contingent on knowledge of the underlying redshift distribution. We discuss the limitations of assuming a fixed redshift distribution, and the effect of not fully removing the low-redshift sources, in Sec. 3.3.3.

Finally, we bin the theoretical cross-spectrum as for the data, accounting for the mode-coupling matrix of each patch (Das et al., 2011). We then compare the model to the data using a Gaussian likelihood, accounting for the full bin-to-bin covariance as derived from Monte Carlo simulations (Sec. 3.2.3). We primarily fit for an overall scaling A to the fiducial bias model, such that $b(z) \rightarrow Ab(z)$.

Bin b	$[l_{\rm min}$, $l_{\rm max}]$	$C^{\kappa g}_{b,\rm ACT}$ $(\times 10^7)$	$C^{\kappa g}_{b,\rm ACTPol}$ $(\times 10^7)$	$C^{\kappa g}_{b,\rm comb}$ $(\times 10^7)$
200	[100,300]	1.76 ± 0.74	2.04 ± 0.80	1.89 ± 0.54
450	[301,600]	0.59 ± 0.32	-0.20 ± 0.42	0.30 ± 0.25
750	[601,900]	0.57 ± 0.25	0.32 ± 0.31	0.48 ± 0.19
1050	[901,1200]	0.28 ± 0.20	0.38 ± 0.26	0.32 ± 0.16
1350	[1201,1500]	0.17 ± 0.16	-0.08 ± 0.22	0.09 ± 0.13
1650	[1501,1800]	0.04 ± 0.15	0.01 ± 0.19	0.03 ± 0.12
1950	[1801,2100]	0.11 ± 0.14	-0.00 ± 0.18	0.07 ± 0.11
2250	[2101,2400]	0.07 ± 0.14	-0.01 ± 0.17	0.04 ± 0.11
2550	[2401,2700]	0.06 ± 0.14	0.19 ± 0.18	0.11 ± 0.11
2850	[2701,2999]	-0.01 ± 0.13	-0.23 ± 0.19	-0.08 ± 0.11

Table 3.1: The measured cross-spectrum $C^{\kappa g}_b$ between FIRST radio source overdensity with ACT and ACTPol lensing convergence. The bins are chosen to be wide enough that correlations are small ($< 10\%$, Sec. 3.2.3), but narrow enough to resolve structure in the cross-spectrum.

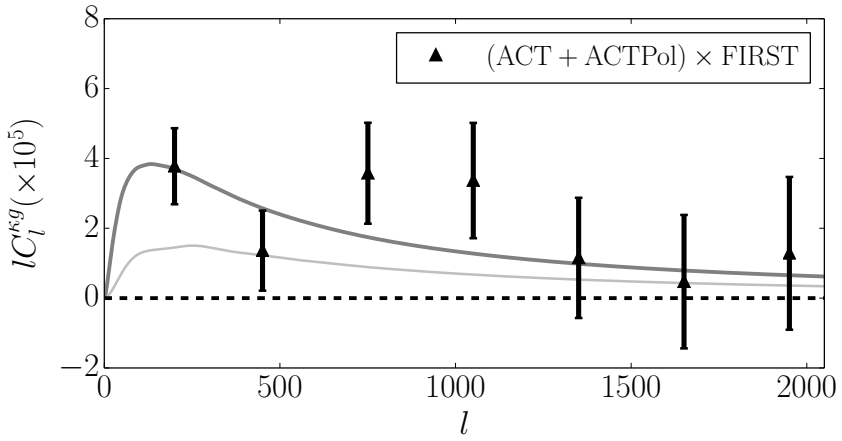

Figure 3.5: Data cross-spectrum $lC^{\kappa g}_l$ for (ACT+ACTPol)×FIRST. *Dark grey, solid*: the best-fit cross-spectrum. *Light grey, solid*: the contribution from $z > 1.5$ sources. We restrict the plot to $l < 2000$ where the signal-to-noise dominates. We show as error bars the diagonal components of the empirical covariance matrix derived from Monte Carlo simulations (Sec. 3.2.3). Scaling the amplitude of the fiducial bias model, the combined significance of the bias detection is 4.5σ (Sec. 3.3).

3.3 Results

The cross-spectra for ACT×FIRST, ACTPol×FIRST, and their combination are shown in Fig. 3.5–3.6 and reported in Table 3.1. We find $A_{\rm ACT} = 1.22 \pm 0.31$ and $A_{\rm ACTPol} = 0.85 \pm 0.36$, with a combined constraint $A = 1.06 \pm 0.24$. The goodness-of-fit statistics for these best-fit models are reported in Table 2, and account for the small (<10%) bin-to-bin correlations (Sec. 3.2.3). This amplitude is consistent with the expected bias from the radio simulations, and we interpret the

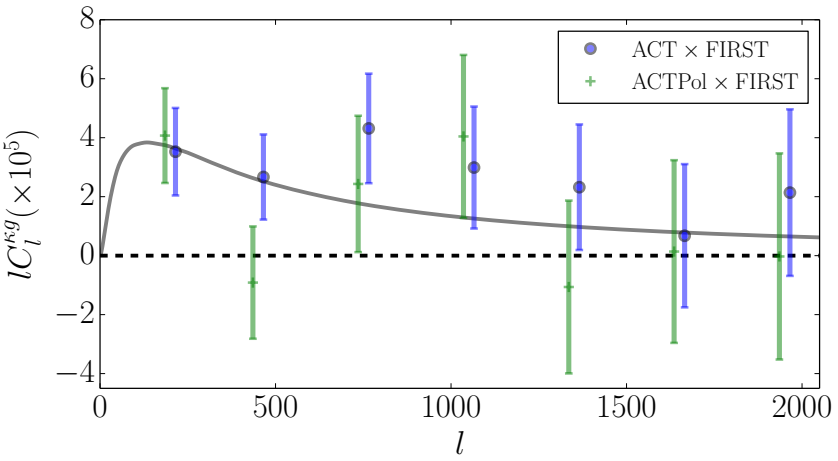

Figure 3.6: Data cross-spectrum $lC_l^{\kappa g}$ for ACT×FIRST (*blue*) and ACTPol×FIRST (*green*). *Dark grey, solid*: the best-fit cross-spectrum for the combined data. *Light grey, solid*: the contribution from $z > 1.5$ sources. The ACT and ACTPol points have been displaced to the right and left by $\Delta l = 15$, respectively, for visual clarity. We show as error bars the diagonal components of the empirical covariance matrix derived from Monte Carlo simulations (Sec. 3.2.3).

	A	S/N	χ^2 (ν)	PTE
ACT	1.22 ± 0.31	3.9	3.2 (9)	0.96
ACTPol	0.85 ± 0.36	2.4	7.2 (9)	0.62
Comb.	1.06 ± 0.24	4.5	11.0 (19)	0.92

Table 3.2: Results showing the bias amplitude A relative to the fiducial model of Fig. 3.3. We also quote the signal-to-noise ratio S/N, chi-squared values at the best-fit χ^2, the number of degrees of freedom ν and the probability to exceed this χ^2 (PTE) under the assumption of the best-fitting model.

result further in Sec. 3.3.1.

The parameter A only scales the bias-dependent part of the theoretical model and not the magnification bias term. To assess the overall detection significance we rescale the amplitude of the total theoretical cross-spectrum by a free parameter α: $C_l^{\kappa g} \to \alpha C_l^{\kappa g}$. This is equivalent to equally rescaling both terms in Eq. 5.3. The combined data require $\alpha = 1.06 \pm 0.24$, and the cross-spectrum is detected at $\sqrt{\chi^2_{\rm null} - \chi^2_{\rm bf}} = 4.4\sigma$ statistical significance. Here $\chi^2_{\rm null} = 31.4$ is the chi-squared value of the fit under the null hypothesis (no cross-correlation) and $\chi^2_{\rm bf} = 11.0$ is the chi-squared value for the best-fit model (number of degrees of freedom $\nu = 19$).

The mean cross-spectrum of ACT and ACTPol Monte-Carlo simulations with the FIRST dataset is shown in Fig. 3.2 and is consistent with null (Sec. 3.2.3). These simulations reproduce the amplitude and statistics of the lensing field but not the true mass distribution on the sky.

We further test our pipeline, checking for spurious correlations present only in the lensing and galaxy data, by performing two additional null tests.

Firstly, we randomly permute the six FIRST patches within the equatorial strip, such that all patches are moved from their true position, with respect to the fixed ACT patches. We recompute the cross-spectrum, shown in Fig. 3.7. Fitting the normalization of the fiducial bias model A to these data, we obtain $A_{\rm ACT,shuffle} = -0.18 \pm 0.31$, consistent with null. The chi-squared value of the null hypothesis is 18.2 for $\nu = 10$ degrees of freedom, or a probability-to-exceed the observed chi-squared of 5%.

Secondly, we make reconstructions of the lensing field where the deflection field has been redefined – as the curl of the lensing potential – and hence the expected 'convergence' Ω is zero, following Sherwin et al. (2012) and van Engelen et al. (2014). These maps contain reconstruction noise but should contain no common signal with the overlapping galaxy field. We recompute the cross-spectrum of the lensing curl maps Ω with the FIRST maps, shown in Fig. 3.8. Fitting the normalization of the fiducial bias model A to these data, we obtain $A_\Omega = 0.19 \pm 0.17$. Error bars are calculated from the data auto-spectra using the Knox formula (Knox, 1995). The chi-squared value of the null hypothesis is 21.8 for $\nu = 19$ degrees of freedom, or a probability-to-exceed of 0.29, consistent with a null result.

Removal of the known $z < 0.2$ sources, which constitute $\approx 5\%$ of the FIRST sample (Sec. 3.2.2), has only a small effect on the inferred bias amplitude: without removal we find a combined constraint $A_{\rm noZcut} = 1.08 \pm 0.24$, consistent with expectations given the shape of the lensing kernel and low bias of SFGs at low redshift.

Restricting the analysis to multipoles $l < 900$, representing linear scales at the effective redshift $z_{\rm eff} = 1.5$, we find $A = 1.01 \pm 0.24$, giving a detection significance of 4.2σ. We conclude that any scale dependence of the bias on non-linear scales is unimportant for this measurement.

3.3.1 AGN bias

The primary results of this section are shown in Figure 3.9. We determine the redshifts to which our measurement is most sensitive by considering the kernel $C_l^{\kappa g}(z) \equiv W_\kappa(z) W_g(z) P(l/\chi(z); z)$ of the theoretical cross-spectrum (Eq. 4.3; shown in Fig. 3.10). At $l = 200$, where the signal-to-noise of the cross-spectrum peaks, the mean redshift of the kernel is $\int z C_{200}^{\kappa g}(z) dz / C_{200}^{\kappa g} = 1.5$.

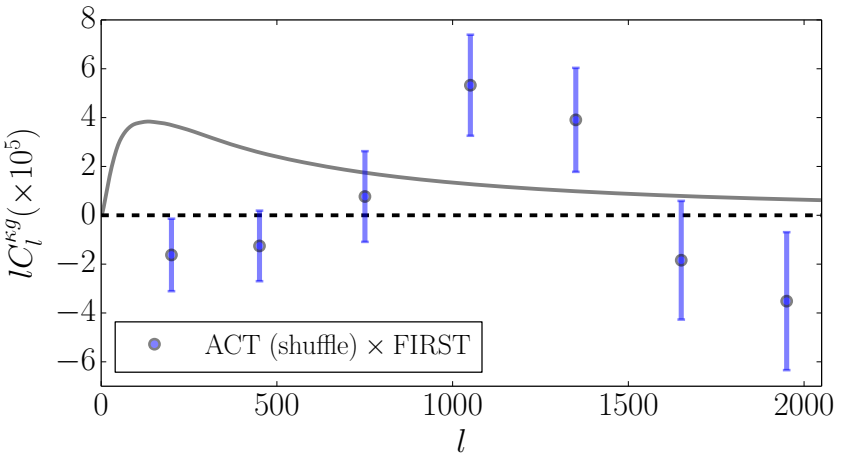

Figure 3.7: Cross-correlation between shuffled FIRST maps with ACT lensing convergence. Fitting the normalization of the fiducial bias model A to these data, we obtain $A_{\rm ACT,shuffle} = -0.18 \pm 0.31$, consistent with null (Sec. 3.3). *Grey solid curve:* Cross-spectrum for the fiducial bias model which best fits the data of Fig. 3.5.

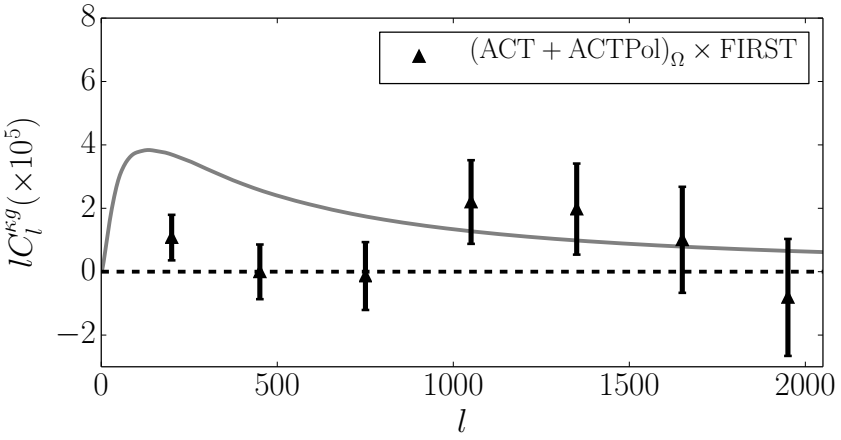

Figure 3.8: Cross-spectrum $lC_l^{\kappa g}$ between ACT+ACTPol lensing curl maps Ω and FIRST (Sec. 3.3). We restrict to $l < 2000$ for comparison with Fig. 3.5. Fitting the normalization of the fiducial bias model A to these data, we obtain $A_\Omega = 0.19 \pm 0.17$ and a chi-squared value of the null hypothesis of 21.8 for $\nu = 19$ degrees of freedom (a probability-to-exceed of 0.29). As expected this cross-correlation is consistent with null. *Grey solid curve:* Cross-spectrum for the fiducial bias model which best fits the data of Fig. 3.5.

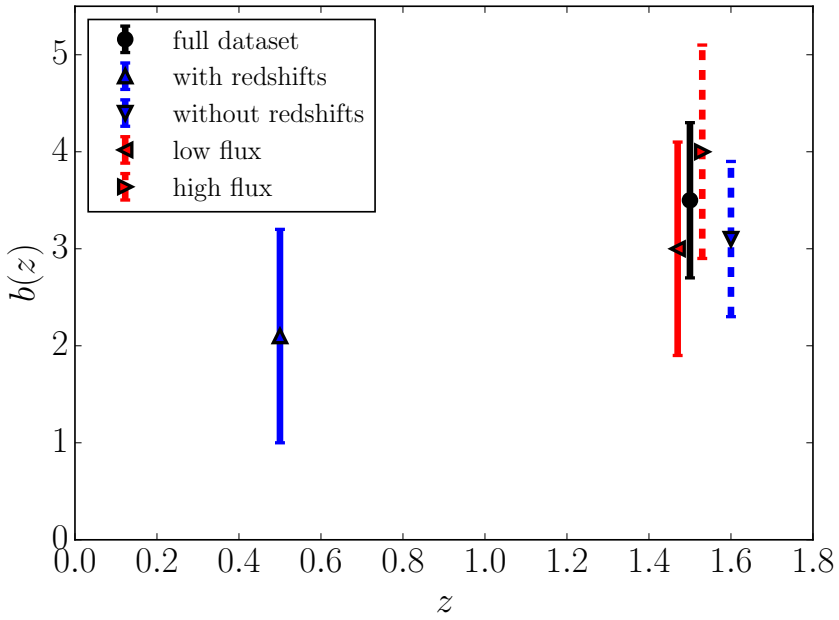

Figure 3.9: Summary of the primary bias results from Sec. 3.3 and Sec. 3.3.1. For various data splits we plot the corresponding constraint on the bias at the effective redshift $z_{\rm eff}$ of the measurement. We show the full dataset constraint (*black, solid, circle*), the constraint for faint (*red, solid, left triangle*) and bright (*red, dashed, right triangle*) sources, and the constraint for sources with (*blue, solid, upper triangle*) and without (*blue, dashed, lower triangle*) individual identified redshifts. We displace the red points slightly (from $z_{\rm eff} = 1.5$) for visual clarity.

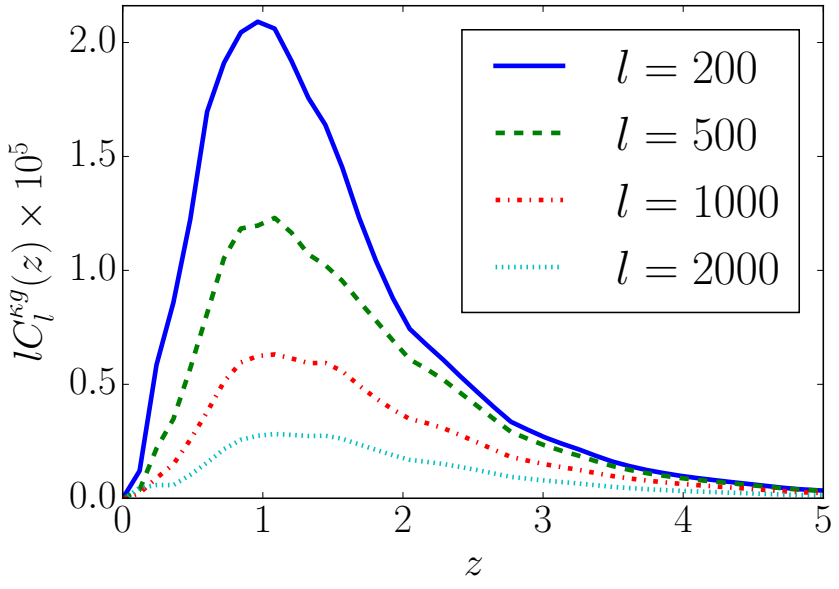

Figure 3.10: Cross-power spectrum kernel $C_l^{\kappa g}(z) \equiv W_\kappa(z) W_g(z) P(l/\chi(z); z)$, demonstrating the scale-dependent sensitivity of the cross-spectrum to source redshift. At $l = 200$, where the signal-to-noise peaks, the mean redshift of the kernel is $\int z C_{200}^{\kappa g}(z) dz / C_{200}^{\kappa g} = 1.5$, which we adopt as the effective redshift $z_{\rm eff}$ of the radio source bias measurement. The spread in the kernel reflects sensitivity to a wide range of redshift. See Sec. 3.2.3 for details.

We adopt $z_{\rm eff} = 1.5$ as the effective redshift of the measurement, estimating $b(z_{\rm eff}) = 3.5 \pm 0.8$. We note, however, that the measurement is sensitive to a range of redshifts: at $l = 200$ the kernel is non-negligible ($> 10\%$ of its peak value) out to redshifts $z > 3$, and the kernel shifts to higher redshifts at smaller scales. High-redshift sources make an important contribution to the small-scale cross-spectrum.

We consider a set of variations to the bias model. First we fit the data with a redshift-independent bias model, varying the amplitude b. We find $b = 2.8 \pm 0.6$, with $\chi^2_{\rm bf} = 11.2$, consistent with the primary result. The data alone cannot distinguish a redshift-independent bias model from an evolving bias model, although our redshift-dependent model is more physically motivated theoretically and by empirical observations (e.g., Wilman et al., 2008; Lindsay et al., 2014a).

We probe the flux dependence of the AGN bias by splitting the FIRST sources into two roughly equal-sized subsamples, with a partition at 2.5 mJy. We create new maps of these FIRST sources, as described in Sec. 3.2.2. From SKADS we estimate the normalised redshift distribution for each subsample, finding they are equal to within $\approx 15\%$ across $0.3 < z < 4$, calculate $z_{\rm eff} = 1.5$ for both subsamples, and expect a bias amplitude $\approx 20\%$ higher in the bright sample. We repeat the analysis of Sec. 3.2.3, finding $b(z_{\rm eff}; F > 2.5 \text{ mJy}) = 4.0 \pm 1.1$ and $b(z_{\rm eff}; F < 2.5 \text{ mJy}) = 3.0 \pm 1.1$. This is consistent with the expectation that the high-flux sample selects preferentially for the most-luminous sources, and these sources lie in the most highly-biased environments (e.g., Mo & White, 1996; Raccanelli et al., 2012).

We investigate whether the data provide information about the (largely unknown) high-redshift bias of radio-loud AGN. Here we fix the bias to the fiducial model at redshifts $z < 1.5$ and to a redshift-independent value above this. We constrain the high redshift bias to be $b(z > 1.5) = 4.1 \pm 1.7$. The detection significance is reduced relative to the full sample because only high-redshift sources ($\approx 1/3$ of the total) constrain this parameter. The increase in bias between low and high redshift samples is consistent with the result of Lindsay et al. (2014a), who show that the bias $b(z)$ continues to increase above redshift $z = 2$, although we note the significance is low.

We divide the source sample into those that have redshift estimates or not. A fraction of 0.27 of the FIRST sources have a reliable optical match as described in Sec. 3.2.2. The redshift distribution of these sources is strongly weighted to low redshifts, peaking around $z = 0.5$. Fol-

lowing Lindsay et al. (2014a) we can estimate the redshift distribution of the remaining sources by comparison with the SKADS simulated radio catalogue used to construct the model redshift distribution for the full sample. We construct independent overdensity maps for these two radio populations (with/without redshift) and recompute the data cross-spectra. We also recalculate the theoretical cross-correlation, accounting for the different source distributions, as a function of a redshift-independent bias term b. We find $b = 2.1 \pm 1.1$ at an effective redshift $z_{\rm eff} = 0.5$ for the sample with redshifts, and $b = 3.1 \pm 0.8$ at an effective redshift $z_{\rm eff} = 1.6$ for the sample without redshifts, shown in Fig. 3.9. Although not formally significantly different, this is again consistent with an increasing bias as a function of redshift.

3.3.2 Comparison to previous bias measurements

Geach et al. (2013) find a constant linear bias $b = 1.61 \pm 0.22$ at an effective redshift $z \approx 1$ for IR-selected quasars from WISE in cross-correlation with the SPT convergence map. At the same redshift our bias amplitude determination corresponds to $b(z = 1) = 2.6 \pm 0.6$. Their quasar sample is shallower (42 sources deg^{-2}) than in the FIRST maps presented here (71 sources deg^{-2}), and the predominant signal comes from $z < 2$ sources (there are expected to be no $z > 3$ sources). The higher bias determination presented here is consistent with a more highly biased population of sources being sampled.

Sherwin et al. (2012) constrain a constant linear bias $b = 2.5 \pm 0.6$ for optically-selected quasars from SDSS totalling 75 sources deg^{-2}. The redshift distribution of these sources peaks at $z = 1.4$. White et al. (2012) determine $b = 3.8 \pm 0.3$ from the two-point correlation function of quasars in the Baryon Oscillation Spectroscopic Survey across the redshift range $2.2 < z < 2.8$. Comparing with Fig. 3.9, this is in good agreement with our result and assumed bias model.

Lindsay et al. (2014b) measure the bias as a function of redshift by auto-correlation of radio sources from the GAMA survey to the same depth (1 mJy) as this analysis. Assuming comoving clustering, their low redshift measurement, $b(z \approx 0.5) = 2.13^{+0.90}_{-0.76}$, is consistent with the results presented here, while at high redshift they find $b(z \approx 1.5) = 9.45^{+0.58}_{-0.67}$, significantly higher than seen in this analysis. However, we note there is no cosmic variance contribution included in their quoted uncertainties, and they show this is a large effect for the small GAMA fields for sources at lower redshift. We suggest the high value of their measurement relative to that presented here

is consistent with cosmic variance.

Our result probes the multipole range $100 < l < 3000$, corresponding to physical scales $\approx$ 2–60 Mpc at the effective redshift $z_{\rm eff} = 1.5$. As seen in Fig. 3.5, low- and high-redshift sources contribute to the cross-spectrum differently as a function of scale. At the scales probed by the *Planck* Collaboration (2013b) lensing cross-correlation analysis with NVSS radio sources ($l < 400$), $z > 1.5$ sources contribute $\sim 1/3$ of the total cross-spectrum, whereas at smaller scales these sources contribute equally alongside the $z < 1.5$ sources. By measuring the cross-spectrum across a wide-range of scales one might distinguish between low- and high-redshift sources. Future high-precision determinations of this cross-spectrum will go further in breaking the degeneracy between source populations and constraining the bias as a function of redshift.

We can translate the constraint on the AGN bias at redshift $z_{\rm eff} = 1.5$ into an inference on the mass of the halo in which the typical AGN source resides. Using the fitting function of Tinker et al. (2010), we find $\log(M_{\rm halo}/M_\odot) = 13.6^{+0.3}_{-0.4}$, assuming that haloes virialize at a density ratio $\Delta = 200$ times that of the universe at the epoch of formation. If we instead adopt the Mo & White (1996) prescription relating bias to halo mass, as used in this analysis and Wilman et al. (2008) to construct the bias models for each radio population, we find a consistent figure of $\log(M_{\rm halo}/M_\odot) = 13.5^{+0.3}_{-0.4}$. This observed mass is higher than seen in, e.g., Sherwin et al. (2012) – $\log(M_{\rm halo}/M_\odot) = 12.9^{+0.3}_{-0.5}$ – and Geach et al. (2013) – $\log(M_{\rm halo}/(h^{-1}M_\odot)) = 12.3^{+0.3}_{-0.2}$. This is consistent with the observations of e.g., Shen et al. (2009) and Hatch et al. (2014) that the environments of radio-loud AGN are significantly denser than those of radio-quiet AGN.

We find a high bias for these sources compared to optically- and IR-selected AGN. This analysis provides complementary information by probing the bias of radio-selected AGN which, in the context of previous work, is indicative of bias evolution and a very large halo mass for these sources. The broad picture is that of an increasing bias as a function of redshift, and of radio-loud AGN occupying more massive haloes than radio-quiet AGN across a similar redshift range.

Our findings are in line with studies of the stellar masses (e.g., Jarvis et al., 2001b; Seymour et al., 2007) and environments (e.g., Wylezalek et al., 2013; Hatch et al., 2014) of powerful radio sources to high redshift. Specifically, we find strong evidence that powerful radio sources are more highly biased tracers of the dark matter density field than other AGN that are detectable to high redshift (e.g., quasars; Sherwin et al., 2012; Geach et al., 2013). As well as being important for

tracing the underlying dark matter distribution with *multi-tracer* techniques such as described in Ferramacho et al. (2014), this also suggests that mechanical feedback from the jets of powerful radio AGN, should only have a significant effect on the level of star formation within the most massive dark matter haloes at all epochs. However, we note that such an effect can not only have an impact on both the AGN host galaxy (e.g., Croton et al., 2006; Bower et al., 2006; Hopkins et al., 2006; Dubois et al., 2013; Mocz et al., 2013), but also the wider cluster environment (e.g., Rawlings & Jarvis, 2004).

3.3.3 Modelling limitations and astrophysical systematics

Because $C_l^{\kappa g}$ depends only on the product $b(z)dn/dz$, and we constrain only its normalisation, we stress that the results presented here are contingent on the assumed redshift evolution of the bias and redshift distribution of the sources.

The SKADS simulation is populated using empirical radio luminosity functions as described in Wilman et al. (2008). Extrapolation of the empirical luminosity functions into unobserved regimes will lead to inaccuracies in the inferred redshift distribution and bias model. To investigate the sensitivity of our measurement to uncertainties about the source redshift distribution, we recompute the theoretical spectra, unrealistically removing all sources above redshift $z > 3$ when calculating dn/dz; at high redshift the underlying dn/dz is most uncertain and likely depends on radio luminosity (e.g., Jarvis & Rawlings, 2000; Jarvis et al., 2001b; Wall et al., 2005; Rigby et al., 2011). Fitting the theoretical cross-spectrum as in Sec. 3.3, we find $b(z_{\rm eff} = 1.2) = 3.2 \pm 0.8$, representing a $\sim 0.25\sigma$ shift from the primary result under this significant perturbation of the theoretical redshift distribution. We thus do not expect that the source distribution uncertainty strongly biases our result, although future analyses with higher statistical power will require careful consideration of this systematic uncertainty.

We fix the cosmology to the *Planck* best-fit values throughout this analysis, which could affect the inference of the AGN bias. However, the significant (40σ) detection of the *Planck* lensing auto-spectrum means that model uncertainty from the cosmology is sub-dominant with respect to astrophysical uncertainties (*Planck* Collaboration, 2015e). Perturbing the *Planck* best-fit cosmological parameters by $+1\sigma$ and recomputing the theoretical cross-spectrum, $C_l^{\kappa g}$, the amplitude is shifted by $< 6\%$ across all relevant scales; we thus neglect this source of systematic

uncertainty.

Potential astrophysical systematic contaminants include infrared sources, Sunyaev-Zel'dovich clusters and Galactic cirrus. Sherwin et al. (2012) show that these constitute small effects on the measured cross-spectrum between quasars and lensing ($< 10\%$ in total), negligible at the level of statistical uncertainty in this analysis. Although the sources studied in Sherwin et al. (2012) are optically-selected AGN, we expect the result to hold for the radio-loud AGN of this analysis given the roughly similar redshift distributions. Furthermore, bright radio sources ($\gtrsim 5$ mJy) in the CMB temperature and polarisation maps are subtracted prior to lensing reconstruction, using a match-filtered source template map, thus mitigating radio-source contamination in the CMB convergence map (Das et al., 2011; van Engelen et al., 2014).

3.4 Conclusions

We present a measurement of the angular cross-spectrum between lensing convergence from ACT and the overdensity of radio sources identified in the FIRST survey, rejecting the null-hypothesis of no correlation at 4.4σ significance. The data are well fit by the *Planck* best-fit ΛCDM cosmological model where we model the source population with a redshift-dependent bias. We interpret the result in terms of a constraint on the bias of AGN, which dominate the FIRST sample, considering various bias models and data splits to probe different redshift regimes and AGN populations, and put these in the context of previous measurements of AGN bias. We translate the bias determination into a constraint on the mass of the host haloes, corroborating previous work showing that the environments of radio-loud AGN are more dense than those of optically-selected AGN.

We consider various sources of systematic uncertainty, both astrophysical contaminants and modelling limitations. We conclude that our results are robust to these effects. As deeper and wider radio surveys and improved lensing maps become available, these systematic effects will become increasingly important to measure and model accurately. The auto- and cross-spectra $\{C_l^{gg}, C_l^{\kappa\kappa}, C_l^{\kappa g}\}$ provide complementary information about the large-scale structure they probe, with the cross-spectrum in particular being robust to systematic biases particular to each dataset. A full analysis will simultaneously estimate the three power spectra, marginalising over uncertainty

in the redshift distribution and cosmology (Pearson & Zahn, 2014). With current data there are strong degeneracies in the cross-spectrum amplitude between sources from different redshifts. The shape of the power-spectra contains information about the bias evolution, and larger, more sensitive surveys will allow us to break these degeneracies.

The measurement of the high bias (and correspondingly large halo mass) of this radio population, relative to other dark matter tracers, indicates that these sources would be useful in the multi-tracer technique of Ferramacho et al. (2014). Using all the information in auto- and cross-correlations between multiple tracers, which differentially trace the dark matter, will provide tight constraints on primordial non-Gaussianity by reducing the impact of cosmic variance at large scales.

The SKA will serve as a deep probe of large-scale structure in the universe, it will be limited by different systematics than optical surveys, and the observed source distribution will be skewed to higher redshifts than either LSST or Euclid (Laureijs et al., 2011; Amendola et al., 2013; Jarvis et al., 2015). Kirk et al. (2015a) show that next-generation CMB lensing experiments, in combination with the SKA, will constrain the amplitude of the lensing-radio density cross-spectrum to the sub-percent level. With tight constraints on cosmology, this translates into $< 1\%$ uncertainty on the bias amplitude, offering broad scope for probing the history and evolution of AGN. Future high-precision measurements of $C_l^{\kappa g}$ will use information about the shape of the cross-spectrum, and source tomography, to constrain the bias as a function of redshift, calibrating galaxy redshift surveys and constraining extensions to ΛCDM.

The cross-correlation of CMB lensing with tracers of large-scale structure will become an increasingly important calibrator for future high-precision galaxy and weak-lensing surveys.

Chapter 4

Towards a cosmological neutrino mass detection

In this chapter we consider how CMB lensing, in particular, can be used to measure the mass of neutrinos. Future cosmological measurements should enable the sum of the neutrino masses to be determined indirectly through their effect on the expansion rate of the Universe and the clustering of matter. We consider prospects for the gravitationally-lensed cosmic microwave background (CMB) anisotropies and baryon acoustic oscillations in the galaxy distribution (BAO), examining how the projected uncertainty of ≈ 15 meV on the neutrino mass sum (a 4σ detection of the minimal mass) might be reached over the next decade. The current 1σ uncertainty of ≈ 103 meV (*Planck*-2015+BAO-15) will be improved by upcoming 'Stage-3' CMB experiments (S3+BAO-15: 44 meV), then upcoming BAO measurements (S3+DESI: 22 meV), and planned next-generation 'Stage 4' CMB experiments (S4+DESI: 15–19 meV, depending on angular range). An improved optical depth measurement is important: the projected neutrino mass uncertainty increases to 26 meV if S4 is limited to $\ell > 20$ and combined with current large-scale polarisation data. Looking beyond ΛCDM, including curvature uncertainty increases the forecast mass uncertainty by $\approx 50\%$ for S4+DESI, and more than doubles the uncertainty with a two-parameter dark energy equation of state. Complementary low-redshift probes including galaxy lensing will play a role in distinguishing between massive neutrinos and a departure from a $w = -1$, flat geometry.

This chapter is organised as follows. We introduce and summarise the measurement of the sum of the neutrino masses to date in Sec. 4.1. In Sec. 4.2 we give a brief review of the cosmological

effects of neutrinos, and in Sec. 4.3 we study how the mass measurement may be reached, step wise, using data collected during the coming decade. In Sec. 4.4 we investigate the dependence on experimental details, and in Sec. 4.5 we explore degeneracies with other cosmological parameters. We conclude in Sec. 4.6.

4.1 Introduction

A central goal in both cosmology and particle physics is to measure the mass of the neutrino particles. The neutrino sector is still poorly understood and the mechanism that gives rise to their mass is unknown. There are thought to be three active neutrino species, with mass differences measured through solar, atmospheric, reactor, and accelerator neutrino oscillation experiments (for reviews see, e.g., Gonzalez-Garcia & Nir, 2003; Maltoni et al., 2004; Smirnov, 2006; Feldman et al., 2012). The results imply a minimum total mass of 60 meV in a normal hierarchy with two lighter neutrinos and one heavier neutrino, or 100 meV in an inverted hierarchy with two massive neutrinos.

Cosmology provides an indirect probe of massive neutrinos (e.g., Doroshkevich et al., 1980, 1981; Hu & Dodelson, 2002; Dolgov, 2002; Elgarøy & Lahav, 2005; Tegmark, 2005; Hannestad, 2005; Ichikawa et al., 2005; Lesgourgues & Pastor, 2006; Fukugita, 2006; Giunti & Kim, 2007; Abdalla & Rawlings, 2007; Komatsu et al., 2009; Ferroni et al., 2009; Hannestad, 2010). At early times massive neutrinos behave like non-interacting relativistic particles and then later like cold dark matter. As such, they affect the expansion rate of the Universe differently than a pure radiation or pure matter component, and they modify the evolution of perturbations at early times. They also modify the growth of structure through a suppression of the clustering of matter on scales that entered the cosmic horizon while the neutrinos were relativistic.

The current 95% upper limit from cosmological data on the sum of the neutrino masses is $\Sigma m_\nu < 230$ meV, from the *Planck* measurements of the cosmic microwave background (CMB) combined with baryon acoustic oscillation (BAO) measurements from the Baryon Oscillation Spectroscopic Survey (BOSS; Anderson et al., 2014; *Planck* Collaboration, 2015c). The limit is $\Sigma m_\nu < 680$ meV from the CMB alone (*Planck* Collaboration, 2015c). Tighter limits have been found including Lyman-α forest measurements from quasars in the BOSS survey, but the result

depends on numerical hydrodynamical simulations which may contribute additional systematic uncertainty ($\Sigma m_\nu < 120$ meV; Palanque-Delabrouille et al., 2013, 2015).

Forecasts of mass limits for upcoming cosmological data sets, including galaxy lensing and clustering, redshift-space distortions, the kinematic Sunyaev-Zel'dovich effect, and counts of galaxy clusters, have been studied extensively, showing the promise of a wide range of future cosmological data to target a neutrino mass measurement (e.g., Kitching et al., 2008; Font-Ribera et al., 2014; Villaescusa-Navarro et al., 2015; Mueller et al., 2015; Errard et al., 2015). In this paper we focus on the combination of lensed CMB and BAO measurements – datasets which do not require detailed modelling of non-linear structure formation or an understanding of galaxy bias. The gravitationally-lensed CMB measures the growth of structure at times typically before the Universe was half its current age, and on angular scales larger than ≈ 100 Mpc, and so it is dominated by linear physics. Studies of this combination have been reported in Hall & Challinor (2012); Abazajian et al. (2013); Wu et al. (2014) and Pan & Knox (2015), with a 4σ detection of neutrino mass forecast for the next generation of experiments. In this paper we investigate this further, exploring the dependence on experimental details and on parameter degeneracies.

4.2 Cosmological effects of neutrinos

Standard Model neutrinos are initially relativistic, following a thermal distribution after decoupling from the primordial plasma when the Universe had a temperature of around $k_B T \approx 1$ MeV. The neutrino temperature decreases as the scale factor grows, until their rest-mass energy dominates and they become indistinguishable from cold dark matter (Fig. 1.9). For a neutrino of mass m_ν the transition occurs at $z \approx 120(\frac{m_\nu}{60\ \mathrm{meV}})$, so current limits indicate a transition epoch of $120 \lesssim z \lesssim 460$ for a normal mass hierarchy (Ichikawa et al., 2005).

This limit implies that the neutrinos were still relativistic when the CMB decoupled, so they are indistinguishable from massless neutrinos in the primary anisotropies. However, higher-mass neutrinos become non-relativistic sooner, which reduces the early-time integrated Sachs-Wolfe (ISW) effect (Sec. 1.3.1). This gravitational redshift of the CMB photons arises when the non-negligible radiation component causes the potentials of the density fluctuations to evolve and affects the anisotropies on scales around the first acoustic peak (Fig. 1.11; Hu & Dodelson, 2002;

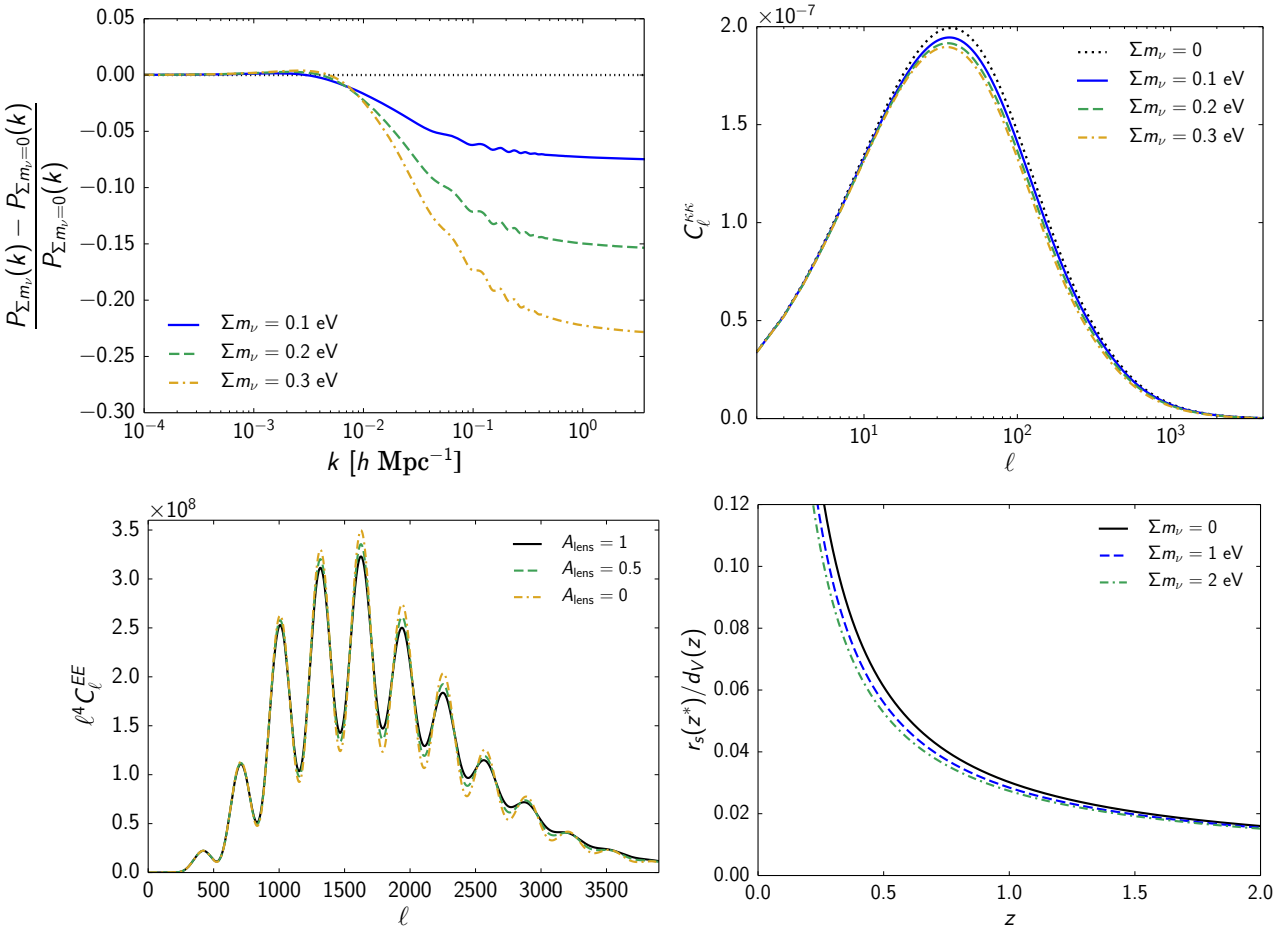

Figure 4.1: Effect of neutrino mass on CMB power spectra and the BAO distance ratio. Fractional change of the linear matter power spectrum today $P(k)$ (*top left*) and CMB convergence power spectrum $C_l^{\kappa\kappa}$ (*top right*) with neutrino mass Σm_ν, for fixed physical dark-matter density today, $\Omega_c h^2 + \Omega_\nu h^2$. Suppression of power is due to neutrino free-streaming, and the oscillatory structure comes from a slight shift in the acoustic scale between models. *Bottom left:* Lensed CMB E-mode power spectrum with varying amplitudes of the lensing potential $A_{\rm lens}$, approximating and exaggerating the effect that massive neutrinos have on the CMB polarisation spectrum. *Bottom right:* BAO distance ratio r_s/d_V for fixed θ_A and $\Omega_c h^2$. Massive neutrinos behave like additional matter in the BAO redshift range, decreasing H_0 and increasing the volume distance d_V.

Ichikawa et al., 2005; Hou et al., 2014).

The early ISW contribution from neutrinos is not sensitive to masses that remain relativistic until well after decoupling (Ichikawa et al., 2005; Komatsu et al., 2011), but further information comes from probes of later-time large-scale structure measurements. Massive neutrinos interact weakly, allowing them to free stream out of overdensities while relativistic, so the growth rate of matter perturbations inside the horizon is suppressed compared to a universe with only cold dark matter. For comoving wave numbers $k \gg k_{\rm FS}$, Hu et al. (1998) show that the suppression of the matter power spectrum today, $P(k)$, is proportional to the sum of the neutrino masses:

$$\frac{P_{\Sigma m_\nu}(k) - P_{\Sigma m_\nu=0}(k)}{P_{\Sigma m_\nu=0}(k)} \approx -0.07 \left(\frac{\Sigma m_\nu}{0.1 \text{ eV}}\right) \left(\frac{\Omega_{\rm m} h^2}{0.136}\right)^{-1}, \tag{4.1}$$

where the comoving free-streaming scale is given by

$$k_{\rm FS} = 0.0072 \left(\frac{\Sigma m_\nu}{0.1 \text{ eV}}\right)^{1/2} \left(\frac{\Omega_{\rm m}}{0.315}\right)^{1/2} h \,\text{Mpc}^{-1}, \tag{4.2}$$

as illustrated in Figure 4.1 and e.g., Abazajian et al. (2013), for models with fixed total matter density. For current limits this scale is estimated to lie in the range $0.005\ h\ \text{Mpc}^{-1} \lesssim k_{\rm FS} \lesssim 0.011\ h\ \text{Mpc}^{-1}$.

The suppression of small-scale power can be probed using galaxy clustering and the gravitational lensing of galaxies. These are promising avenues for neutrino mass measurements (e.g., Font-Ribera et al., 2014), although these observables are sensitive to non-linearities in the matter power spectrum and scale-dependent galaxy and shape biases (Bird et al., 2012; Mandelbaum, 2015). An alternative route is through the gravitational lensing of the CMB (see, e.g., Lewis & Challinor, 2006, for a review). Here the CMB photons are deflected by the large-scale structure, integrated over the photon path after decoupling.

Following Lewis & Challinor (2006), the CMB convergence angular power spectrum, $C_\ell^{\kappa\kappa}$, is a weighted projection of the matter power spectrum $P(k, \chi)$; under the Limber approximation,

$$C_\ell^{\kappa\kappa} = \int_0^{\chi_H} d\chi \frac{W^2(\chi)}{f_k^2(\chi)} P\left(\frac{\ell}{f_k(\chi)}, \chi\right), \tag{4.3}$$

where χ_H is the comoving horizon size, $f_k(\chi)$ relates line-of-sight comoving distances and trans-

verse comoving distances in a curved universe, and the window function $W(\chi)$ is

$$W(\chi) = \frac{3\Omega_{\rm m} H_0^2}{2c^2} \frac{f_k(\chi) f_k(\chi^* - \chi)}{a(\chi) f_k(\chi^*)} \tag{4.4}$$

for $\chi < \chi^*$ and zero otherwise. Here $a(\chi)$ is the scale factor and χ^* is the radial comoving distance to the last-scattering surface. This angular power spectrum is sensitive to Σm_ν, as shown in Figure 4.1, and does not depend on galaxy bias. CMB lensing probes structures at higher redshift than galaxy clustering or weak lensing, and is correspondingly less sensitive to detailed non-linear modelling of the matter power spectrum. In practice it is reconstructed from CMB temperature and polarisation maps using a four-point function (Sec. 1.4 and, e.g., Das et al., 2011).

The CMB temperature and polarisation angular power spectra, $\{C_\ell^{TT}, C_\ell^{TE}, C_\ell^{EE}, C_\ell^{BB}\}$, are also modified by lensing, which smears the acoustic peaks by adding variance to the apparent scale of a mode, converts E-mode polarisation into B-mode polarisation, and adds small-scale power in T, E, and B (e.g., Lewis & Challinor, 2006). The approximate effect of massive neutrinos is shown in Figure 4.1 for the E-mode polarisation, where we artificially amplify the effects of neutrinos on the CMB lensing, rather than the primary CMB, by varying the amplitude of the lensing potential. Increasing the neutrino mass has a similar effect to decreasing the lensing amplitude. Compared to the power spectra, the reconstructed convergence field contains more information on the neutrino mass (Benoit-Lévy et al., 2012).

Massive neutrinos also affect angular diameter distances $d_A(z)$ and the expansion rate $H(z)$, as their evolution differs from a pure radiation or pure matter component (e.g., Pan & Knox, 2015). These can be measured using a *standard ruler* method that is relatively free of systematic uncertainties: the primordial oscillations in the photon-baryon fluid are imprinted in the galaxy distribution as BAOs. The comoving scale of the oscillations is fixed by the sound horizon at decoupling, r_s, which is not significantly affected by neutrino masses given current limits and is in the linear regime of density perturbations (≈ 150 Mpc). The observed spherically-averaged BAO angular scale for galaxies at redshift z is sensitive to the parameter combination $r_s/d_V(z)$; d_V is the *volume distance* (Eisenstein et al., 2005),

$$d_V(z) \equiv \left[cz(1+z)^2 d_A^2(z) H^{-1}(z)\right]^{1/3}. \tag{4.5}$$

For fixed cold-dark-matter density, more massive neutrinos increase the total late-time non-relativistic matter content, which increases the volume distance, as shown in Fig 4.1.

4.3 Improvements in the next decade

We consider how upcoming and planned CMB and BAO experiments will improve the current limits on the sum of the neutrino masses, building on previous analyses (Hall & Challinor, 2012; Abazajian et al., 2013; Font-Ribera et al., 2014; Wu et al., 2014; Pan & Knox, 2015).

4.3.1 New data

The current state of the art for the CMB is the *Planck* and *WMAP* satellite data, including the first analysis of the full-mission *Planck* data (*Planck* Collaboration, 2015b,c). Improved small-scale CMB measurements are currently being made by the "Stage 2" ground-based experiments: ACTPol, SPTpol, and POLARBEAR (Naess et al., 2014; The POLARBEAR Collaboration, 2014a; Keisler et al., 2015). These will soon be upgraded to "Stage 3" (hereafter, S3) with new detectors and sensitivity to multiple frequencies. Here we consider a S3 "wide" experiment that maps 40% of the sky, and a "deep" experiment that maps 6% of the sky. The wide experiment is similar to AdvACT specifications (Calabrese et al., 2014) and the deep experiment to SPT-3G (Benson et al., 2014). These experiments are expected to take data from 2016 to 2019, and the specifications we adopt are given in Table 4.1. We also anticipate data from S3 experiments targeting larger angular scales (e.g., CLASS), but do not consider these specifically. In addition, we can expect a complete analysis of the *Planck* polarisation data, optimistically including reliable large-scale polarisation data.

Experiment	$f_{\rm sky}$	Beam (arcmin)	ΔT (μK-arcmin)	ΔP (μK-arcmin)
S3-wide	0.4	1.4	8.0	11.3
S3-deep	0.06	1	4.0	5.7
S4	0.4	3	1	1.4
CV-low	0.4	60	1	1.4

Table 4.1: Upcoming (S3) and proposed (S4, CV-low) CMB experiments that will add to *Planck*. The *Planck* specifications we use are in the Appendix.

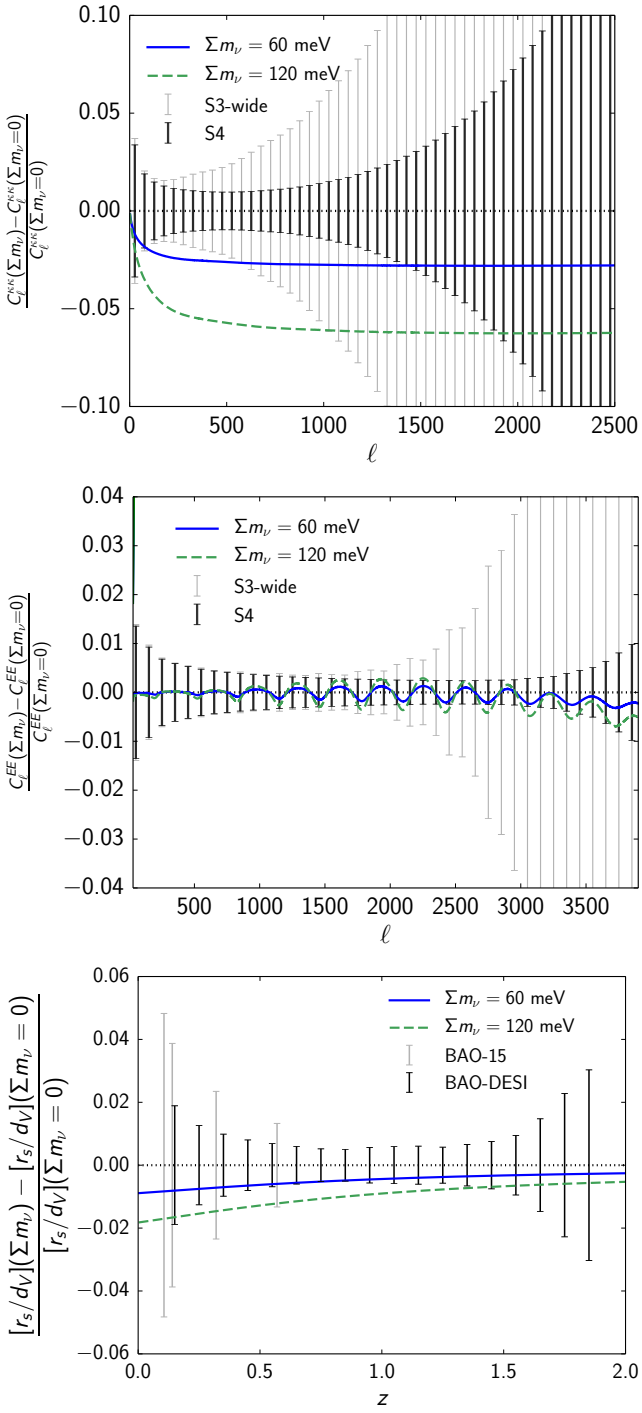

Figure 4.2: Fractional change in the convergence κ (*top panel*) and E-mode power spectrum (*middle panel*) with neutrino mass, for fixed $\Omega_c h^2 + \Omega_\nu h^2$, with expected uncertainties for S3-wide and S4 CMB data. A higher neutrino mass gives rise to less lensing, decreasing the E-mode peak smoothing. *Bottom panel*: Fractional change in distance ratio r_s/d_V, with uncertainties from current (BAO-15; Anderson et al., 2014) and forecast (DESI; Font-Ribera et al., 2014) BAO data. In this bottom panel $\Omega_c h^2$ is fixed.

Beyond S3, a "Stage-4" (S4) experiment – or set of experiments – is being developed by the CMB community that may cover at least half the sky to typical noise levels of 1 μK-arcmin (Abazajian et al., 2013). There are also proposed space-based experiments including LiteBIRD and PIXIE (Matsumura et al., 2014; Kogut et al., 2011), and we approximate their role with a cosmic variance-limited large-scale polarisation measurement ("CV-low") that could be combined with S3 ground-based data.

On the BAO front, current state-of-the-art measurements come from the BOSS "LowZ" and "C-MASS" galaxy samples at $z = 0.32$ and $z = 0.57$ (Anderson et al., 2014). These are supplemented by data from the Six-Degree Field Galaxy Redshift Survey (6dFGRS) at $z = 0.11$ (Beutler et al., 2011) and the SDSS MGS sample at $z = 0.15$ (Ross et al., 2015). Improved measurements are being made by the eBOSS survey which will survey a deeper sample (Comparat et al., 2015). A significant advance should be made with the Dark Energy Spectroscopic Instrument (DESI), due to begin surveying in 2018, which is expected to measure the BAO distance ratio from redshifts $0.15 < z < 1.85$ in bins of width $\Delta z = 0.1$ to percent-level precision (Levi et al., 2013; Font-Ribera et al., 2014).

4.3.2 Forecasting methods

We use a Fisher-matrix forecasting method to predict the neutrino mass uncertainties (Sec. 1.8). For a model defined by parameters $\boldsymbol{\theta}$, the Fisher matrix is

$$F_{ij}(\boldsymbol{\theta}) = \left\langle -\frac{\partial^2 \ln p(\boldsymbol{\theta}|\mathbf{d})}{\partial\theta_i \partial\theta_j} \right\rangle, \tag{4.6}$$

where $p(\boldsymbol{\theta}|\mathbf{d})$ is the posterior distribution for $\boldsymbol{\theta}$ given data $\mathbf{d}$. The forecast parameter covariance is then given by the inverse of the Fisher matrix, $\mathbb{C}_{ij} = (F^{-1})_{ij}$. Here our data are the lensed TT, TE, and EE CMB power spectra, reconstructed CMB convergence power spectrum $\kappa\kappa$, and BAO distance ratio measurements $r_s/d_V(z)$. Our parameters are the standard six ΛCDM parameters (the physical baryon and cold-dark-matter densities $\Omega_\mathrm{b}h^2$ and $\Omega_\mathrm{c}h^2$, the angular size of the sound horizon at recombination θ_A, the optical depth to reionisation τ, the amplitude of primordial scalar perturbations A_s and the tilt of the primordial scalar power n_s) plus the neutrino mass sum Σm_ν, as well as possible extension parameters including curvature and dark energy. Our

methods are summarised in the Appendix, including choices made about the fiducial model, the choice of parameter basis, and step sizes for calculating derivatives. We also describe validation of our numerical code.

We use the *lensed* CMB power spectra and the convergence power spectrum as our CMB observables, which differs from the approach in Hall & Challinor (2012); Abazajian et al. (2013); Font-Ribera et al. (2014); Wu et al. (2014), but more closely follows the real data analysis: the CMB sky we see is lensed, and it is a difficult inverse problem to infer the unlensed sky (e.g., Anderes et al., 2015). Using unlensed spectra in forecasts removes information contained in the lensed temperature and polarisation fields. However, it is challenging to construct the full covariance matrix for the lensed power spectra and convergence power spectrum: the T, Q and U fields are all lensed by the same lensing potential, which correlates the power spectra and adds additional non-Gaussian covariance. This is explored in detail in Benoit-Lévy et al. (2012) and Schmittfull et al. (2013). In this analysis we make the approximation of discarding BB information and assuming Gaussian uncertainties in TT, TE, EE and $\kappa\kappa$. We account for the correlation between the temperature and polarisation signals and the unlensed T–κ correlation (from the ISW effect), but neglect other correlations with κ. This is likely a good approximation for S3 data, but could underestimate certain parameter errors for S4-type data by up to $\approx 20\%$ (Benoit-Lévy et al., 2012). We use CAMB for evaluation of all relevant CMB and lensing power spectra (Lewis et al., 2000).

For the noise levels of *Planck*, we consider two cases, "*Planck*-2015" (P15), which produces cosmological constraints which closely match the published results (*Planck* Collaboration, 2015c), and "*Planck*-pol", which includes TE and EE data coming from the polarisation measurements of the High Frequency Instrument (HFI), including large scales. The specifications are given in the Appendix. For *Planck*-pol, noise levels are approximated by taking temperature sensitivities from P15 and assuming the per-channel noise scalings from temperature to polarisation in the *Planck* Blue Book (*Planck* Collaboration, 2005). This is likely to be over-optimistic at the largest scales.

For the CMB power spectra, we set a maximum multipole for the recoverable information: $\ell^T_{\rm max} = 3000$, $\ell^P_{\rm max} = 4000$, and $\ell^\kappa_{\rm max} = 3000$ for the future S3 and S4 experiments. Smaller scales are likely hard to extract due to extragalactic foreground contamination. We assume white noise, and do not include additional foreground uncertainty beyond the multipole cuts outlined above,

although the expected S3-wide white-noise level includes some foreground inflation (Calabrese et al., 2014). We also set a minimum multipole of $\ell_{\rm min} = 50$ for S3 due to the challenge of recovering large scales from the ground, and consider two options for S4: $\ell_{\rm min} = 50$ and $\ell_{\rm min} = 5$. For S3 and S4 we include *Planck* data for $2 < \ell < \ell_{\rm min}$, and our nominal analysis uses *Planck*-pol unless stated otherwise. We consider the importance of the large-scale polarisation measurements in Sec. 4.4.1.

We use the quadratic-estimator formalism of Hu & Okamoto (2002) to calculate the CMB convergence noise spectrum $N_\ell^{\kappa\kappa}$ (Sec. 1.4). This uses the coupling of otherwise uncorrelated modes in temperature and polarisation to reconstruct the lensing potential. Iterative delensing procedures are able to reduce the effective noise level of the estimated lensing field, particularly for the low-noise ($\Delta P \lesssim 4$ μK-arcmin) future experiments considered here (Smith et al., 2012). We consider the impact of this process in Sec. 4.3.3.

For the BAO measurements we use the published uncertainties on the distance ratio r_s/d_V for the current BAO data described in Sec. 4.3, assigning the label 'BAO-15' (Beutler et al., 2011; Anderson et al., 2014; Ross et al., 2015). For DESI we use the forecast uncertainties on $d_A(z)$ and $H(z)$ given in Font-Ribera et al. (2014) to estimate the expected r_s/d_V uncertainties, summarized in the Appendix. We do not use broadband shape information in the galaxy power spectra.

In Figure 4.2 we show the fractional changes in lensing and E-mode power spectrum and the BAO distance ratio for physically relevant variations in the neutrino mass. We overplot projected uncertainties for S3, S4, BAO-15 and DESI to illustrate the constraining power of these experiments.

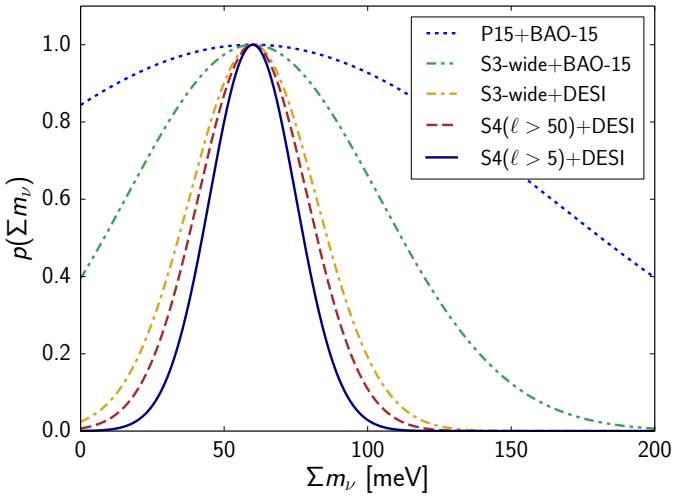

Figure 4.3: Forecast marginal posterior constraints on the sum of the neutrino masses Σm_ν within a ΛCDM+Σm_ν model, assuming Gaussian uncertainties. The current uncertainties (P15+BAO-15) are expected to improve rapidly, with S3 CMB data and DESI BAO data expected by ∼2020.

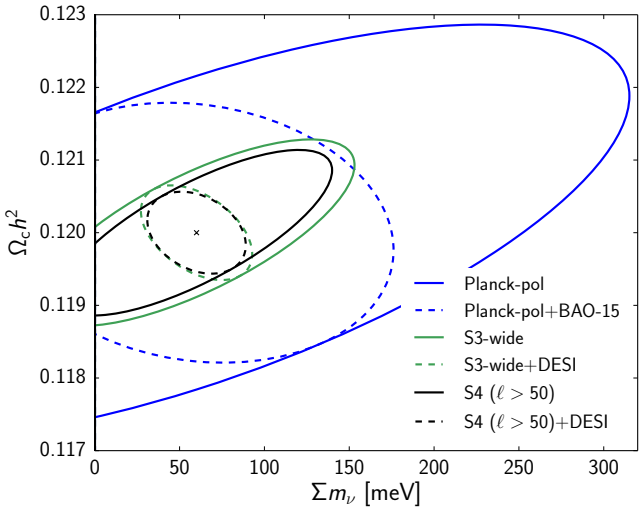

Figure 4.4: Expected joint constraint (68% C.L.) on the neutrino mass sum Σm_ν and physical cold-dark-matter density $\Omega_c h^2$ within a ΛCDM+Σm_ν model. The BAO constraint, sensitive to the total late-time cold-dark matter-density, is almost orthogonal to the CMB lensing constraint, breaking the degeneracy.

4.3.3 Expected constraints

We find forecast marginal 1σ uncertainties on the sum of the neutrino masses Σm_ν of

$$\frac{\sigma(\Sigma m_\nu)}{\text{meV}} = \begin{cases} 103 & \text{(P15 + BAO-15)} \\ 44 & \text{(S3-wide + BAO-15)} \\ 22 & \text{(S3-wide + DESI)} \\ 19 & \text{[S4 } (\ell > 50) \text{ + DESI]} \\ 15 & \text{[S4 } (\ell > 5) \text{ + DESI]} \end{cases} \tag{4.7}$$

where S3 and S4 include *Planck*-pol at large scales. If we replace S3-wide with the S3-deep survey we find $\sigma(\Sigma m_\nu) = 53$ meV when combined with BAO-15, and 25 meV with DESI. The expected constraints are summarized in Figure 4.3, and are consistent with findings in Wu et al. (2014). These forecasts imply that if the neutrino hierarchy is inverted, with mass sum > 100 meV, we may have $> 2\sigma$ evidence for nonzero mass in the next few years from S3 data, and an almost 5σ detection in $\approx$5 years with DESI. If instead it is the minimal mass of 60 meV, a 2–3σ indirect measurement should be reachable in five years, with stronger evidence from the subsequent experiments.

As illustrated in Figure 4.4, there is a strong positive correlation between the neutrino mass and cold-dark-matter density in the CMB observables. This arises predominantly from the competing influence of these parameters on the lensing signal. Increased neutrino mass suppresses small-scale power, while increasing cold-dark-matter density *boosts* small-scale power by shifting matter-radiation equality to earlier times; this shortens the radiation-domination epoch in which sub-horizon modes of the gravitational potential decay, enhancing the small-scale amplitude of structure (Namikawa et al., 2010; Hall & Challinor, 2012; Pan et al., 2014). Conversely, BAO data constrain the *sum* of the CDM and massive neutrino density, since it is this combination that affects the angular diameter distance and expansion rate at the redshifts probed by BAO experiments.

The power of the BAO and CMB data lie in their combination. BAO measurements alone cannot constrain the primordial parameters A_s and n_s, nor the optical depth to reionisation τ;

the constraints lie in the four-dimensional subspace spanned by $\Omega_c h^2, \Omega_b h^2, \Sigma m_\nu$, and θ_A (being most constraining in the directions corresponding approximately to $\Omega_m h^2$ and θ_A). S4 alone could achieve a neutrino mass error of $\sigma(\Sigma m_\nu) = 53$ meV; with DESI BAO this is expected to tighten to $\sigma(\Sigma m_\nu) = 19$ meV, enough for a 3σ detection of the minimal mass.

For comparison, we note that the forecast conditional neutrino mass uncertainty (the uncertainty when *fixing* the ΛCDM parameters) is $\sigma(\Sigma m_\nu) = 3$ meV for S4($\ell > 5$)+DESI, compared to $\sigma(\Sigma m_\nu) = 15$ meV in the marginalised case. This highlights the effect of degeneracies even in the base model.

Comparing to previous results, the S4($\ell > 5$)+DESI forecast matches previous results (Abazajian et al., 2013; Wu et al., 2014; Pan & Knox, 2015), and the "current data" P15 + BAO-15 forecast is compatible with the published result: $\sigma(\Sigma m_\nu) < 0.23$ eV at 95% confidence (*Planck* Collaboration, 2015c), with the forecast errors on other ΛCDM parameters also matching closely.

Here we choose by default a fiducial mass sum of 60 meV, approximating the normal hierarchy as one massive neutrino and two massless neutrinos, with a modified temperature to account for QED and non-instantaneous decoupling effects (such that $N_{\rm eff} = 3.046$). Given the known mass splittings from oscillation experiments, this is an excellent approximation to the normal hierarchy in the minimal-mass scenario (Feldman et al., 2012). For comparison, we compute the constraint on the neutrino mass sum in the inverted hierarchy, with two degenerate massive neutrinos and one massless neutrino, with a neutrino mass sum of $\Sigma m_\nu = 120$ meV. For these two cases we find $\sigma(\Sigma m_\nu) = 53$ (54) meV for S4 alone for a fiducial mass of 60 (120) meV, and with DESI included this reduces to $\sigma(\Sigma m_\nu) = 19$ (20) meV. We therefore do not expect our results are sensitive to the adopted fiducial model. We describe further tests of the code in the Appendix.

We consider the impact of using an improved, likelihood-based lensing estimator (going beyond the first-order quadratic estimator of Hu & Okamoto, 2002) to reduce the effective noise power $N_l^{\kappa\kappa}$ of the lensing reconstruction. Hirata & Seljak (2003) show that $N_l^{\kappa\kappa}$ can be reduced by a factor of 2 for an S4-like experiment; under this modification, we find only a 3% tightening of the neutrino mass constraint for S4+DESI. The improvement is small due to the contribution of cosmic variance (CV) to the lensing power spectrum uncertainties, and the degeneracy of neutrino mass with other ΛCDM parameters; the S4 lensing power spectrum derived from the quadratic estimator is already CV-limited out to $\ell \approx 800$.

As in Benoit-Lévy et al. (2012), we find that there is useful information in the lensed power spectra. We show the relative impact on forecasted constraints for S3-wide and S4 in Table 4.2 (including *Planck*-pol at large scales but no BAO). Lensing reduces the neutrino mass uncertainty by a factor of ≈ 6 compared to the unlensed CMB. As such, the two-point functions will provide increasingly more information as E-mode polarisation measurements improve. Checking for consistency between the lensed observables will be an important systematic test for new CMB data. For example, marginalising over the lensing effect in the two-point functions, by introducing a variable lensing amplitude parameter $A_{\rm lens}$, would isolate the impact of neutrinos on the four-point function (Calabrese et al., 2008).

	Unlensed	Lensed TT, TE, EE (two-point only)	Unlensed+$\kappa\kappa$ (four-point only)
S3, $\sigma(\Sigma m_\nu)$:	435	75	61
S4, $\sigma(\Sigma m_\nu)$:	363	64	53

Table 4.2: Impact of lensing information on the neutrino mass constraint (in units of meV). Constraining power comes from both the lensed spectra (two-point) and the reconstructed lensing potential (four-point). Gaussian uncorrelated errors are assumed.

4.4 Dependence on experimental details

Since future CMB experiments are currently under development, we investigate the importance of certain experimental details on the mass constraint.

4.4.1 Importance of the reionisation bump

The amplitude of primordial power, A_s, is partially degenerate with Σm_ν, since A_s increases the amplitude of clustering at small scales, and Σm_ν decreases it in an approximately scale-independent fashion (e.g., Fig. 4.2). The amplitude A_s is not well determined by the primordial CMB temperature anisotropy; an increased optical depth to reionisation lowers the signal such that the normalization of the anisotropy measures the combination $A_s e^{-2\tau}$ (*Planck* Collaboration, 2013a). This leads to a degeneracy between τ and Σm_ν that can be broken with precision measurements of the reionisation bump at multipoles $\ell < 20$ in polarisation (Sec. 1.3.3; Smith et al., 2006; Font-Ribera et al., 2014).

Here we explore the importance of making a robust optical depth measurement, considering three cases for S4: current *WMAP* measurements (Hinshaw et al., 2013; *Planck* Collaboration, 2015b), optimistic future *Planck*-pol measurements (see Appendix), and a future S4 measurement that reaches the largest scales ($\ell_{\rm min} = 5$). We find forecast constraints of

$$\frac{\sigma(\Sigma m_\nu)}{\rm meV} = \begin{cases} 27 & [\text{S4 } (\ell > 50) + \textit{WMAP}\text{-pol} + \text{DESI}] \\ 19 & [\text{S4 } (\ell > 50) + \textit{Planck}\text{-pol} + \text{DESI}] \\ 15 & [\text{S4 } (\ell > 5) + \text{DESI}] \end{cases} \tag{4.8}$$

with the uncertainty on τ reducing from 0.008 to 0.005 to 0.003, respectively. This is compared to 0.013 for *WMAP*-pol from EE alone, i.e., improved CMB lensing data helps constrain τ even when the neutrino mass is varied. Figure 4.5 shows the expected correlation between τ and neutrino mass.

Figure 4.5 also shows the impact of reducing the minimum multipole of the S4 experiment on the neutrino mass constraint, supplemented with *Planck*-pol or the current *WMAP*-pol at the largest scales. There is a limiting plateau for S4 at $\ell_{\rm min} > 20$, and a clear improvement as the polarisation is better measured at increasingly large scales. The S4($\ell > 5$) + DESI limit reaches the CV limit for CMB data[1].

We also consider the relative importance of making a higher sensitivity small-scale measurement versus a new large-scale polarisation measurement. We start with a S3-type $\ell > 50$ experiment, and then either increase the $\ell > 50$ sensitivity, or supplement it with a new CV-low large-scale measurement at $\ell < 50$. We find forecast constraints of

$$\frac{\sigma(\Sigma m_\nu)}{\rm meV} = \begin{cases} 22 & [\text{S3 } (\ell > 50) + \textit{Planck}\text{-pol} + \text{DESI}] \\ 19 & [\text{S4 } (\ell > 50) + \textit{Planck}\text{-pol} + \text{DESI}] \\ 17 & [\text{S3 } (\ell > 50) + \text{CV-low} + \text{DESI}] \end{cases} \tag{4.9}$$

This indicates that a cosmic-variance-limited measurement of optical depth could be more valuable than more sensitive small-scale data, especially given that *Planck*-pol large-scale polarisation data

[1] Pan & Knox (2015) found that an S4 experiment combined with a CV-limited BAO experiment could tighten the neutrino mass constraint further, to $\sigma(\Sigma m_\nu) = 11$ meV.

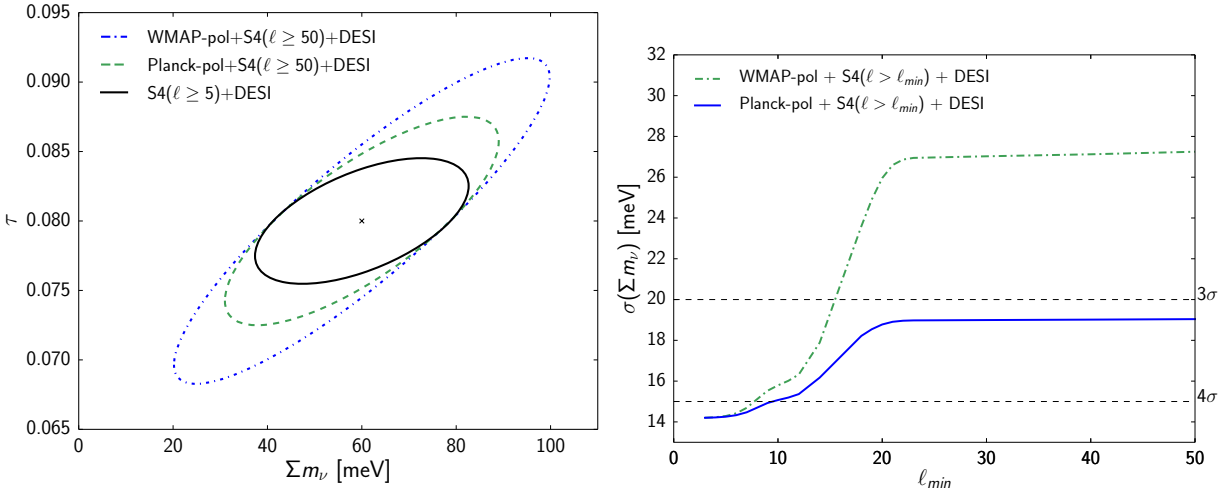

Figure 4.5: *Left panel:* The neutrino mass Σm_ν is correlated with the optical depth to reionisation τ (forecast 68% CL). Current data at $\ell < 20$ (*WMAP*-pol) would leave a degeneracy between Σm_ν and τ that could be broken with improved large-scale polarisation data. *Right panel:* the expected neutrino mass constraint as a function of the minimum multipole accessible to S4, indicating the benefit of reaching large scales.

is itself not yet demonstrated to be free of systematic errors.

We note that 21-cm experiments, which map the brightness temperature of neutral hydrogen as a function of redshift, will probe the epoch of reionisation (Koopmans et al., 2015); the combination of this information with CMB+BAO would break the Σm_ν–τ degeneracy and improve the neutrino mass constraint (investigated in Liu et al., 2015).

4.4.2 Importance of sensitivity and angular range

For the particular case of an $\ell > 50$ experiment covering 40% of the sky at 3-arcmin resolution, combined with *Planck*-pol, we vary the white-noise sensitivity. The forecast neutrino mass limits are shown in Figure 4.6 for CMB-only, CMB+BAO-15, and CMB+DESI. There is clearly an improvement as the noise is reduced, and a significant gain is expected over current *Planck* measurements, but below white-noise levels of ≈ 10 μK-arcmin there does not appear to be a substantial gain (as also seen in Wu et al., 2014).

It is not yet certain whether this ≈ 10 μK-arcmin noise level, over half the sky, will be achieved in practice from the upcoming S3 CMB experiments, or whether the lensing reconstruction will achieve the expected noise levels. Atmospheric, ground and foreground emission are typical contaminants that would increase the effective noise in the maps and in the lensing reconstruction.

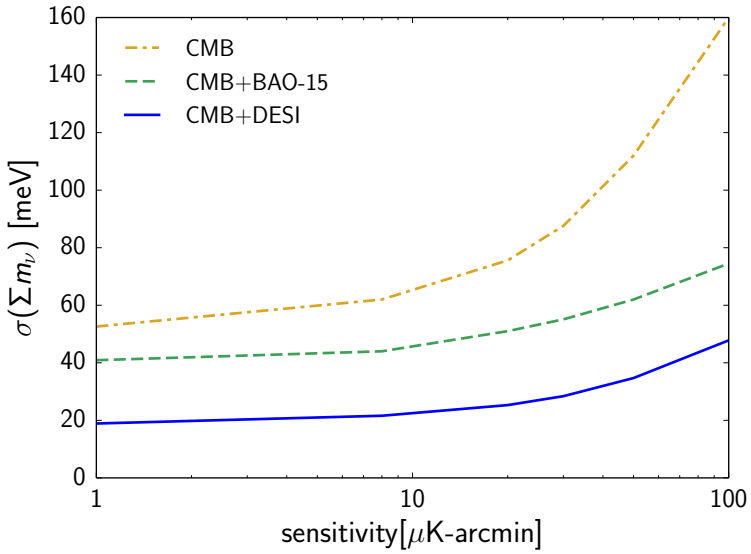

Figure 4.6: The dependence of the neutrino mass constraint $\sigma(\Sigma m_\nu)$ on CMB map sensitivity, for a 3-arcmin resolution experiment covering 40% of the sky.

New data from the current S2 experiments will help clarify the impact of non-white noise on the lensing reconstruction performance.

Here we have continued to restrict our analysis to 'clean' scales at $\ell < 3000$ in TT, TE and $\ell < 4000$ in EE. A foreground-free S3 or S4 experiment would contain information from $\ell > 3000$; the TT spectrum is signal dominated to small scales ($\ell \approx 4800$ for S4) where there are a large number of modes. However, uncertainty about extragalactic foregrounds will likely make this information inaccessible. We find that including $\ell > 3000$ scales would tighten the neutrino mass error by $\approx 10\%$.

We restrict the modes available for reconstruction of the lensing potential to these same scales, and to $\ell > 50$, since reconstruction on the largest scales has yet to be demonstrated. There are difficulties such as mean-field subtraction for masked fields (e.g., *Planck* Collaboration, 2015e). We find that the information, quantified by the term $F_\ell^{\Sigma m_\nu}$ in the Fisher matrix, is concentrated in the multipole range $100 \lesssim \ell \lesssim 1000$ for S3 and $200 \lesssim \ell \lesssim 2000$ for S4. This also means that detailed non-linear modelling should not be required, since at redshift $z \approx 2$ at the peak of the CMB lensing kernel (Eq. 4.4), the information peak corresponds to physical scales $\gtrsim 200$ Mpc.

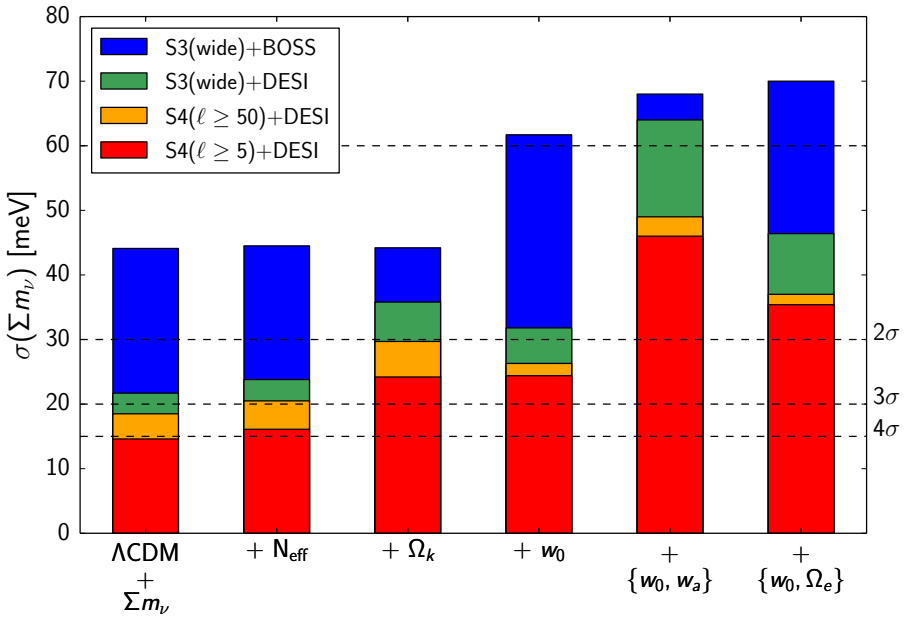

Figure 4.7: Neutrino mass constraints forecast for different data combinations and simple one- or two-parameter extensions to the ΛCDM+Σm_ν model (all except S4($\ell \geq 5$) include *Planck*-pol information at low ℓ). With CMB+BAO data, neutrino mass is not correlated with the number of neutrino species, $N_{\rm eff}$, but is partly correlated with spatial curvature, Ω_k, and with the dark energy equation of state, w_0. Expected confidence levels are shown assuming the minimal total neutrino mass Σm_ν= 60 meV.

4.5 How unique is the massive neutrino signal?

The measurable effect of neutrinos can be partly mimicked by changes in other cosmological parameters as explored in, e.g., Font-Ribera et al. (2014); Benoit-Lévy et al. (2012); Hamann et al. (2012). Massive neutrinos affect the expansion rate and angular diameter distance, but changes in curvature, dark-energy history or the Hubble parameter can compensate to keep the well-constrained acoustic peak structure and BAO distance ratio essentially unchanged.

Here we consider changes in the spatial curvature, and changes in the energy density of dark energy with time. Dark energy is invoked to explain the acceleration of the universe, and may take the form of a cosmological constant with constant energy per unit proper volume, although an evolving dark-energy equation of state ($w \neq -1$) is not ruled out by current observations (*Planck* Collaboration, 2015c). We consider two parameterizations for dark energy: these are the usual Taylor expansion in the scale factor for dynamical dark energy (Chevallier & Polarski, 2001;

Linder, 2003),

$$w(a) = w_0 + w_a(1-a), \tag{4.10}$$

with two free parameters w_0 and w_a, and also the Doran & Robbers (2006) model for early dark energy,

$$\Omega_\Lambda(a) = \frac{\Omega_\Lambda - \Omega_e\left(1-a^{-3w_0}\right)}{\Omega_\Lambda + \Omega_\mathrm{m} a^{3w_0}} + \Omega_e\left(1-a^{-3w_0}\right), \tag{4.11}$$

with parameters Ω_e and w_0. This has a background expansion similar to a massive neutrino for periods of the evolution of the universe. Previous work has considered neutrino mass constraints within this model (e.g., Calabrese et al., 2011; Joudaki & Kaplinghat, 2012). We use the CAMB Parameterized Post-Friedmann module (Fang et al., 2008) and a modified version of CAMB from Calabrese et al. (2011) to compute the power spectrum within these models.

We take as our baseline the S4 ($\ell > 50$)+DESI experiment. Marginalising over simple extensions to the ΛCDM+Σm_ν model, we find

$$\frac{\sigma(\Sigma m_\nu)}{\mathrm{meV}} = \begin{cases} 19 & (\Lambda\mathrm{CDM}{+}\Sigma m_\nu) \\ 30 & (\Lambda\mathrm{CDM}{+}\Sigma m_\nu{+}\Omega_k) \\ 27 & (\Lambda\mathrm{CDM}{+}\Sigma m_\nu{+}w_0) \\ 46 & (\Lambda\mathrm{CDM}{+}\Sigma m_\nu{+}w_0{+}w_a) \\ 37 & (\Lambda\mathrm{CDM}{+}\Sigma m_\nu{+}\Omega_e{+}w_0) \\ 64 & (\Lambda\mathrm{CDM}{+}\Sigma m_\nu{+}w_0{+}w_a{+}\Omega_k) \end{cases} \tag{4.12}$$

We discuss these parameter degeneracies in the remainder of this section. These findings are summarised in Figure 4.7, which also includes the corresponding S3 forecasts. We find that marginalising over the neutrino number N_eff (and other extension parameters that modify the primordial CMB spectrum such as a running spectral index) have a $< 10\%$ effect on the predicted neutrino mass uncertainties.

At first sight, the degradation of the neutrino mass estimate in the case of varying w_0, w_a, and Ω_k simultaneously appears severe, more than tripling the error bar. However, this model has three extra parameters compared to ΛCDM, and within the Bayesian framework, one can rigorously ask whether additional parameters are required by the data, quantifying the trade-off

between improving the fit against an increased complexity of the model (see, e.g., Trotta, 2007, for discussion in the context of cosmological data). This model selection approach would quantify the need for additional extension parameters, and would disfavor an overparameterized model if it is not required by the data. In practice our challenge is likely to lie in distinguishing between different one-parameter extensions to ΛCDM: are we seeing non-zero neutrino mass, or could the data be similarly well-explained by a small amount of curvature, or a deviation from a $w = -1$ equation of state?

4.5.1 Physical degeneracies

To help understand these degeneracies, we note the Hubble parameter $H(z)$ is given by

$$\frac{H^2(z)}{H_0^2} = \Omega_{\rm r}(1+z)^4 + \Omega_{\rm m}(1+z)^3 + \Omega_k(1+z)^2 + \Omega_\Lambda(z), \tag{4.13}$$

at times after the neutrinos become non-relativistic, where $\Omega_\Lambda(z) \equiv \rho_\Lambda(z)/\rho_{\rm crit}$ is the dark-energy density (a constant if $w = -1$), $\rho_{\rm crit}$ is the critical density today, $\Omega_{\rm r}$ is the radiation density today (e.g., photons and massless neutrinos), $\Omega_{\rm m} = \Omega_{\rm c} + \Omega_{\rm b} + \Omega_\nu$ is the matter density today (CDM, baryons and massive neutrinos) and $\Omega_\nu h^2 = \Sigma m_\nu/94.1$ eV is the physical massive neutrino density today (Eq. 1.46). The angular diameter distance is given by

$$d_A(z) = \frac{c}{H_0(1+z)} \begin{cases} \frac{1}{\sqrt{-\Omega_k}} \sin\left(\sqrt{-\Omega_k}\int_0^z \frac{H_0 dz'}{H(z')}\right) & \Omega_k < 0, \\ \int_0^z \frac{H_0 dz'}{H(z')} & \Omega_k = 0, \\ \frac{1}{\sqrt{\Omega_k}} \sinh\left(\sqrt{\Omega_k}\int_0^z \frac{H_0 dz'}{H(z')}\right) & \Omega_k > 0. \end{cases} \tag{4.14}$$

Photons propagating in a non-flat universe follow curved geodesics, changing the angular diameter distance to an object of fixed proper size, at a given comoving distance, relative to a flat universe. Varying curvature shifts the angular scale of the acoustic peaks; to remain consistent with the CMB data, the matter density and Hubble constant H_0 must vary to keep the peak structure unchanged. This is the well-known *geometric degeneracy* (Bond et al., 1997).

This degeneracy is partially broken by CMB lensing measurements (e.g., Sherwin et al., 2011), which are sensitive to the growth of structure in the late-time universe and, therefore, to the

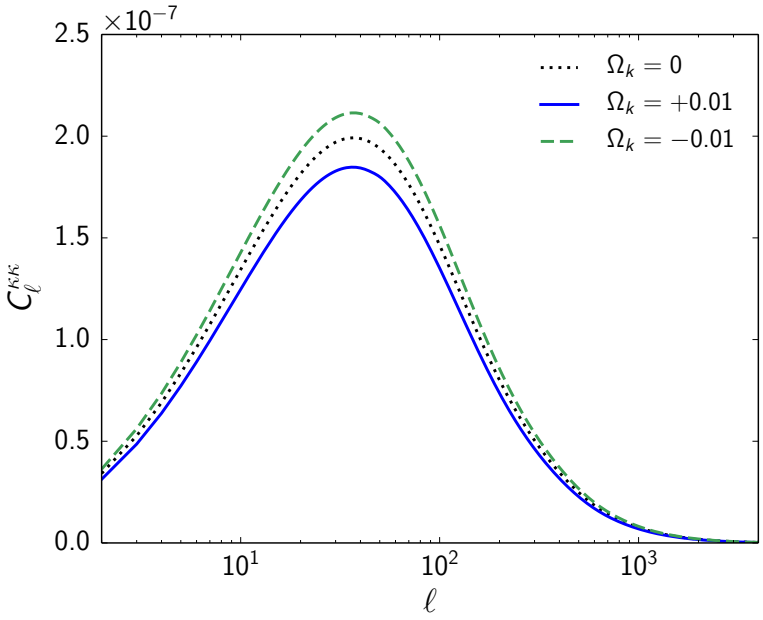

Figure 4.8: CMB convergence power spectrum for varying Ω_k, with other parameters varied to hold the primary CMB fixed. Decreasing Ω_k requires a smaller Hubble constant H_0 and increased growth rate. This has a similar effect to decreasing the neutrino mass.

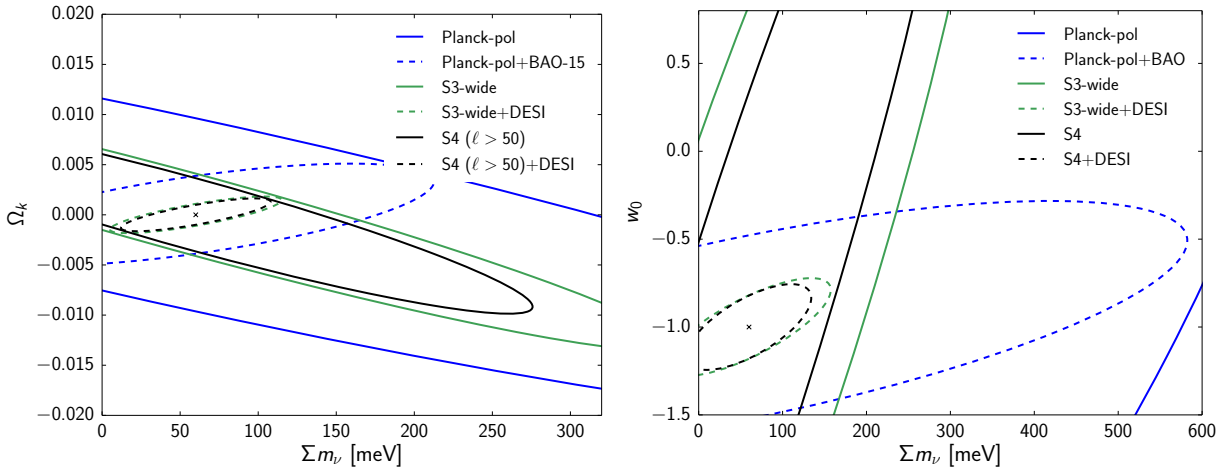

Figure 4.9: *Left panel:* Forecast joint constraint on the neutrino mass, Σm_ν, and spatial curvature, Ω_k, within a ΛCDM+Σm_ν+Ω_k model. The BAO data breaks the anti-correlated degeneracy in the CMB data. *Right panel:* Forecast constraint on Σm_ν and the dark energy equation of state, w_0, marginalised over the ΛCDM and w_a parameters.

matter density $\Omega_{\rm m}$ and dark energy density Ω_Λ. For a fixed CMB acoustic peak scale, the effect of decreasing the curvature parameter Ω_k, moving to a more closed universe, is to decrease the Hubble constant and increase $\Omega_{\rm m}$. This enhances the amplitude of the CMB lensing power spectrum, as illustrated in Figure 4.8, which can be compensated by increasing Σm_ν. This leads to an anti-correlation between Σm_ν and Ω_k when using CMB measurements alone, as shown in Figure 4.9.

The BAO constraint in the Σm_ν–Ω_k plane is approximately orthogonal to the CMB-only constraint, because decreasing Ω_k decreases the volume distance to a given redshift (Eq. 4.5). This can be compensated by a smaller matter density, which can be achieved by lowering the neutrino mass. These data are powerful in combination, and for S3+BOSS the neutrino mass constraint is independent of curvature. However, the neutrino mass constraint from S3+DESI or S4+DESI is expected to degrade by $\approx 50\%$ when allowing for curvature, as illustrated in Figure 4.7.

Similar arguments apply to dark energy, which modifies the background evolution according to its equation of state, but does not contribute to clustering. Increasing the dark-energy equation-of-state parameter w_0 leads to an increased expansion rate, shifting the angular scale of the acoustic peaks; to remain consistent with the CMB data, the Hubble constant H_0 decreases to keep the peak structure unchanged. Similar to the curved model, this increases the clustering, which can be compensated with larger neutrino masses. This gives the positive correlation between w_0 and Σm_ν in CMB data, illustrated in Figure 4.9 and reported in, e.g., Hannestad (2006); Benoit-Lévy et al. (2012). The BAO degeneracy is, however, also positively correlated, so the neutrino mass uncertainty is inflated more than when allowing for curvature. Increasing Σm_ν increases the contribution of neutrinos to $\Omega_{\rm m}$, requiring a smaller Ω_Λ (in a flat universe); the volume distance to a given redshift, and, hence, the BAO peak position, can then be preserved by increasing w_0. The early-dark-energy density parameter Ω_e, is anticorrelated with neutrino mass, due to their similar effects on the background expansion[2] (Calabrese et al., 2011; Joudaki & Kaplinghat, 2012).

We find that the neutrino mass constraints from S3+DESI or S4+DESI are degraded by more than a factor of 2 when allowing for a non-minimal ($w \neq -1$) dark-energy equation-of-state in the cosmological model, as shown in Figure 4.7. Distinguishing a non-zero neutrino mass parameter from a universe with dark energy beyond the cosmological constant will be difficult

[2]Here we adopted a fiducial $\Omega_e = 0.007$, allowed by current data (*Planck* Collaboration, 2015d). A smaller Ω_e would improve the neutrino mass constraint, as the parameters are anticorrelated and Ω_e cannot be negative.

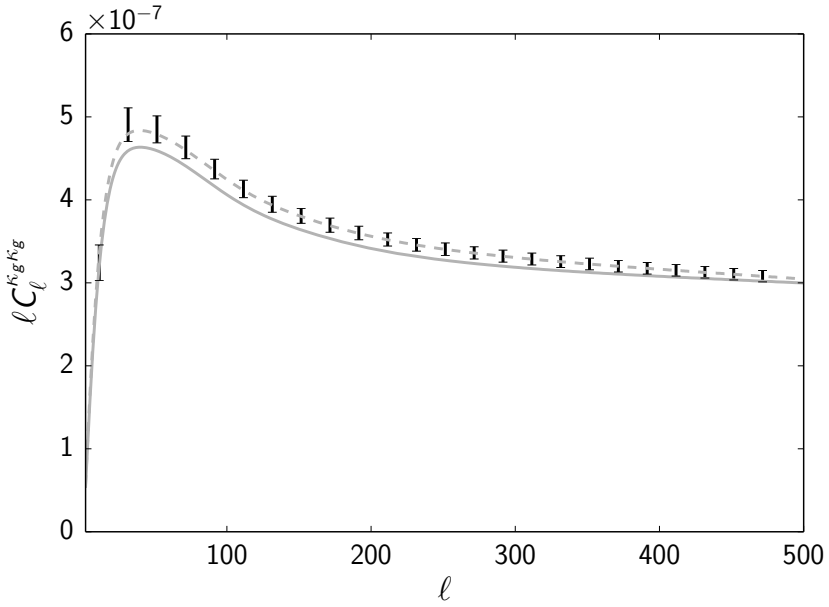

Figure 4.10: Galaxy lensing power spectra for two models degenerate in CMB spectra and BAO distance ratios. They could be distinguished using a future galaxy weak lensing survey ($f_{\rm sky} = 0.5$, $n_{\rm gal} = 10$ arcmin^{-2}, and source redshift distribution $n(z)$ as for the LSST 'gold sample'; LSST Science Collaboration et al., 2009).

with the CMB+BAO data combinations considered here. Allowing for freedom in both dark energy and curvature (i.e. a ΛCDM+Σm_ν+Ω_k+w_0+w_a model) degrades the constraint further to $\sigma(\Sigma m_\nu) = 64$ meV for S4+DESI, but would include three new parameters.

4.5.2 Breaking degeneracies with complementary measurements

We have focused on the minimal combination of future CMB and BAO data. To break these remaining degeneracies between neutrino mass and curvature/dark-energy parameters, we would turn to other large-scale structure probes such as the galaxy power spectrum, redshift-space distortions, galaxy weak lensing, the kinematic Sunyaev-Zel'dovich effect, and galaxy cluster counts, which measure the growth of structure at later times, and Type Ia supernovae, which better constrain the expansion rate.

We have not included these and other data in our baseline forecasts as they are, arguably, more prone to systematic uncertainties such as tracer bias and the shape of the nonlinear power spectrum, which contains unknown baryonic feedback effects, source distribution uncertainties and multiplicative bias (Das et al., 2013; Mandelbaum, 2015; *Planck* Collaboration, 2015c). How-

ever, a promising path would be to examine a suite of different complementary probes, each in combination with the CMB and BAO data, to distinguish neutrino mass from a non-flat or non-Λ model.

For example, in Figure 4.10 we illustrate schematically how galaxy weak lensing can break remaining degeneracies between two example models. Here the CMB spectra and BAO distances are indistinguishable with S4+DESI data (see parameters in the Appendix).

Their galaxy weak lensing signals, which probe the growth of structure during the dark-energy-dominated era are distinguishable with a realistic future weak lensing survey. With a 2% difference in σ_8, robust cluster count measurements should also help discriminate between these models (e.g., Carbone et al., 2012; Costanzi Alunno Cerbolini et al., 2013).

4.6 Discussion and conclusions

We have demonstrated that an indirect detection of the sum of the neutrino masses should be possible with upcoming CMB and BAO surveys. In the next decade a 4σ detection should be attainable, within the ΛCDM+Σm_ν model, even in the minimal mass ($\Sigma m_\nu = 60$ meV) scenario. We have found that this is contingent on obtaining improved large-scale polarisation measurements from the CMB, which may be the hardest experimental challenge. It will also be necessary to exclude degeneracies with other plausible extensions such as curvature and dark energy. We find that allowing for these extensions degrades the expected neutrino mass constraint, and that use of other large-scale structure probes will be necessary to rule out other departures from ΛCDM.

Our forecasts make a number of assumptions. We have neglected non-Gaussian terms in the CMB covariance due to correlation of the temperature, polarisation and lensing fields, which should have a small impact on imminent data but will become more important as noise levels reduce. We assume Gaussianity of the posterior distribution and rely on data independence of the covariance matrix. These assumptions become increasingly accurate for the precision future experiments considered here, but remain an approximation. Our analysis also assumes little foreground contamination. This will be most important for the large-scale CMB signal, but is not expected to significantly degrade the smaller-scale lensing signal.

We have assumed white noise in the new CMB data. Typically, this has not yet been achieved

in practice from the ground, due to atmospheric and scan-synchronous emission that induces additional variance at large scales. However, polarisation measurements are cleaner than temperature in this respect, and new experiments have sophisticated designs to modulate the polarisation signal. Performance of current Stage-2 experiments will help refine noise projections for the future experiments. We have also assumed ideal performance of the future DESI BAO experiment.

In terms of theoretical scope, beyond an evolving dark-energy equation-of-state, we have not considered other possible new physics. For instance, theoretically-motivated axion contributions to the background expansion and perturbations might offer an alternative explanation for a neutrino mass detection. The cosmological effects are not identical (Marsh et al., 2012), but further investigation will be useful.

Finally, neutrino mass is the next beyond-ΛCDM parameter that we know will be needed to fit data, so it is valuable that competitive constraints are expected to come from different combinations of cosmological data sets beyond the CMB and BAO data considered here. Their complementarity will aid in convincingly excluding systematic effects and alternative cosmological models.

Appendix: Forecasting parameter constraints

We forecast posterior parameter constraints and degeneracy directions by evaluating the relevant Fisher matrix as defined in Eq. 4.6. We assume uniform priors on all parameters, allowing the posterior $p(\boldsymbol{\theta})$ to be replaced with the likelihood $\mathcal{L}(\boldsymbol{\theta}) = \mathbb{P}(\mathbf{d}|\boldsymbol{\theta}, M)$ in Eq. 4.6. The Fisher matrix is evaluated at *fiducial parameters*; the choice is unimportant under the assumption of a Gaussian posterior and data-independent covariance, although in practice they are chosen to match the current best-fit parameters of the model, as the real covariance does have a term that scales with the signal. Our choices are shown in Table 4.3.

The observables entering the Fisher matrix are the CMB power spectra and the BAO distance ratio measurements. For the CMB, we insert the likelihood in Eq. 1.56 into Eq. 4.6; neglecting parameter dependence in the power spectrum covariance matrix, one obtains

$$F_{ij} = \sum_{\ell} \frac{\partial C_l^{\top}}{\partial \theta_i} \mathbb{C}_{\ell}^{-1} \frac{\partial C_l}{\partial \theta_j}. \tag{4.15}$$

Parameter	Fiducial value	Step size
$\Omega_{\rm b}h^2$	0.0222	0.0008
$\Omega_{\rm c}h^2$	0.1197	0.0030
$10^2\theta_A$	1.0409	0.0050
$10^9 A_s$	2.196	0.1
n_s	0.9655	0.010
τ	0.078	0.020
Σm_ν (meV)	60	20
Ω_k	0	0.01
w_0	-1	0.3
w_a	0	0.6
$N_{\rm eff}$	3.046	0.080
Ω_e	0.007	0.002

Table 4.3: Fiducial values and step sizes for the numerical derivatives, chosen to be small enough to minimize error in the Taylor expansion of the two-sided derivative, while keeping numerical stability in the derivatives from the CAMB power spectra. The fiducial value of $\Omega_e = 0.007$ is at the upper 95% confidence level given current data (*Planck* Collaboration, 2015d).

Experiment	$f_{\rm sky}$	ν/GHz	$l_{\rm min}$	$l_{\rm max}$	FWHM/arcmin	$\Delta T/\mu$K-arcmin	$\Delta P/\mu$K-arcmin
Planck-2015	0.44	30,44,70,100	2	2500	33,23,14,10	145,149,137,65	...,...,450,...
		143,217,353			7,5,5	43,66,200	...,...,...
Planck-pol							...,...,450,103
							81,134,406
WMAP-pol	0.74	33,41,64,94	2	1000	41,28,21,13	...,...,298,296	425,420,424,...

Table 4.4: Specification for the *Planck* and *WMAP* experiments used in the analysis, assuming white-noise properties. We define *Planck*-2015 to reproduce the constraints from *Planck* 2015 data; for *Planck*-pol we use the *Planck* Blue Book scaling factors to convert to polarisation. For *WMAP*-pol we recover an optical depth uncertainty that matches the WMAP9 data. When combining with S3 and S4, we include *Planck*-pol data across the full $f_{\rm sky} = 0.44$ at large scales ($\ell < \ell_{\rm min}^{\rm S3/S4}$) and across $f_{\rm sky} = 0.2$ in the multipole range $\ell_{\rm min}^{\rm S3/S4} < \ell < 2500$, as the useful *Planck* data and S3/S4 will likely not overlap completely. Our results are insensitive to the exact choice of this non-overlapping region size. When using *WMAP*-pol data we substitute it in at large scales ($\ell < \ell_{\rm min}^{\rm S3/S4}$) over $f_{\rm sky} = 0.74$.

Experiment	Redshift	$\frac{\sigma(r_s/d_V)}{(r_s/d_V)}$ (%)	$\sigma(r_s/d_V)$	Ref
6dFGRS	0.106	4.83	0.0084	(Beutler et al., 2011)
SDSS MGS	0.15	3.87	0.015	(Ross et al., 2015)
LowZ	0.32	2.35	0.0023	(Anderson et al., 2014)
C-MASS	0.57	1.33	0.000 71	(Anderson et al., 2014)
DESI	0.15	1.89	0.00410	(Font-Ribera et al., 2014)
	0.25	1.26	0.00170	
	0.35	0.98	0.000 88	
	0.45	0.80	0.000 55	
	0.55	0.68	0.000 38	
	0.65	0.60	0.000 28	
	0.75	0.52	0.000 21	
	0.85	0.51	0.000 18	
	0.95	0.56	0.000 18	
	1.05	0.59	0.000 17	
	1.15	0.60	0.000 16	
	1.25	0.57	0.000 14	
	1.35	0.66	0.000 15	
	1.45	0.75	0.000 16	
	1.55	0.95	0.000 19	
	1.65	1.48	0.000 28	
	1.75	2.28	0.000 41	
	1.85	3.03	0.000 52	

Table 4.5: Specification for current BAO-15 data (*top*), and forecast DESI data (*bottom*). We derive the expected fractional uncertainties on r_s/d_V for DESI from the fractional errors on d_A/r_s and $H(z)$ forecast in Font-Ribera et al. (2014). Using the r_s/d_V observable discards information in redshift-space distortions of the galaxy correlation function by treating the clustering as isotropic. The absolute values correspond to a ΛCDM model with $\Sigma m_\nu = 60$ meV.

Applying Wick's theorem, the covariance matrix for the power spectra has elements

$$\mathbb{C}(\hat{C}_l^{\alpha\beta}, \hat{C}_l^{\gamma\delta}) = \frac{1}{(2l+1)f_{\rm sky}}\left[(C_l^{\alpha\gamma} + N_l^{\alpha\gamma})(C_l^{\beta\delta} + N_l^{\beta\delta}) + (C_l^{\alpha\delta} + N_l^{\alpha\delta})(C_l^{\beta\gamma} + N_l^{\beta\gamma})\right], \quad (4.16)$$

where $\alpha, \beta, \gamma, \delta \in \{T, E, B, \kappa_c\}$, and $f_{\rm sky}$ accounts for the loss of information due to partial sky coverage (Knox, 1995; Hobson & Magueijo, 1996; de Putter et al., 2009). Noise spectra are generated for each observable given input noise properties such as CMB map sensitivities. We assume additive white noise for the CMB,

$$N_\ell^{\alpha\alpha} = (\Delta T)^2 \exp\left(\frac{\ell(\ell+1)\theta^2_{\rm FWHM}}{8\ln 2}\right), \quad (4.17)$$

for $\alpha \in \{T, E, B\}$, where ΔT (ΔP for polarisation) is the map sensitivity (in μK-arcmin) and $\theta_{\rm FWHM}$ is the beam width. This is an optimistic approximation: real noise-spectra from ground-based experiments have a dominant contribution from atmospheric variance at large scales (see, e.g., Fig. 4 of Das et al., 2013). The atmosphere is weakly polarized, and hence the white-noise approximation is better in E and B than T. The CMB lensing reconstruction noise is calculated using the Hu & Okamoto (2002) quadratic-estimator formalism. As described in the main text, we neglect non-Gaussian terms in the power spectrum covariance, and we also neglect the BB spectrum as it does not contribute significantly to upcoming constraints and has a highly non-Gaussian covariance (Benoit-Lévy et al., 2012).

We add information from BAO experiments by computing the BAO Fisher matrix:

$$F_{ij}^{\rm BAO} = \sum_k \frac{1}{\sigma^2_{f,k}} \frac{\partial f_k}{\partial \theta_i} \frac{\partial f_k}{\partial \theta_j}, \quad (4.18)$$

where $f_k \equiv f(z_k) = r_s/d_V(z_k)$ is the sound horizon at photon-baryon decoupling r_s over the volume distance d_V to the source galaxies at redshift z_k. These real and forecast data are reported in Table 4.5.

The total Fisher information matrix is given by the sum of the CMB and BAO Fisher matrices, and is inverted to forecast parameter covariances. An alternative Monte Carlo Markov Chain approach, using simulated data, can be taken to account for non-Gaussianity of the posterior (e.g., Hall & Challinor, 2012), but the Gaussian approximation is likely increasingly good as the

data quality improve from *Planck* through S3 to S4.

Our forecasting code, OxFish, has been developed for this analysis and is used to forecast parameter covariance matrices in one coherent python package. The code interfaces with the CAMB code for evaluation of power spectra. We compare with real data or previous work where possible. We construct the *Planck*-2015 (P15) specification, given in Table 4.4, to produce constraints that match the ΛCDM uncertainties from the *Planck* Collaboration (2015c), with the beam sizes and noise levels matching the detector sensitivities in *Planck* Collaboration (2015a).

We also forecast the neutrino mass constraint from P15+BAO, finding $\sigma(\Sigma m_\nu) = 103$ meV. Placing the peak of the posterior at the fiducial $\Sigma m_\nu = 60$ meV, this corresponds to $\Sigma m_\nu < 245$ meV at 95% confidence, comparable to the actual result of $\Sigma m_\nu < 230$ meV (*Planck* Collaboration, 2015c). This includes joint light-curve supernovae data and an H_0 prior, but these are expected to have a small impact. We also agree with Abazajian et al. (2013); Wu et al. (2014); Pan & Knox (2015) on the neutrino mass constraint for the S4+DESI data combination, finding $\sigma(\Sigma m_\nu) = 15$ meV if we assume that the reionisation bump is measured.

We use the following parameters for the curves in Figure 4.10: solid curve; $\{\Omega_\mathrm{b}h^2 = 0.0222$, $\Omega_\mathrm{c}h^2 = 0.0120$, $\Sigma m_\nu = 3$ meV, $\tau = 0.067$, $10^2\theta_A = 1.0417$, $10^9 A_s = 2.15$, $n_s = 0.9647$, $\sigma_8 = 0.835$, $H_0 = 68.0\}$; dashed curve, $\{\Omega_\mathrm{b}h^2 = 0.0222$, $\Omega_\mathrm{c}h^2 = 0.0119$, $\Sigma m_\nu = 117$ meV, $\tau = 0.089$, $10^2\theta_A = 1.0399$, $10^9 A_s = 2.24$, $n_s = 0.9663$, $\sigma_8 = 0.822$, $H_0 = 66.6\}$.

Chapter 5

Weak-lensing cross-correlations

We considered the correlation of CMB lensing and the clustering of radio galaxies in chapter 3. Here we present a forecasting and analysis pipeline for the cross-correlation between CMB lensing and cosmic shear, focusing on applications to the Atacama Cosmology Telescope (ACT). We introduce and motivate the joint analysis of CMB lensing and galaxy lensing in Sec. 5.1, including a discussion of measurements to date and key systematic effects in lensing data. We argue that simultaneous analysis provides a robust approach to calibration of these effects. In Sec. 5.2 we review the theory of weak-lensing cross-correlations, and generate simulated maps to illustrate the properties of the observables. In Sec. 5.3 we detail how the shear signal is extracted from the shapes of galaxies and show how this is combined with CMB lensing data to estimate their cross-spectrum. The pipeline we have developed – from map making to parameter estimation – is presented and demonstrated on both real and simulated data in Sec. 5.4 . In Sec. 5.5 we forecast the ability of future experimental combinations to detect the lensing cross-correlation, and we conclude in Sec. 5.6.

5.1 Introduction

Gravitational lensing traces the integrated matter fluctuations along the line of sight to the photon source (Sec. 1.4). Measurements allow us to constrain the growth of structure, and the contents and expansion rate of the universe (e.g. Heymans et al., 2013; van Engelen et al., 2014; Story et al., 2015; *Planck* Collaboration, 2015e). Lensing is unbiased, in that it traces the *total* matter

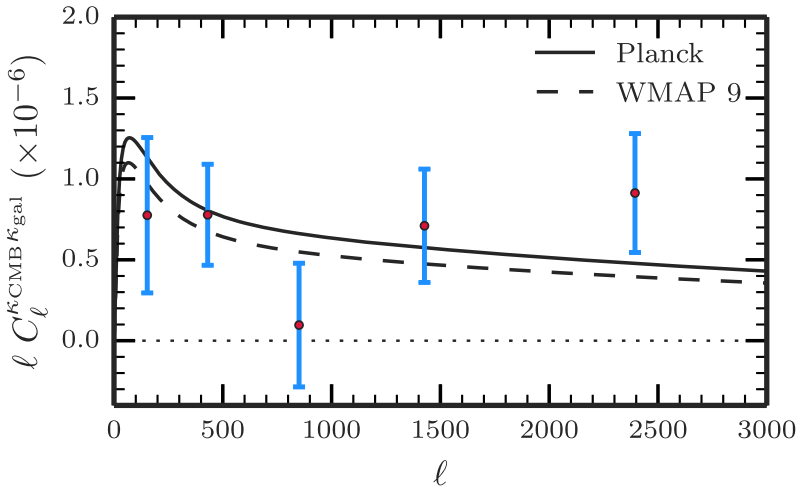

Figure 5.1: *Source: Hand et al. (2015).* Cross-spectrum between ACT lensing and CS82 galaxy lensing (*red points*). The cross-spectrum is detected at 4.2σ significance over a 120 deg^2 area, and is well-fit by a ΛCDM model with a best-fit *Planck* or *WMAP* 9 cosmology. This represents the first measurement of the galaxy lensing – CMB lensing cross-correlation.

distribution of the universe. It is therefore not affected by unknown tracer bias – which enters into clustering measurements of luminous matter – and is hence a more direct probe of the growth of structure (Sec. 1.5).

CMB lensing and galaxy lensing are complementary cosmological probes. CMB lensing is most sensitive to structure at higher redshifts (its kernel peaks around $z \approx 2$), while the dominant galaxy lensing signal traces structure at later times, typically around $z \approx 0.5$ (Hand et al., 2015; Liu & Hill, 2015; Kirk et al., 2015b). In some sense, CMB lensing is a clean probe; we know the source distribution extremely well (a δ-function at last scattering), and measurements are currently thought to be dominated by statistical, rather than systematic, uncertainties. It is therefore an excellent calibrator for galaxy lensing measurements.

CMB lensing – galaxy lensing cross-spectra contain cosmological information over and above the auto-spectra of either measurement alone, and combining the datasets can allow for the calibration of systematic effects particular to each measurement (Vallinotto, 2012; Das et al., 2013; Liu et al., 2016).

We now continue this introduction with a discussion of the state-of-the-art results in CMB lensing – galaxy lensing joint analysis.

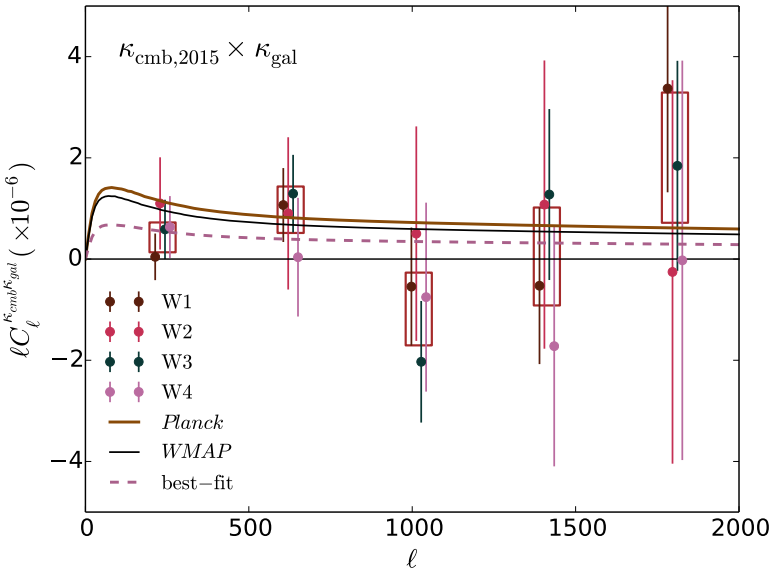

Figure 5.2: *Source: Liu & Hill (2015).* Cross-spectrum between *Planck* lensing and CFHTLenS galaxy lensing. The best fit amplitude is a factor $A = 0.44 \pm 0.22$ lower than that expected for a ΛCDM model and assumed source redshift distribution (*pink-dashed cf. red-solid*).

5.1.1 Measurements to date

The first measurement of the CMB lensing – galaxy lensing cross-spectrum was made by Hand et al. (2015), representing a 4.2σ detection of non-zero cross-correlation, using data from the CFHT Stripe 82 (CS82) survey and the Atacama Cosmology Telescope (ACT) over ≈ 160 deg^2. They report a cross-correlation amplitude $A = 0.78 \pm 0.18$ relative to that expected for a *Planck* best-fit cosmology, and given an assumed source redshift distribution. The measured cross-spectrum and best-fit model are shown in Fig. 5.1.

In Liu & Hill (2015) the galaxy convergence maps from CFHTLenS are cross-correlated with *Planck* CMB lensing measurements over 154 deg^2. Given the noise levels of *Planck* and CFHTLenS, a detection significance of $\approx 4.6\sigma$ was forecast – comparable to the result of Hand et al. (2015). However, the resulting amplitude was measured to be low by a factor of around 0.44 ± 0.22 relative to the ΛCDM expectation, with some small sensitivity to the cosmological parameters adopted (*Planck* vs. *WMAP*) and data used (*Planck* 2013 vs. *Planck* 2015). The data cross-spectra for each patch, and a combined model fit, are shown in Fig. 5.2. We discuss possible reasons for this anomalous result in Sec. 5.1.2.

The latest measurement – the cross-correlation between Dark Energy Survey (DES) lensing from science verification data and CMB lensing from *Planck* and SPT – was recently announced

in Kirk et al. (2015b). In both cases the inferred cross-correlation amplitude was found to be consistent with the ΛCDM expectation ($A_{\rm SPT} = 0.88 \pm 0.30, A_{\rm Planck} = 0.86 \pm 0.39$), with no modelling of instrinsic alignments. Including a simplistic model of intrinsic alignments (one which does not distinguish by galaxy type), gives an order 20% correction to the inferred amplitude (e.g., $A_{\rm SPT}^{\rm withIAs} = 1.08 \pm 0.36$), although this may be an overestimation of the alignment effect (Chisari et al., 2015; Larsen & Challinor, 2015).

5.1.2 Systematics in lensing measurements

Motivated by the results of Liu & Hill (2015), Hand et al. (2015) and Kirk et al. (2015b), we now discuss various systematic effects which affect weak-lensing measurements, and argue that the CMB lensing cross-correlation is an important calibration measurement, as well as adding cosmological information beyond galaxy-lensing measurements alone.

Galaxy lensing relies on the principle that, with no intervening structure, galaxy shapes would be randomly orientated when projected on the sky. The essential idea is then to search for coherent distortions in the shapes of nearby galaxies which would be generated by lensing from intervening matter, or *cosmic shear*.

The point-spread function (PSF) of an optical telescope – due to the finite size of the telescope dish, atmospheric variations and the telescope optics – is convolved with every pixel in a galaxy image; this smears the flux coming from a given direction across a small patch of sky around that direction. PSF spatial variation can artificially boost the observed galaxy lensing signal (mimicking Ω_m or σ_8), and must be carefully measured – using the shapes of nearby stars – and accounted for in the generative data model for the galaxy image. The bright stars used to infer the PSF can also broaden the profile relative to the field average due to the extra build-up of electrons on the detectors. This can lead to 'over-correction' for the PSF in a faint galaxy image (Mandelbaum, 2015). These effects must be handled rigorously to avoid inference biases and under-estimation of uncertainties. Use of the CMB lensing cross-correlation is particularly useful, then, since any PSF contamination of the estimated shear maps will be uncorrelated with the CMB lensing signal. Any inconsistency of the signal measured through the auto- and cross-spectra might then hint at PSF systematics in the galaxy data.

A *noise bias* can arise because the posterior for the galaxy shape parameters (e.g., centroid,

ellipticity and size) is, generally, highly non-Gaussian; the effect of these parameters on the galaxy image is non-linear, with galaxies typically only detected at all with a flux signal-to-noise in the range $10 \lesssim \nu_{\rm SNR} \lesssim 100$ (Miller et al., 2013). The signal-to-noise on the *shape* parameters of an individual galaxy will be correspondingly smaller than this. The use of a point estimator (such as the maximum likelihood or posterior mean) to summarise the inference of the galaxy properties will then naturally lead to a bias; the estimated properties, averaged over data realisations, do not coincide with the underlying quantity. This bias will depend on the properties of the source, such as size and flux signal-to-noise, and with current approaches must be corrected for empirically with simulations (Heymans et al., 2012; Jarvis et al., 2015).

We can write the data-averaged ellipticity $\mathbf{e}$ in any particular direction as

$$\langle \mathbf{e} \rangle = (1+m)\mathbf{g} + \alpha \mathbf{e}_{\rm PSF} + c, \tag{5.1}$$

where $\mathbf{g}$ is the reduced cosmological shear, $\mathbf{e}_{\rm PSF}$ is the shape of the local PSF, α represents the PSF-to-shape leakage, m is an induced *multiplicative bias* and c an *additive bias* arising from the effects described above (Heymans et al., 2006; Mandelbaum, 2015; Jarvis et al., 2015). At first order these biases can be considered to be scale-independent quantities, and can be controlled (calibrated) by inclusion of the CMB lensing cross-correlation in the analysis. For example, a constraint $\sigma(m) \approx 2\%$ on the multiplicative bias is forecast for the survey specification given in Das et al. (2013). This is possible because a multiplicative bias enters linearly in the cross-spectrum, but quadratically for the galaxy-lensing auto-spectrum.

Galaxies that are physically close to each other are expected to be intrinsically aligned at some level due to their common density or tidal environment. This gives rise to the so-called II term in the galaxy two-point correlation function: there exists a (scale-dependent) correlation between intrinsic galaxy shapes which must be modelled in order to make accurate inferences about the corresponding cosmic shear (GG) term. Foreground galaxies can additionally be correlated with background galaxies because they form in the same large-scale structure which lenses the background galaxy (the GI term). Recent measurements from CFHTLenS show early-type galaxies are intrinsically aligned, while late-types show a signal consistent with null, although we note current data are not yet sufficiently constraining to make a precise statement in this regard

(Heymans et al., 2013).

Various studies have shown that intrinsic alignments are also expected to contaminate the CMB lensing – galaxy lensing cross-correlation; this is a GI-type contribution, in which the large-scale structure which lenses the CMB also sources the intrinsic alignment of the galaxy (Troxel & Ishak, 2014; Hall & Taylor, 2014; Chisari et al., 2015; Larsen & Challinor, 2015). Intrinsic alignments affect galaxy lensing, but not an isolated CMB lensing measurement, so the cross-spectrum acts as a useful calibrator for a given galaxy population. A joint analysis of cosmic shear and CMB lensing, which combines all the auto- and cross-spectra, could allow for 'self-calibration' of the intrinsic alignment amplitude, and a full propagation of this uncertainty into cosmological inference (Hall & Taylor, 2014).

As demonstrated in Liu & Hill (2015), Hand et al. (2015) and Kirk et al. (2015b), the inferred amplitude of the empirical cross-spectrum is sensitive to the cosmological model adopted for the analysis. This is beneficial in that it shows, in principle, the power for this cross-spectrum to be used for precision cosmological analysis. Physically, different cosmological models have different late-time growth rates and therefore different levels of structure formation over the redshift range probed by the measurement. However, with current data, other systematic effects limit this interpretation. The choice of cosmological model can lead to a $\sim 15\%$ difference in the inferred amplitude A, since this is measured relative to the chosen model.

There is some tension in the *Planck*- and CFHTLenS-derived values of Ω_m and σ_8 (MacCrann et al., 2015; Kitching et al., 2016). These are particularly important parameters for determining the level of structure formation at late times. This could go some way to explaining the low cross-correlation amplitude found in Liu & Hill (2015); the same (unknown) effect which leads to the 'low' growth parameters in the CFHTLenS analysis might appear at some level in the cross-correlation with *Planck*.

A more correct approach would be to marginalise over cosmological parameter uncertainties when fitting the cross-spectrum; this would remove this systematic effect for future measurements (Pearson & Zahn, 2014). If other systematics can also be handled to a sufficient level, the cross-spectrum could in principle be folded in as an additional (and important) observable for cosmological parameter inference, with the amplitude and other cosmological parameters fit simultaneously (Vallinotto, 2012; Das et al., 2013).

An additional systematic effect comes from uncertainties on the redshift distribution of the source galaxies dn/dz. Because the source redshift distribution is inherent in the theory power spectra (e.g., Eq. 5.7), photometric-redshift uncertainties propagate into modelling uncertainty. Fixing the source distribution for analysis purposes will lead to biases on the inferred amplitude or cosmological parameters (e.g., Allison et al., 2015). Bonnett et al. (2015) consider the bias induced on cosmological parameters in using an incorrect dn/dz in the analysis, applying this to the science verification sample of DES galaxies. In particular, they find a 3% bias on σ_8, which is nevertheless less than the 1σ uncertainty in their cosmological analysis. Their approach is to introduce nuisance parameters (and corresponding priors) which quantify plausible variations in the shape of the underlying redshift distribution. These nuisance parameters are then simultaneously sampled with the primary parameters to perform marginalisation. Full propagation of photo-z uncertainties (quantified by a probability distribution $p(z)$ output by photo-z extraction codes) has yet to be implemented in a real cosmological analysis, but will become important for future deep and wide weak-lensing studies such as Euclid and LSST. This systematic effect is particularly significant for the CMB lensing cross-correlation because the underlying source redshift distribution is most uncertain at high redshifts, where the CMB lensing kernel peaks ($z \sim 2$), and therefore where much of the signal is generated.

Further systematic model uncertainties arise from the unknown amplitude and shape of the small-scale matter power spectrum. At low redshifts and small scales ($k \gtrsim 0.2$ Mpc), contributions to the matter power spectrum from non-linearly evolved structure are non-negligible; perturbation theory breaks down and baryonic feedback effects become important, making modelling the theory power-spectra complicated (Sec. 1.2.5; Smith et al., 2003). Again, a joint analysis of CMB lensing and galaxy lensing is beneficial here, since CMB lensing generally probes structure formation at higher redshifts where contributions from non-linear scales are less important than for galaxy-lensing alone.

Some combination of these systematic effects may explain the low amplitude for the cross-spectrum observed in Liu & Hill (2015). Recently, the effect of multiplicative bias on these results is investigated in more detail in Liu et al. (2016). We also note that, given the uncertainties, there is a non-negligible possibility that the result is simply a statistical fluctuation.

Many of the systematic effects discussed above are common to both the galaxy auto- and

cross-spectra with CMB lensing, but they may contribute as different factors in the different power spectra. This emphasises the power of the cross-correlation as a calibrator.

Although the level of systematic uncertainty in CMB lensing data is thought to be small – the source plane is well-understood, the kernel peaks at high redshift where non-linear effects are less important, and there is no intrinsic alignment-type contamination – there is still the possibility that other sources of non-Gaussianity in the CMB maps, or the adopted reconstruction approach, leave residual systematic effects in the lensing estimators. For instance, in the *Planck* 2015 data, a null-test for the TT curl estimator over the multipole range $400 \leq L \leq 2048$ is failed at the 3σ level, suggestive of an unknown systematic propagating into the lensing estimator (*Planck* Collaboration, 2015e). A powerful test – yet to be performed – would be a consistency analysis, at both the map and power spectrum level, between multiple CMB lensing experiments.

5.2 Weak-lensing cross-correlations in ΛCDM

Lensing convergence is defined by $\kappa = -\frac{1}{2}\nabla \cdot \boldsymbol{d}$, where $\boldsymbol{d}$ is the projected deflection on the sky, and in the weak-lensing regime is directly related to magnification: $\mu(\boldsymbol{\theta}) \approx 1 + 2\kappa(\boldsymbol{\theta})$. It represents an integrated measure of matter fluctuations δ between the source and observer (Bartelmann & Schneider, 2001):

$$\kappa(\boldsymbol{\theta}) = \int_0^{\chi_H} d\chi W(\chi)\delta(\boldsymbol{\theta}, \chi), \tag{5.2}$$

where the window function, or *kernel*, is given by

$$W(\chi) = \frac{3\Omega_{\rm m}H_0^2}{2c^2}\frac{f_k(\chi)}{a(\chi)}\int_\chi^{\chi_H} d\chi' p(\chi')\frac{f_k(\chi'-\chi)}{f_k(\chi')}, \tag{5.3}$$

a is the scale factor as a function of radial comoving distance χ, $f_k(\chi)$ is the transverse comoving distance, and p is the source density distribution. For the CMB,

$$p_{\rm c}(\chi) = \delta(\chi - \chi^*), \tag{5.4}$$

where χ^* is the comoving distance to last-scattering, while for galaxies

$$p_{\rm g}(\chi) = \frac{dn}{dz}\frac{dz}{d\chi}, \tag{5.5}$$

where dn/dz is the normalised redshift distribution of sources, which in practice is measured photometrically or spectroscopically. Statistical isotropy of the convergence fields allows us to write the 2-point correlation function in Fourier space as follows[1]:

$$\langle \kappa_{\rm c}(\boldsymbol{\ell})\kappa_{\rm g}^{*}(\boldsymbol{\ell}')\rangle = C_{\ell}^{\kappa_c\kappa_g}\delta(\boldsymbol{\ell}-\boldsymbol{\ell}'). \tag{5.6}$$

Using Eq. 5.2–5.3, the cross-spectrum can be written in the Limber approximation as

$$C_{\ell}^{\kappa_c\kappa_g} = \int_0^{\chi_H} d\chi \frac{W_{\kappa_c}(\chi)W_{\kappa_g}(\chi)}{f_k(\chi)^2} P\left(k=\frac{\ell}{f_k(\chi)},\chi\right) \tag{5.7}$$

(Eq. 1.20; Bartelmann & Schneider, 2001). The cross-spectrum is therefore a weighted integral of the matter power spectrum, and encodes cosmological parameter dependence through its sensitivity to growth (via P) and geometry. The signal is sensitive to structure at the redshifts z over which there is significant overlap of the two kernels W_i; typically $0.2 \lesssim z \lesssim 1.5$ for current measurements. The cross-spectrum also quantifies the (scale-dependent) correlation ρ between the fields:

$$\rho(\ell) = \frac{C_{\ell}^{\kappa_c\kappa_g}}{\sqrt{C_{\ell}^{\kappa_c\kappa_c}C_{\ell}^{\kappa_g\kappa_g}}}. \tag{5.8}$$

For any given set of cosmological parameters, one can compute the theory matter power spectrum $P(k,\chi)$ using a Boltzmann solver such as CAMB or CLASS (Lewis et al., 2000; Lesgourgues, 2011). Using this output, we use our own python code to compute the relevant distances, and window functions, to perform the line-of-sight integral in Eq. 5.7. We plot $C_{\ell}^{\kappa_c\kappa_g}$ and $\rho(\ell)$ for the fiducial cosmology, for various source redshift distribution, in Fig. 5.3–5.4.

The amplitude of the cross-spectrum increases for a source population at higher redshift. Physically, this is because the photons being emitted by the source galaxies traverse more dark matter fluctuations and are deflected by more potential wells, increasing the variance in the deflection angle and the correlation with dark matter fluctuations which also lens the CMB.

In the large-angular-scale part of the cross spectrum (the low-ℓ regime), the non-negligible contributions to the integral above typically require that the comoving distance χ be small so

[1] Eq. 5.6 is appropriate in the flat sky approximation, but is easily generalised to the exact curved sky expression by replacing the Fourier modes $\kappa(\boldsymbol{\ell})$ with spherical harmonic coefficients $\kappa_{\ell m}$ and setting $\delta(\boldsymbol{\ell}-\boldsymbol{\ell}') \to \delta_{\ell\ell'}\delta_{mm'}$ (Sec. 1.1).

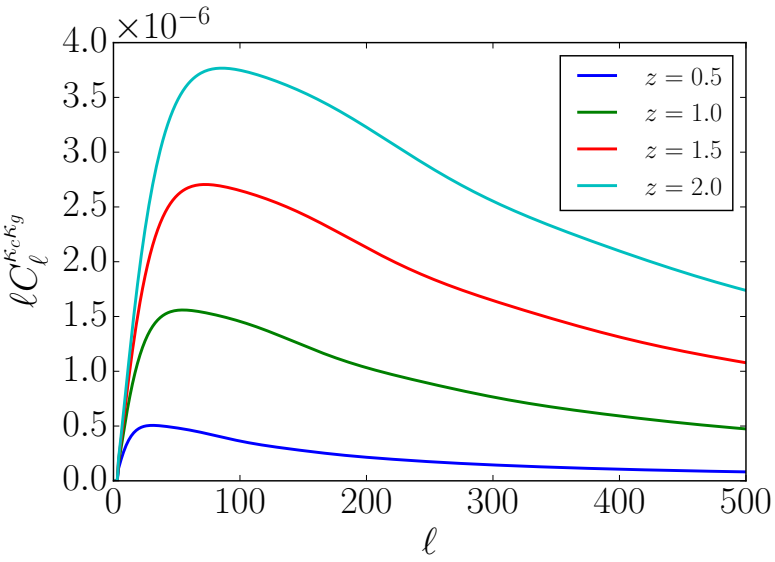

Figure 5.3: The theoretical cross-spectrum between the CMB lensing and galaxy lensing fields for a narrow source distribution centred at various redshifts z. Contributions to the small-scale cross-spectrum from the non-linear evolution of the matter power spectrum are significant from $\ell \gtrsim 100$ (Liu & Hill, 2015). The peak moves to the right for sources at higher redshift as the typical lenses are more distant, subtending a smaller angle on the sky. The amplitude also increases both because the galaxy lensing power increases (photons traverse more structure from source to telescope), and because the correlation coefficient between fields also increases as more of that structure is common between galaxies and the CMB (Fig. 5.4).

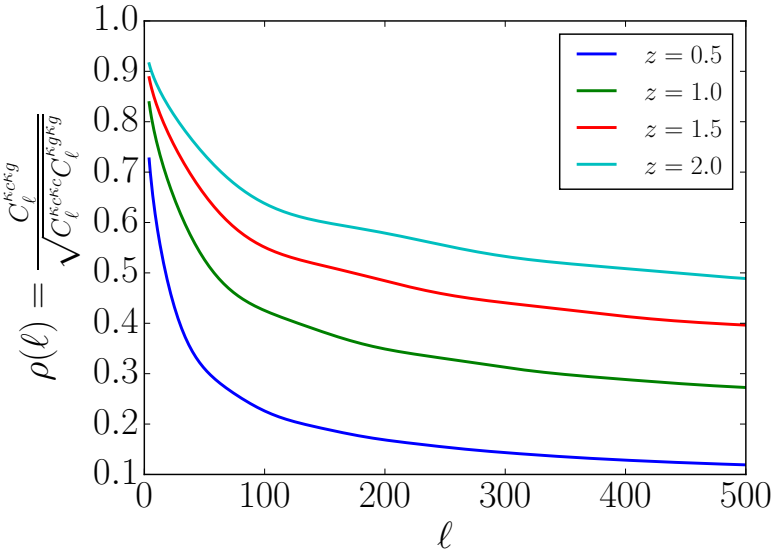

Figure 5.4: The cross-correlation coefficient between the CMB and galaxy lensing fields. The correlation between the fields increases with angular scale (decreasing ℓ), as more weight in the integral (Eq. 5.7) is given to the common, low-redshift matter fluctuations.

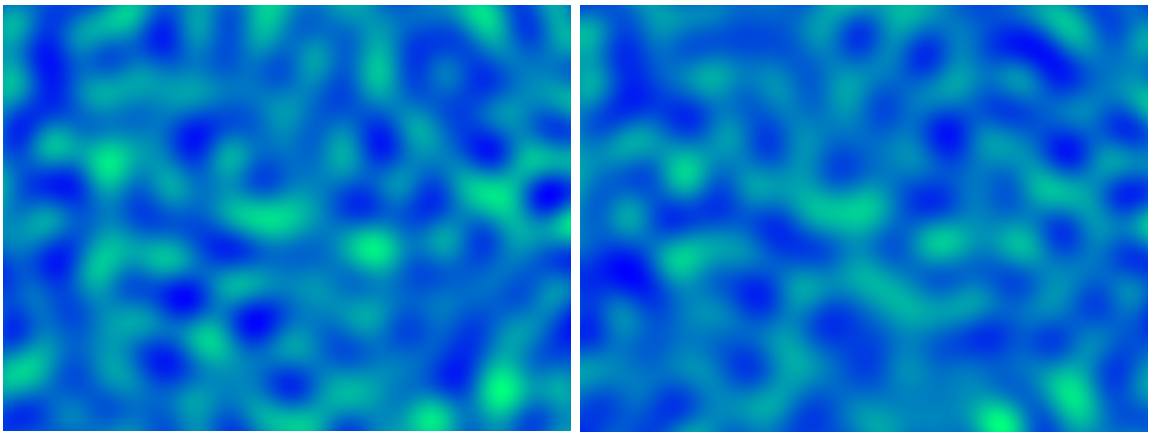

Figure 5.5: Simulation of a CMB lensing convergence map (*left*) and the corresponding galaxy convergence field (*right*), for a source distribution of galaxies peaking at redshift $z = 2$. The maps cover a 40 deg × 30 deg area and are low-pass filtered for visual clarity (removing power on angular scales $\theta \lesssim 2$ deg). The correlation between these dark-matter tracers is clearly visible by eye; there is a large fraction of common structure in both probes (*cf.* Fig. 5.6).

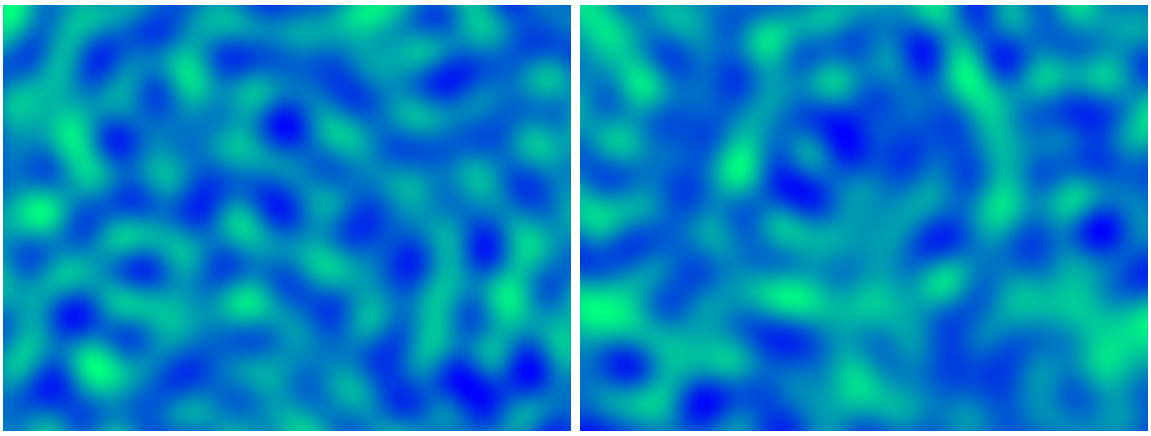

Figure 5.6: Simulation of a CMB lensing convergence map (*left*) and the corresponding galaxy weak-lensing field (*right*), for a source distribution of galaxies peaking at redshift $z = 0.5$. The maps cover a 40 deg × 30 deg area and are low-pass filtered for visual clarity (removing power on angular scales $\theta \lesssim 2$ deg). The low-redshift sources pick out larger-scale matter fluctuations than in Fig. 5.5. The correlation between the fields is also less obvious; there is a lower fraction of intervening structure in common (*cf.* Fig. 5.5).

that the modes picked out are near the peak of the matter power spectrum (Eq. 5.7). In this regime, nearby lenses carry more weight, and these are common to the galaxies and the CMB (Fig. 5.4). At large scales there is therefore increased correlation between the galaxy and CMB lensing fields.

To further illustrate the properties of the signal, in Fig. 5.5–5.6 we show simulated maps of the correlated CMB lensing and galaxy lensing fields for two different redshift distributions.

A modified theory of gravity can affect all terms in Eq. 5.7. Typically the Poisson equation of a modified gravity theory, which relates the gravitational potentials to density perturbations δ, is different than in a ΛCDM model, or there is a non-trivial *slip relation* between the gravitational potentials ϕ and ψ. The growth of (typically large-scale) fluctuations is modified, which affects the matter power spectrum $P(k, z)$, and there is a modification of cosmological distances, which affects the geometry and kernel terms. For a detailed study of the effects of a modified theory of gravity on weak-lensing observables, see Schmidt (2008) and Leonard et al. (2015).

5.3 Joint analysis of CMB lensing and galaxy lensing

The aim of this chapter is to motivate and develop the analysis tools for weak-lensing cross-correlations with ACTPol and AdvACT. In this section we continue to investigate the signal and noise properties of the observables, using AdvACT as an example. We show the projected lensing noise power spectra for AdvACT for each of the different lensing estimators and their minimum variance (MV) combination in Fig. 5.7. We covered the physics and extraction of the CMB lensing signal in detail in Sec. 1.4, where we detailed how these curves are derived. The TT reconstruction continues to provide the most lensing information at AdvACT noise levels, but the TE, EE and EB estimators are also important, with the MV combination significantly improved over the temperature estimator alone. The lensing modes can be mapped with $SNR > 1$ at scales $L \lesssim 350$.

In Fig. 5.8–5.9 we show simulated AdvACT convergence maps, given these noise properties. In Fig. 5.10 we plot the CMB lensing auto-spectrum signal-to-noise, as a function of scale, for ACTPol and AdvACT. The measured power spectrum will be signal dominated below $L \lesssim 700, 1500$ respectively, except at the largest scales (say $L \lesssim 50$) where – in practice – the variance of the mean-field induced by windowing will suppress the signal-to-noise.

We now outline the construction of a (noisy) map of galaxy convergence from a shear catalogue which has been previously inferred from the raw photometry. Typically the catalogue records angular sky position, photometric redshift, inferred ellipticity $e_{j,i}$ ($j \in \{1, 2\}$ labels the two components of ellipticity) and weight w_i accorded to each galaxy i, which is defined to maximise the signal-to-noise of the shear measurement (Miller et al., 2013).

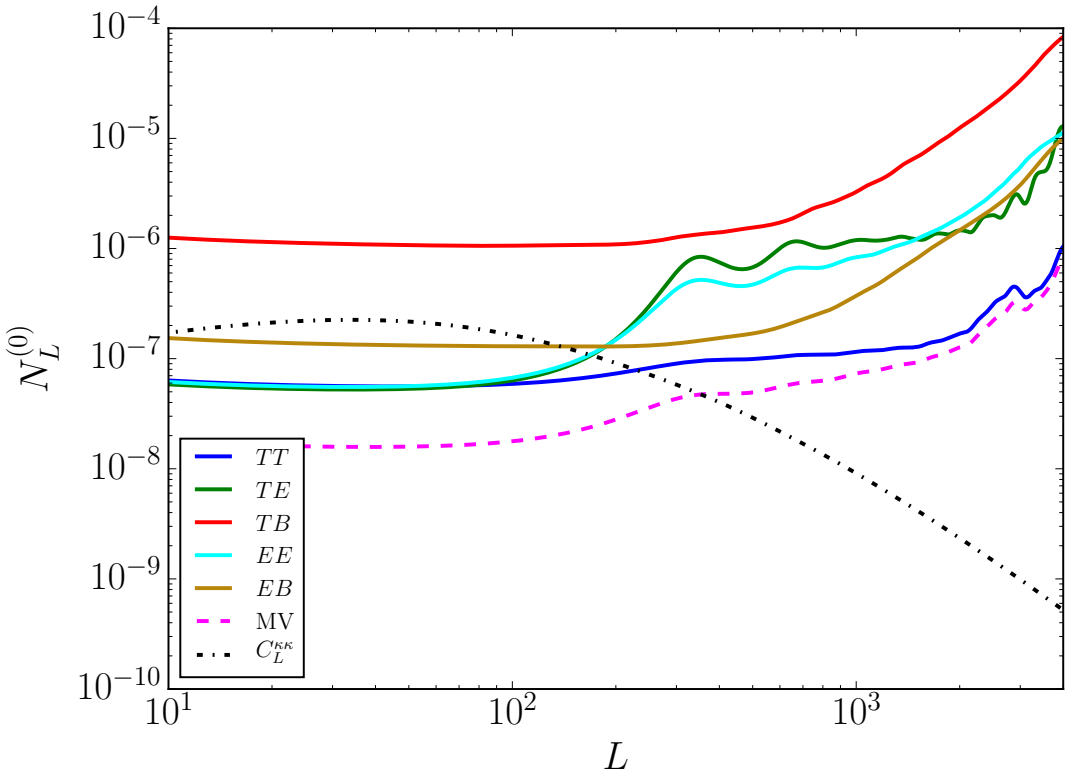

Figure 5.7: Projected reconstruction noise variance $N_L^{(0)}$ for each quadratic estimator, for AdvACT, assuming white-noise only. The BB quadratic estimator has negligible signal-to-noise in allowed cosmologies ($r < 0.2$). The signal power-spectrum $C_L^{\kappa\kappa}$ is the black dot-dashed curve. Reconstruction noise for the minimum-variance combination of estimators is shown as the pink dashed curve. The TT reconstruction provides the most lensing information at AdvACT noise levels, but the TE, EE and EB estimators are also important. The lensing modes can be mapped with signal-to-noise > 1 at scales $L \lesssim 350$.

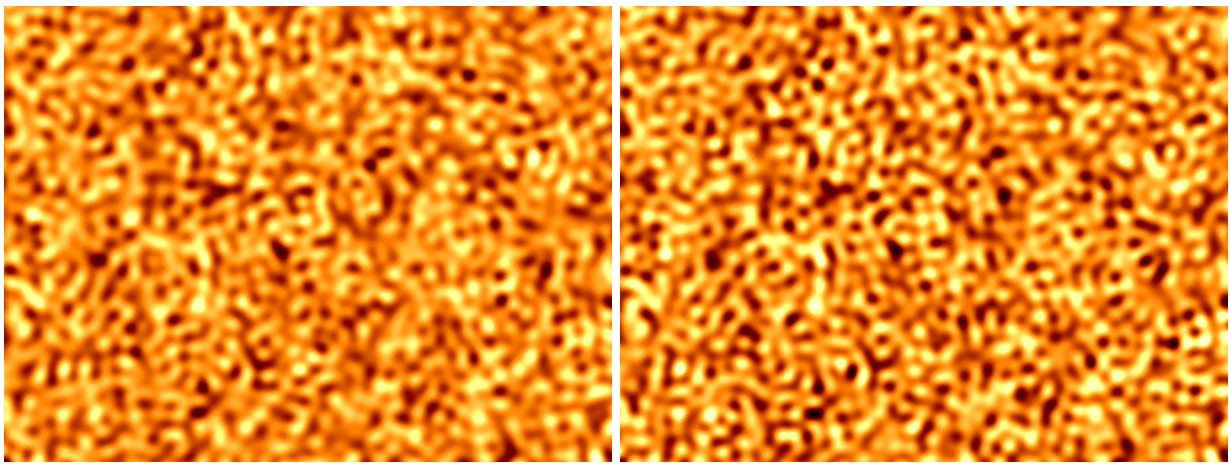

Figure 5.8: *Left panel:* Simulated CMB lensing convergence map over a 40 deg × 30 deg patch. *Right panel:* Simulated AdvACT minimum-variance reconstruction of this signal, produced by adding to the signal a realisation of the noise power spectrum in Fig. 5.7. Both maps are smoothed to remove fluctuations on scales $L > 350$. Large-scale fluctuations are mapped precisely and features are visible by eye. At the smallest visible scales (≈ 0.5 deg), the signal and noise have comparable power (*cf.* Fig. 5.9).

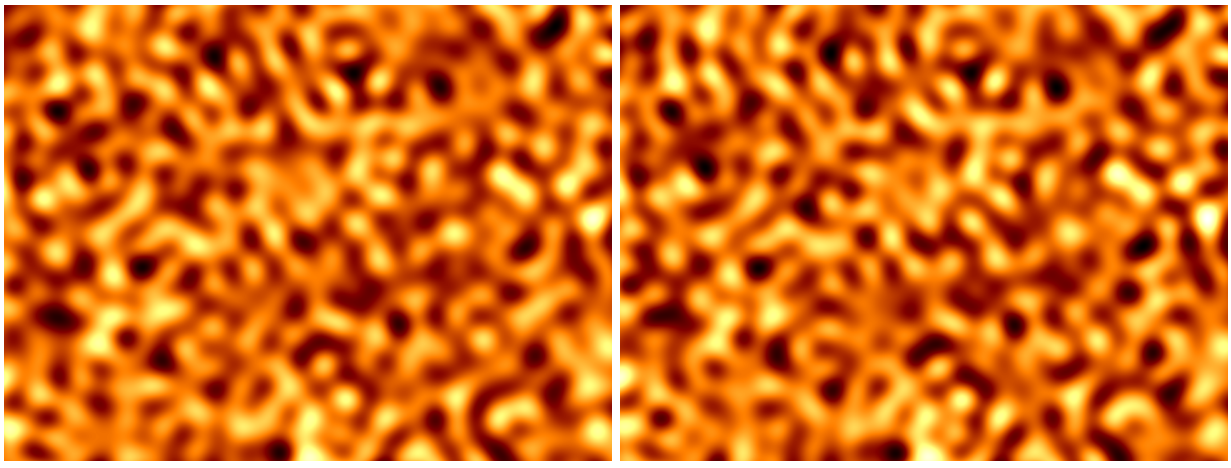

Figure 5.9: *Left panel:* Simulated CMB lensing convergence map over a 40 deg × 30 deg patch. *Right panel:* Simulated AdvACT minimum-variance reconstruction of this signal, produced by adding to the signal a realisation of the noise power spectrum in Fig. 5.7. Both maps are smoothed to remove fluctuations on scales $L > 180$ (angular scales < 1 deg). The map is signal-dominated at all visible scales (*cf.* Fig. 5.8)

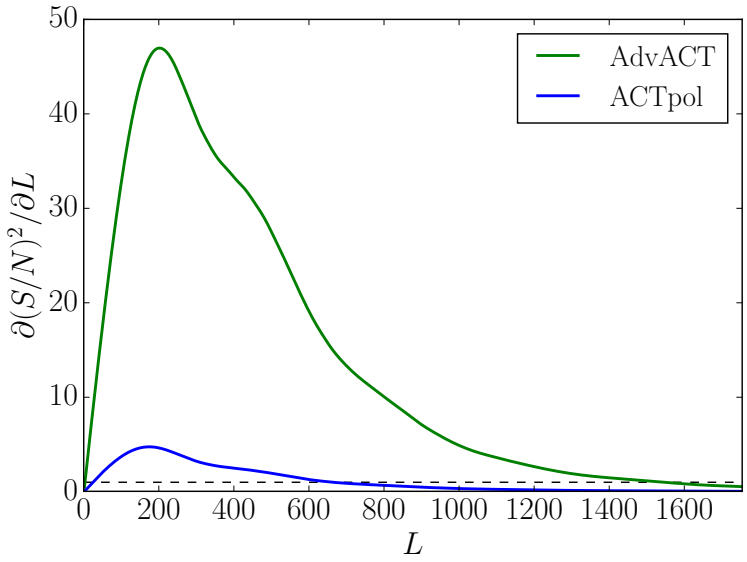

Figure 5.10: Forecast signal-to-noise (SNR) on the CMB lensing power spectrum $C_L^{\kappa_c\kappa_c}$ per angular multipole L. The overall detection significance is given by the square root of the integral under each curve. For the full ACTPol survey, this equates to a $\approx 40\sigma$ detection, while the lensing power spectrum is expected to be measured by AdvACT at $\approx 150\sigma$. The dashed horizontal line indicates a signal-to-noise of unity. Both forecasts assume white noise and the specifications given in Table 5.2.

Shear maps are then constructed by averaging ellipticities within each pixel $\boldsymbol{n}$:

$$\hat{\gamma}_j(\boldsymbol{n}) = \frac{\sum_{i\in\boldsymbol{n}} w_i e_{j,i}}{\sum_{i\in\boldsymbol{n}} w_i(1+m_i)}. \tag{5.9}$$

The multiplicative bias m_i assigned to each galaxy is calibrated through simulations and is a function of the properties of the source, e.g., flux signal-to-noise and shape (Miller et al., 2013). A minimum variance estimate of the convergence field is then inferred as a weighted linear-combination of the shear maps in Fourier space, by Kaiser-Squires inversion (Kaiser & Squires, 1993):

$$\hat{\kappa}(\boldsymbol{\ell}) = \hat{\gamma}_1(\boldsymbol{\ell})\frac{\ell_x^2-\ell_y^2}{\ell^2} + \hat{\gamma}_2(\boldsymbol{\ell})\frac{2\ell_x\ell_y}{\ell^2}. \tag{5.10}$$

This can then be low pass-filtered, ensuring the estimator has finite variance in real-space, by suppressing the (spectrally-flat) shot noise in the high-frequency part of the spectrum (Kaiser & Squires, 1993; Van Waerbeke et al., 2013; Hand et al., 2015).

Using Eq. 1.70, $\hat{\kappa_c} = \ell(\ell+1)\hat{\phi}_c/2$ and Eq. 5.10, we define the binned *pseudo-spectrum*

$$\tilde{C}_b^{\kappa_c\kappa_g} \equiv \sum_{\boldsymbol{\ell}} P_{b\boldsymbol{\ell}} \Re[\hat{\kappa_c^*}(\boldsymbol{\ell})\hat{\kappa_g}(\boldsymbol{\ell})] \tag{5.11}$$

where the binning matrix $P_{b\boldsymbol{\ell}}$ weights modes $\boldsymbol{\ell}$ which fall in each bin; we take $P_{b\boldsymbol{\ell}} = 1/n_b$ if $\boldsymbol{\ell} \in b$ and zero otherwise, where n_b is number of modes lying in bin b, although more sophisticated binning schemes such as inverse-variance weighting are possible (e.g., Das et al., 2011). We define the binned *mode-coupling matrix* $M_{bb'}$ by

$$M_{bb'} \equiv \sum_{\boldsymbol{\ell}\boldsymbol{\ell}'} P_{b\boldsymbol{\ell}} M_{\boldsymbol{\ell}\boldsymbol{\ell}'} Q_{\boldsymbol{\ell}'b}, \tag{5.12}$$

where the inverse binning operator $Q_{\boldsymbol{\ell}'b} = 1$ if $\boldsymbol{\ell} \in b$ and zero otherwise, and $M_{\boldsymbol{\ell}\boldsymbol{\ell}'} = W_{\kappa_c}(\boldsymbol{\ell} - \boldsymbol{\ell}')W_{\kappa_g}(\boldsymbol{\ell} - \boldsymbol{\ell}')$ is the mode coupling due to the window W_i applied to each field, which includes any masking, apodisation, real-space signal-to-noise weighting or beam smoothing. The binned CMB lensing – galaxy lensing cross-spectrum is then estimated as

$$\hat{C}_b^{\kappa_c\kappa_g} = \sum_{b'} M_{bb'}^{-1} \tilde{C}_{b'}^{\kappa_c\kappa_g}. \tag{5.13}$$

Eq. 5.13 is an unbiased estimator of the underlying (binned) cross-spectrum (Das et al., 2013; Hand et al., 2015; Liu & Hill, 2015). Inclusion of the mode-coupling correction is an essential step, particularly given the complicated masking and filtering operations required for real data; e.g., Fig. 5.11 shows the highly-structured window function for the CS82 galaxy survey.

5.4 Cross-correlation pipeline

In this section we present the full procedure from map making to parameter estimation utilised in our cross-correlation pipeline. This is currently the main pipeline for ACT cross-correlation analysis, and is built upon that used in Hand et al. (2015) for the ACT×CS82 analysis. We consider two illustrative examples to demonstrate the pipeline in action.

The first example is the cross-correlation between ACTPol lensing convergence measurements, from season one observations in the D6 field, with CFHTLenS observations in the W1 field (Hey-

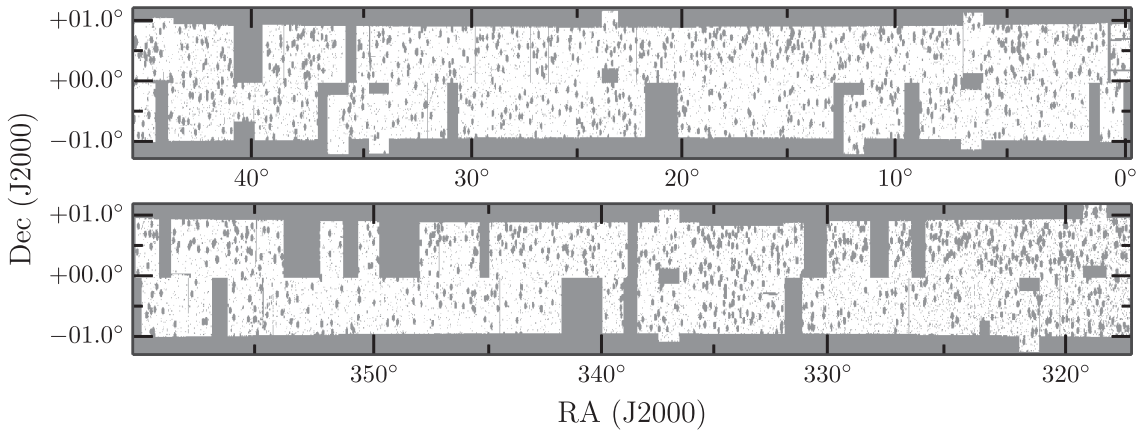

Figure 5.11: *Source: Hand et al. (2015).* The CS82 mask defines the observed footprint and excludes pixels contaminated by bright stars and other image artifacts (grey regions masked). The masking of a Gaussian field correlates modes, mimicking the effects of lensing. In power-spectrum estimation the mode-coupling matrix $M_{bb'}$ must accounted for to obtain an unbiased estimate of the underlying cross-spectrum (Eq. 5.13).

mans et al., 2012; Naess et al., 2014). We show these two convergence maps in Fig. 5.12. The overlap area is ≈ 30 deg^2. Given the noise properties of each map, and applying the Knox formula (Eq. 5.15), the expected detection significance is around 1σ (i.e., extremely marginal); nevertheless this example serves as a useful test case for the cross-correlation pipeline.

As a second example – to demonstrate the capability of this pipeline with high-fidelity data – we consider simulated observations of the full ACTPol survey cross-correlated with a hypothetically complete KiDS-North survey, a measurement planned by the two collaborations. This would correspond to ≈ 750 deg^2 across the ACTPol BOSS-N field. The simulated ACTPol and KiDS convergence maps, and their apodised mask, are shown in Fig. 5.13. The specifications for the two surveys assumed for this analysis are given in Tables 5.1–5.2. The window function is constructed by masking out 250 circles of radius u arcmin, where $u \sim U[0, 20]$, at random positions over the 750 deg^2 field. To create the final window function, this mask is then smoothed with a Gaussian of FWHM = 5 arcmin ($\sigma \approx 2$ arcmin), the map edges padded by 20 arcmin and then apodised by a Gaussian of FWHM = 10 arcmin. The CMB and galaxy convergence maps are drawn as correlated Gaussian realisations with zero mean and covariance matrix defined by the auto- and cross-power spectra in the fiducial model: $C_\ell^{\kappa_c\kappa_c}, C_\ell^{\kappa_c\kappa_g}$ and $C_\ell^{\kappa_g\kappa_g}$. Scale-dependent Gaussian noise is added to each according to their noise spectra $N_\ell^{\kappa_c\kappa_c}$, $N_\ell^{\kappa_g\kappa_g}$ which are derived from realistic survey specifications given in Tables 5.1–5.2. All these spectra are shown in Fig. 5.18.

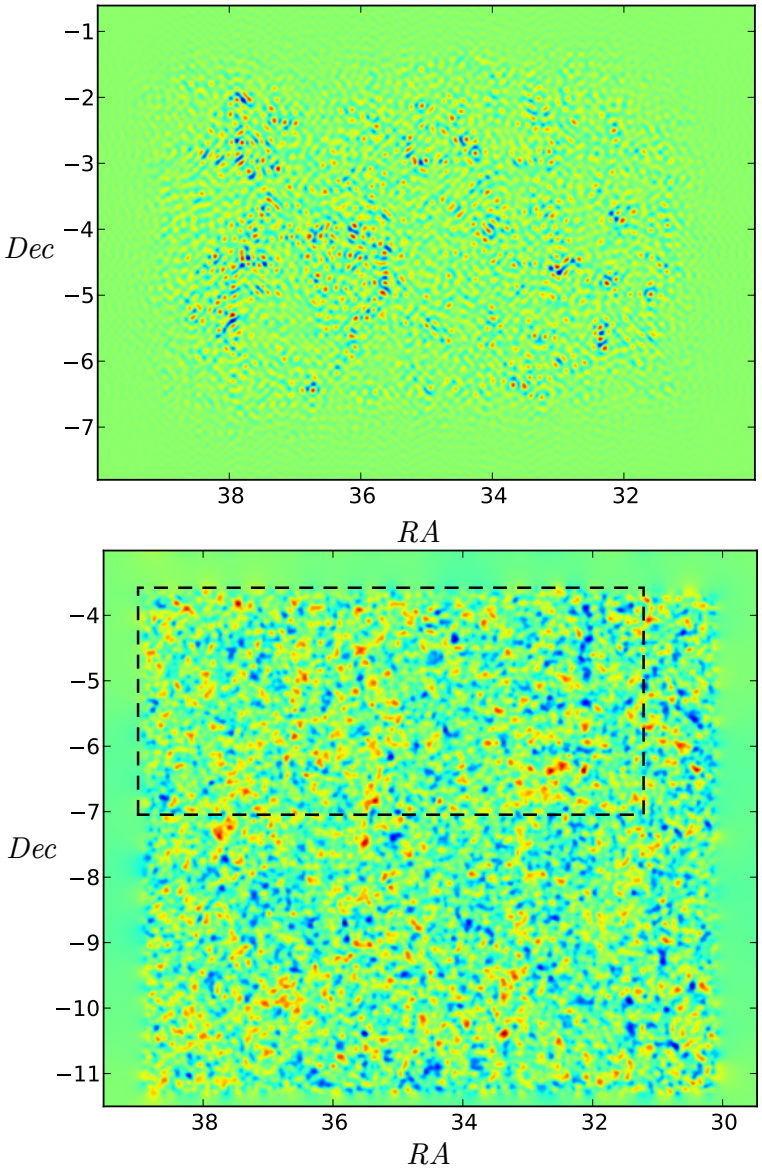

Figure 5.12: *Top panel:* ACTPol lensing convergence map across the D6 field. The map is modulated by two factors of the pixel hitmap – which counts the number of observations in each pixel – explaining the suppression of the signal towards the map edge. The inverse mode-coupling matrix accounts for the total window function to define an unbiased estimator of the cross-spectrum (Eq. 5.11). *Bottom panel:* CFHTLenS galaxy convergence map across the W1 field, constructed as described in Sec. 5.3. The dashed rectangle approximates the area of overlap of the ACTPol and CFHTLenS observations, totalling ≈ 30 deg^2.

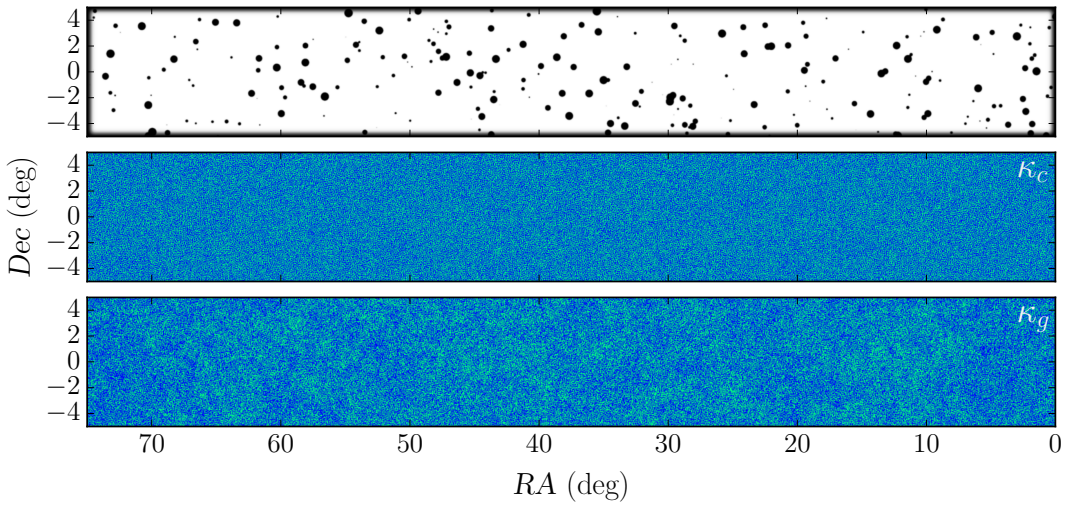

Figure 5.13: The window function (*top panel*), ACTPol convergence (*middle panel*) and KiDS convergence fields (*bottom panel*) used for the cross-correlation simulation example of Sec. 5.4. These maps are unfiltered to demonstrate the raw data. The CMB convergence correlation function is dominated by the noise reconstruction, while large-scale modes are visible in the galaxy convergence field on degree scales which correspond to fluctuations in the lensing signal.

We now outline the recipe followed by our pipeline from galaxy catalogue to parameter estimation:

1. Construction of the galaxy convergence field as described in Sec. 5.3. The map used in the analysis is constructed over the same area and at the same pixelisation as the CMB lensing map; this ensures the same number and size of Fourier pixels in each map, allowing for direct multiplication when constructing the pseudo-C_ℓ's.

2. Application of the apodised real-space galaxy lensing mask to discount regions over which the shear estimator is not well-defined (e.g. nearby bright stars). Apodisation (edge smoothing) removes the high-frequency content of the mask which would otherwise lead to the coupling of arbitrarily separated modes, making the cross-spectrum estimator ill-defined.

3. Apodisation of the map edges, in both maps, for the same reasons as above; ensuring periodicity and preventing *ringing* when moving to Fourier space.

4. Calculation of the mode-coupling matrix $M_{bb'}$, as in Eq. 11 of Das et al. (2011), accounting for the effects of binning, windowing, apodisation and beams. We show the mode-coupling matrix for the simulated ACTPol×KiDS analysis in Fig. 5.14.

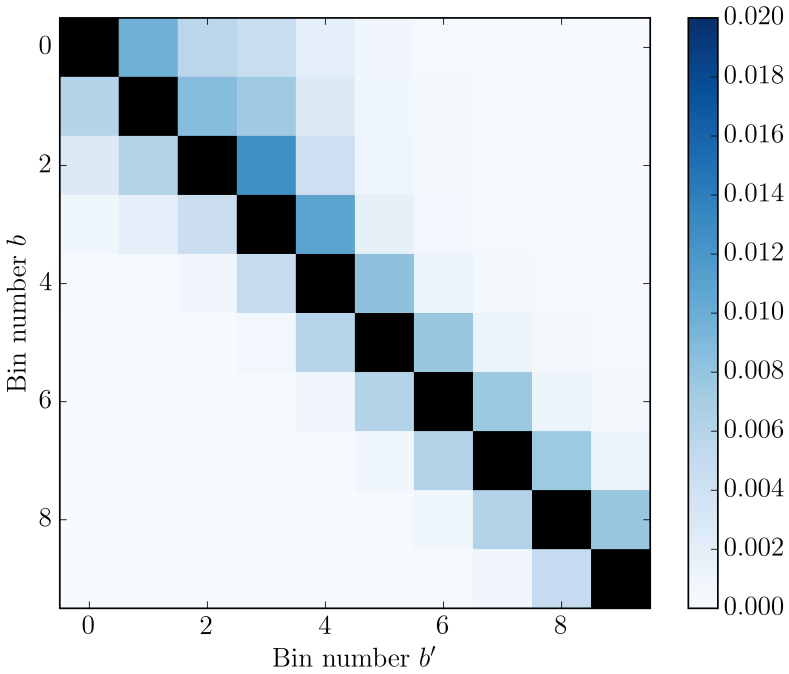

Figure 5.14: Mode-coupling matrix $M_{bb'}$ for the simulated ACTPol×KiDS analysis, which quantifies the contribution of an underlying power spectrum bandpower b to an observed bandpower b'. The matrix has been normalised by its diagonal elements; the off-diagonal terms then represent the *correlations* between the bandpowers. In this case, correlations are small (of order 1%) since the bins are wide ($\Delta\ell \approx 300$), the patch is large ($f_{\rm sky} = 0.02$), and the mask reasonably well-behaved (Fig. 5.13).

5. Inversion of the mode-coupling matrix and calculation of the binned power spectrum estimator $\hat{C}_b^{\kappa_c\kappa_g}$ by Eq. 5.11. The power spectra for our working examples are shown in Figs. 5.15–5.16, along with the ΛCDM expectation.

6. Empirical calculation of the variance of the power spectrum estimator through simulations. Simulated CMB lensing maps, with the same properties as the data map (e.g., the same 2D noise power), are cross-correlated with the galaxy lensing data and the resulting power spectrum estimated as in the step 5. The dispersion of these estimators $\hat{C}_b^{\kappa_c^{\rm sim}\kappa_g}$ corresponds to the variance of the sampling distribution, directly setting the posterior uncertainties on the estimated power spectrum given a uniform prior on each bandpower. This also serves as a pipeline null test, since the average cross spectrum $\langle\hat{C}_b^{\kappa_c^{\rm sim}\kappa_g}\rangle$ should be consistent with zero. This approach is superior to using analytical uncertainties (Eq. 5.15) since it naturally accounts for the additional variance induced by the windowing and beam operations.

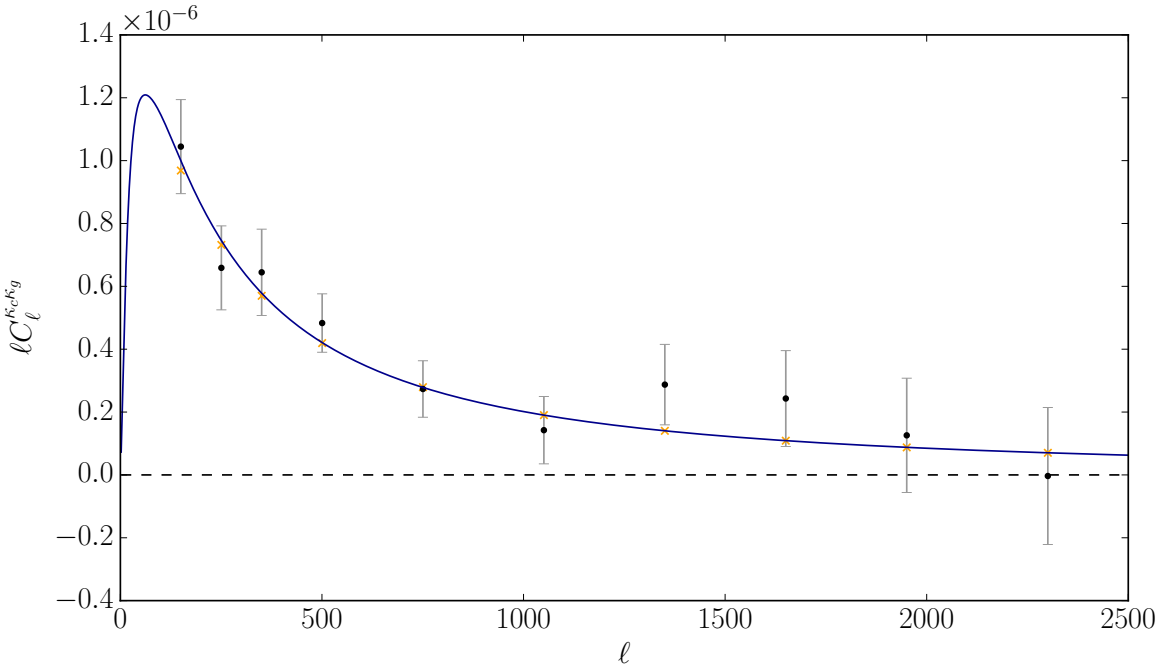

Figure 5.15: *Black data points:* Lensing cross-spectrum inferred for the ACTPol×KiDS analysis – note that the input maps are *simulated* data. The error bars are calculated using Eq. 5.15 and the data-derived auto- and cross-power spectra. We show the input cross-spectrum in blue and the corresponding theory bandpowers as orange crosses. The recovered cross-spectrum is visibly consistent with the input.

7. A binned version of the theory cross spectrum $C_b^{\kappa_c\kappa_g}$ can be computed through Eq. 5.7, and the inferred amplitude of the data A relative to the fiducial ΛCDM model readily constrained. We show this inference for the ACTPol×KiDS example in Fig. 5.17. With higher signal-to-noise, the cross-spectrum likelihood can be combined with the auto-spectra for more comprehensive cosmological parameter inference.

The recovered ACTPol×KiDS cross-spectrum is seen to be consistent with the input data (Fig. 5.15). This statement can be quantified by computing the χ^2 statistic, relevant for Gaussian distributed data d_i scattering around a model f_i:

$$\chi^2(A) \equiv \sum_i \left(\frac{d_i - f_i(A)}{\sigma_i} \right)^2 , \tag{5.14}$$

where i labels the data point and here A is the amplitude of the cross-spectrum relative to the fiducial model. This formula can easily be generalised for correlated data points. For the

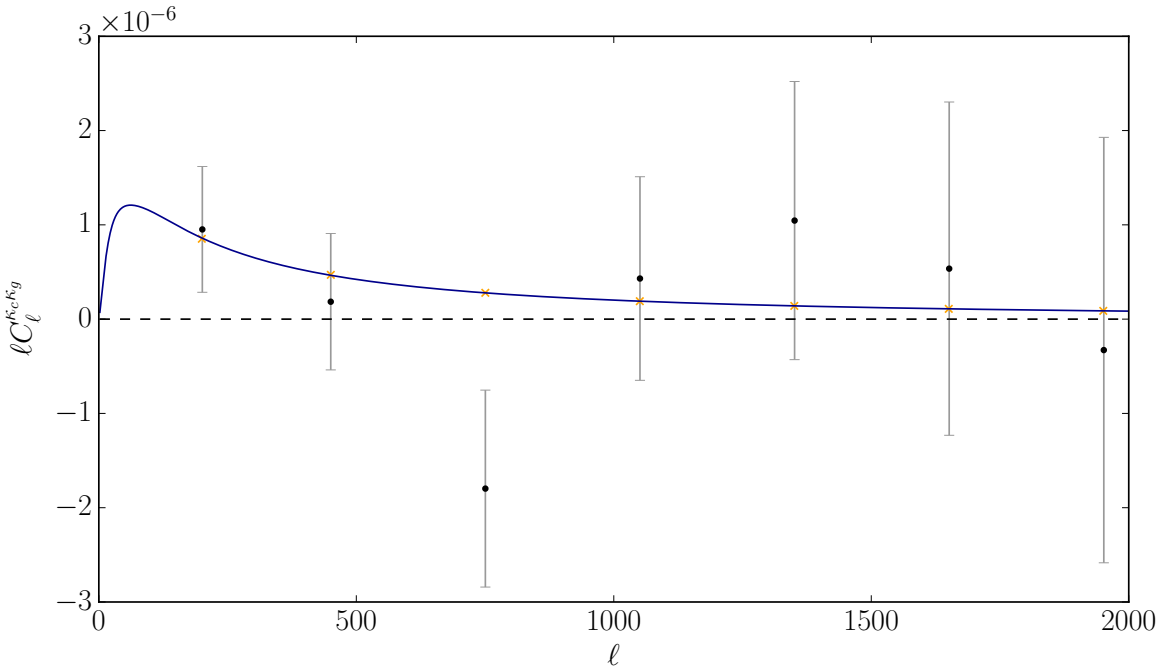

Figure 5.16: *Black data points:* Lensing cross-spectrum inferred for the ACTPol×CFHTLenS analysis. The error bars are calculated using Eq. 5.15 and the data-derived auto- and cross-power spectra. We show the ΛCDM expectation for cross-spectrum in blue and the corresponding theory bandpowers as orange crosses. The measured cross-spectrum is consistent with the theoretical expectation ($\chi^2 = 4.6$, d.o.f. $= 7$), but also a null correlation hypothesis ($\chi^2_{\rm null} = 5.8$).

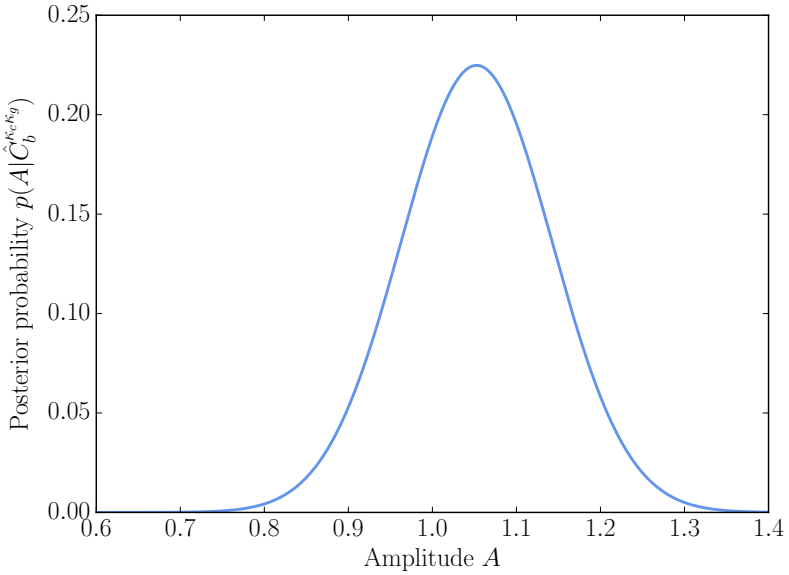

Figure 5.17: Posterior probability distribution for the amplitude A of the ACTPol×KiDS cross-spectrum, relative to the input cross-spectrum ($A = 1$, Fig. 5.15). We assume a flat prior on A, and this Gaussian posterior can be summarised as $A = 1.05 \pm 0.09$.

ACTPol×KiDS example, the uncertainties σ_i are derived using Eq. 5.15 (below) and the data-derived auto- and cross-power spectra. For the fiducial model ($A = 1$), we find $\chi^2 = 3.8$ for 10 degrees of freedom; the probability-to-exceed (PTE) this χ^2 value is 0.95.

The amplitude inferred for the ACTPol×KiDS cross-spectrum, relative to the fiducial model, is $A = 1.05 \pm 0.09$ (Fig. 5.17). Naively computing the empirical power spectrum directly from the windowed (i.e. observed) fields, but neglecting the effect of the mode-coupling matrix, gives an inferred amplitude almost 2σ lower than the correct analysis: $A_{\rm naive} = 0.88 \pm 0.09$. This demonstrates the necessity of using an unbiased estimator, even in this case of a relatively large field, a wide binning scheme and a non-pathological mask (Fig. 5.13).

In the case of the ACTPol×CFHTLenS cross-spectrum, we find $A = 0.77 \pm 0.68$, consistent with the ΛCDM expectation ($\chi^2 = 4.6$, PTE = 0.7), but also consistent with a null signal ($\chi^2_{\rm null} = 5.8$, PTE = 0.56). The detection significance is comparable to that expected for a patch of this size and given the noise properties of each map.

For the purpose of demonstrating the capabilities of the cross-correlation pipeline, we begin the simulated ACTPol×KiDS analysis by drawing noisy, correlated Gaussian realisations given the convergence signal and noise power spectra. We note that, instead, we might have started the analysis with simulations of the noisy CMB and a realistic galaxy catalogue. As described in Sec. 1.4 and Sec. 5.3, these could then be pushed through our map reconstruction pipeline to produce more representative simulated observations. This would capture physical effects which are missing from the analysis presented here, such as the Poisson sampling of the shear field, and the non-Gaussian effect of lensing on the CMB fields. Although we have the machinery to include this extra step of sophistication, it is unnecessary for demonstrating the essential performance of the cross-correlation analysis pipeline.

5.5 Forecasts for future measurements

We now discuss the measurement of the CMB lensing – galaxy lensing cross-spectrum with future data, forecasting the detection significance expected for various data combinations given their noise and survey properties.

The basis for the forecasting approach is to compute the uncertainties on the theory cross-

spectrum as a function of angular multipole ℓ, comparing this to the magnitude of the expected signal. Since the CMB and galaxy lensing fields are approximately Gaussian, using Wick's theorem the estimated cross-spectrum will have sampling variance

$$\sigma^2(\hat{C}_\ell^{\kappa_c\kappa_g}) = \frac{1}{(2l+1)f_{\rm sky}}\left[(C_\ell^{\kappa_c\kappa_c}+N_\ell^{\kappa_c\kappa_c})(C_\ell^{\kappa_g\kappa_g}+N_\ell^{\kappa_g\kappa_g})+(C_\ell^{\kappa_c\kappa_g})^2\right], \tag{5.15}$$

where the signal is estimated by $\hat{C}_\ell^{\kappa_c\kappa_g} = \frac{1}{2\ell+1}\sum_m \Re[\kappa_c^{\ell m *}\kappa_g^{\ell m}]$ (or Eq. 5.11 in the flat-sky approximation) for measured harmonic modes $\kappa^{\ell m}$, and $f_{\rm sky}$ accounts for the finite sky coverage of the overlapping patch. This is commonly called the *Knox formula* (Knox, 1995).

The CMB reconstruction noise power spectrum is calculated from Eq. 1.71, while the galaxy lensing noise power – coming from the Poisson sampling of the shear field by galaxies of instrinsic ellipticity dispersion σ_e – is given by

$$N_\ell^{\kappa_g\kappa_g} = \frac{\sigma_e^2}{n_{\rm gal}}, \tag{5.16}$$

where $n_{\rm gal}$ is the mean galaxy surface density (i.e., source number per unit solid angle). We plot example signal and noise power spectra for the ACTPol and KiDS experiments in Fig. 5.18.

The forecast signal-to-noise ratio SNR on the detection of the cross-spectrum is

$$SNR = \sqrt{\sum_\ell \left(\frac{C_\ell^{\kappa_c\kappa_g}}{\sigma(\hat{C}_\ell^{\kappa_c\kappa_g})}\right)^2}, \tag{5.17}$$

where the denominator is given by Eq. 5.15, and all power spectra are evaluated in the fiducial cosmology (this expression is derived in Eq. 1.92). We summarise the forecast detection significance for several potential CMB lensing – galaxy lensing analyses in Table 5.1, with the assumed experimental specifications give in Table 5.2. The adopted redshift distributions for the DES, KiDS and LSST surveys are shown in Fig. 5.19, and the footprint of ACTPol, AdvACT and some relevant optical surveys are shown in Fig. 5.20.

The ACTPol×KiDS analysis is planned for 2016, and DES cross-correlations will soon be possible, both with the three season ACTPol data and also data from the AdvACT survey (beginning this year; Sec. 1.9). With its geographical position, ACT has access to parts of the sky which are well-populated with current and upcoming optical surveys; KiDS, DES and HSC all overlap with the ACT footprint. LSST is due for first light in 2021, and will run for 10 years; these synergies

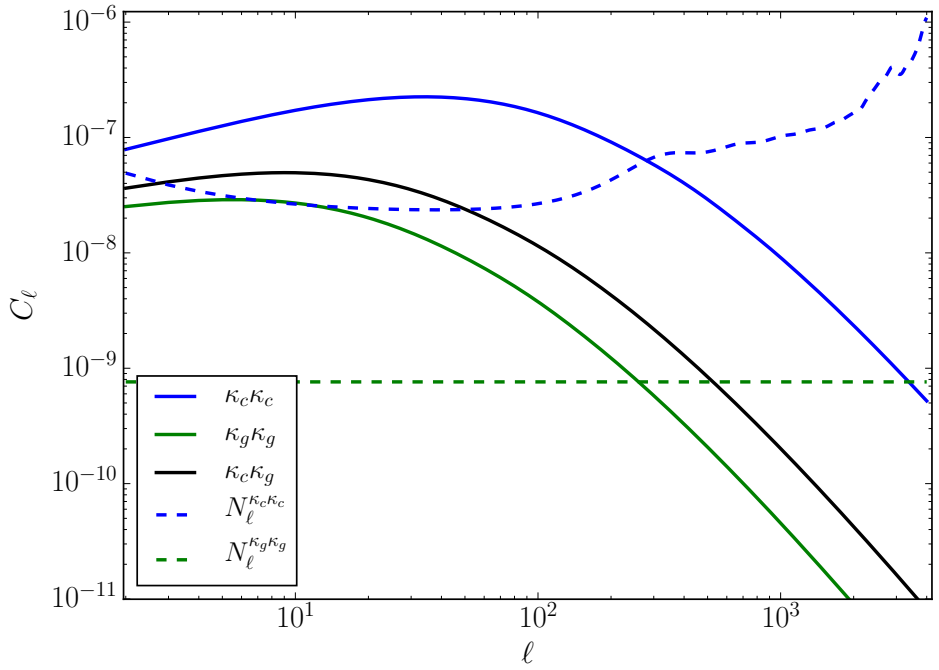

Figure 5.18: Signal (*solid lines*) and noise (*dashed lines*) spectra for ACTPol CMB lensing (*blue*), KiDS galaxy lensing (*green*) and their cross-correlation (*black*).

Data	$f_{\rm sky}$	$n_{\rm gal}$ [arcmin^{-2}]	S/N	Source
ACTPol × KiDS	0.02	10	13	This work
AdvACT × DES	0.10	10	32	This work
AdvACT × LSST	0.25	26	84	This work
AdvACT × LSST	0.40	26	108	This work
Planck × DES	0.10	10	19	Liu & Hill (2015)
Planck × HSC	0.05	20	16	Liu & Hill (2015)
Planck × Euclid	0.20	35	31	Liu & Hill (2015)
Planck × LSST	0.25	40	40	Liu & Hill (2015)

Table 5.1: Forecast signal-to-noise (SNR) on the CMB lensing – galaxy lensing cross-spectrum for different data combinations. Forecasts new to this work are calculated as in Eq. 5.17 with the CMB survey specifications given in Table 5.2. We use the source distributions of Lahav et al. (2010); Hand et al. (2015); LSST Science Collaboration et al. (2009) for DES, KiDS and LSST respectively (Fig. 5.19), and assume an intrinsic ellipticity dispersion $\sigma_e = 0.3$ for all forecasts.

CMB survey	ΔT [μK arcmin]	ΔP [μK arcmin]	Beam FWHM [arcmin]
ACTPol	11	16	1.4
AdvACT	7	10	1.4

Table 5.2: CMB survey specifications used for the forecasts presented in Table 5.1. We show the (assumed white-noise) map sensitivity in temperature T and polarisation P and the beam width at a central frequency $\nu = 150$ GHz (Henderson et al., 2015). The variance of the lensing estimator (Eq. 1.70) then gives the CMB lensing reconstruction noise spectrum $N_\ell^{\kappa\kappa}$.

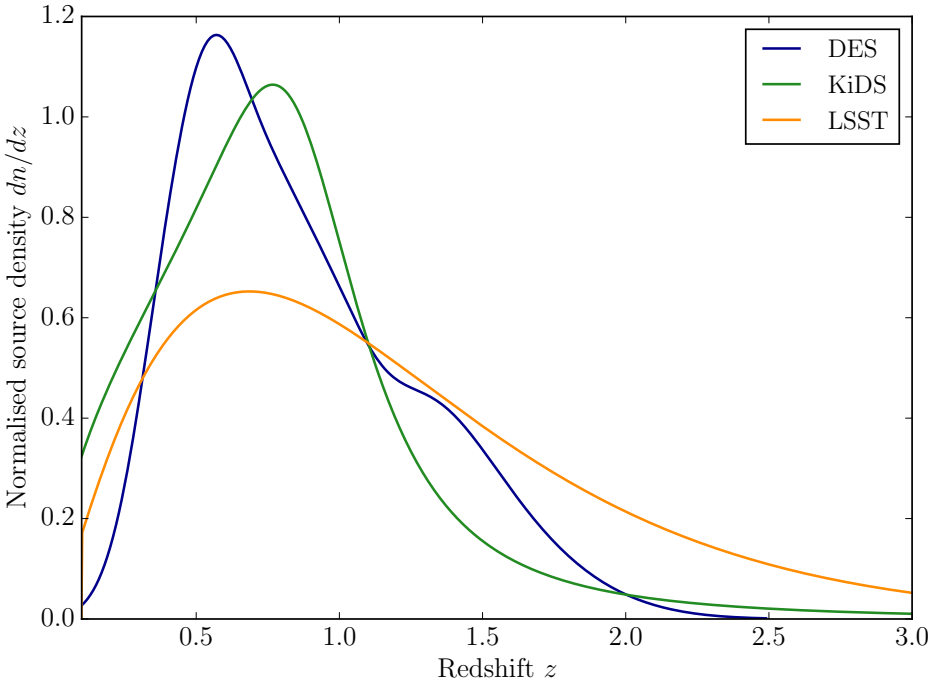

Figure 5.19: The source redshift distributions used for the cross-correlation forecasting analysis in this section, constructed from the expected distributions given in Lahav et al. (2010); Hand et al. (2015); LSST Science Collaboration et al. (2009). Each curve is normalised such that $\int (dn/dz) dz = 1$. The high-redshift tail of LSST boosts the cross-correlation signal above the shallower current surveys.

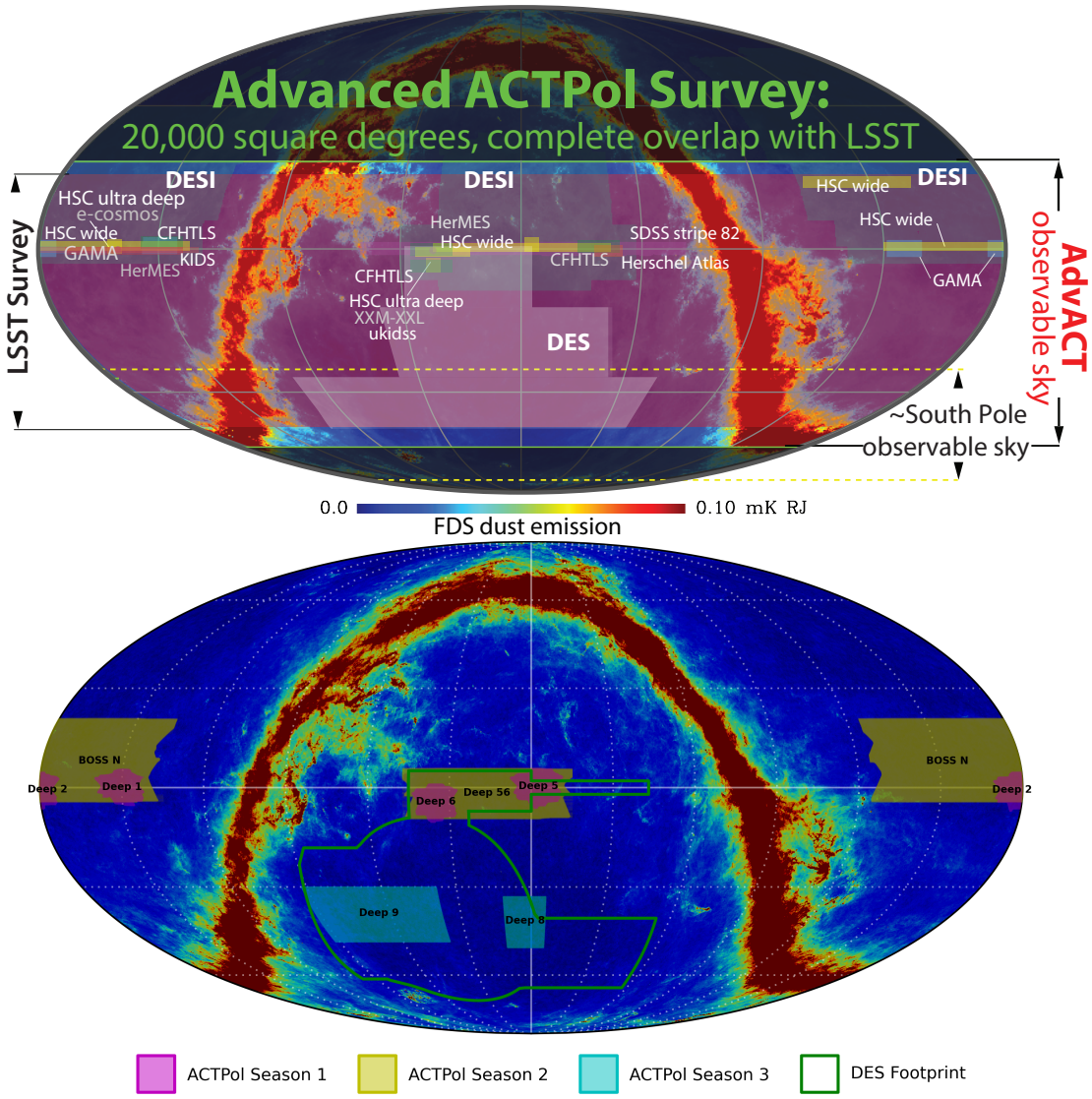

Figure 5.20: *Top panel*: Footprint of various optical surveys within the AdvACT observable sky. Ongoing surveys such as KiDS, HSC and DES will fully overlap with AdvACT, and already have partial overlap with the current three seasons of ACTPol coverage (*bottom panel*). LSST will cover the same 20,000 deg^2 as AdvACT; their lensing cross-correlation will be detected at $\approx 100\sigma$ (Table 5.2). These plots were made by the ACT collaboration.

are hence expected on a longer timescale. The South Pole Telescope (SPT) also has extensive overlap with DES, and the analysis of Kirk et al. (2015b) will likely be extended as improved optical weak-lensing data comes in and SPTpol and SPT-3G lensing data become available.

Given the significance, clearly a tomographic measurement of the signal will be possible, providing more cosmological information than an integrated measurement alone. The forecast detection significance figures should be compared to the current state-of-the-art measurement – the 4.2σ detection of Hand et al. (2015). CMB lensing can be thought of as an additional tomographic bin, with independent noise, with which the galaxy lensing redshift tomographic slices can be cross-correlated to probe the history of structure formation via $P(k, z)$. This emphasises the exciting science potential of this cosmological observable: as a systematic calibrator, and as provider of additional cosmological leverage both in ΛCDM and modified gravity models (Liu & Hill, 2015; Das et al., 2013; Kirk et al., 2015a). For instance, we argue in Chapter 4 that the combination of galaxy weak-lensing information with CMB and BAO data will improve constraints on the sum of the neutrino masses. This cross-correlation would naturally be included in a joint analysis of this kind, and would bring both direct cosmological information and indirectly improve precision of the neutrino mass determination by constraining the multiplicative bias.

We note that ground-based CMB experiments are limited at large scales not by detector shot noise (which gives an essentially white noise contribution on all but the largest scales) but by atmospheric variations and ground pickup (e.g., Naess et al., 2014). Typically, forecasts (including those in Fig. 5.7–5.10 and Table 5.1) assume that the small-scale noise spectrum extrapolates to all scales, ignoring these other contributions. How the additional noise variance at large scales and foreground components will affect future ground-based lensing measurements is an active area of research.

5.6 Conclusions

CMB lensing measurements naturally complement galaxy lensing data. A careful handling of systematic biases will be required for future analyses involving parameter extraction. Combination of the two datasets will improve cosmological constraints beyond auto-spectra alone, and allow for 'self-calibration' of weak-lensing systematics such as multiplicative bias. With current data,

there are strong degeneracies in the cross-spectrum shape between sources from different redshift; upcoming larger and more sensitive surveys will allow us to break these degeneracies. The covariance of the two dark matter probes, which probe common structure, must be accounted for in any likelihood analysis.

A comprehensive joint-analysis of these cosmological probes would allow for the optimal extraction of cosmological information in the presence of systematic biases and uncertainties in the redshift distribution and cosmological model. The dominant systematic effects could then be modelled, quantified and marginalised out. Experiments such as ACT, SPT, DES, LSST and Euclid are expected to perform just such an analysis, combining the covariant CMB lensing, cosmic shear and galaxy clustering data into a joint likelihood.

Chapter 6

Conclusions

During the course of this D.Phil. a major breakthrough for the field of cosmology was made with the release and first analysis of the *Planck* satellite mission data. This has provided the most stringent test to date of the ΛCDM cosmological model, confirming a deviation from scale invariance of the primordial power spectrum and providing a measurement of the dark energy equation-of-state w consistent with a simple cosmological constant. The observed lensing of the cosmic microwave background (CMB) by intervening matter fluctuations is in agreement with ΛCDM predictions given the *Planck* temperature and polarisation power spectra, and is detected at a staggering 40σ significance. This consistency of early- and late-time cosmological probes is a remarkable feature of our universe, increasing our confidence in the current cosmological paradigm. In combination with BICEP-Keck data, there is now a tight upper bound on the primordial gravitational wave amplitude ($r < 0.07$ at 95% confidence), yet searches for the tensor-to-scalar ratio continue in earnest.

The use of gravitational lensing to measure the dark universe has advanced rapidly – both observationally and data analytically – in recent years. CMB and galaxy lensing observations are now being used routinely for precision cosmology. This trend will only continue over the next 5–10 years given the variety of ongoing and upcoming lensing surveys such as DES, KiDS, HSC, LSST, Euclid, AdvACT, SPT-3G and CMB S4. On the data analysis side, improved techniques for the extraction of the lensing signal are being developed, such as those which go beyond simple quadratic estimator techniques for CMB lensing, or those which involve the sampling of the full posterior of the cosmic shear field (e.g., Anderes et al., 2015; Alsing et al., 2016; Heavens et al.,

2016). These considerations suggest an exciting future for lensing measurements.

With a view to analysing these data, we have quantified and described various sampling techniques which can be used, quite generically, for parameter estimation, applying them to toy-model problems and the particular case of parameter inference with the *WMAP* 7 CMB likelihood. We found that nested sampling delivers high-fidelity estimates for posterior statistics at low computational cost, and should be adopted in favour of Metropolis-Hastings in many cases. Affine-invariant MCMC is competitive when computer clusters can be utilised for massive parallelisation. Affine-invariant MCMC and existing extensions to nested sampling naturally probe multi-modal and curving distributions.

We have demonstrated the power of cross-correlations to constrain astrophysical effects by correlating the positions of radio galaxies in the FIRST survey with the CMB lensing convergence estimated from the Atacama Cosmology Telescope. This allowed us to determine the bias of these galaxies, which were preferentially selected as active galactic nuclei (AGN). We measured the angular cross-spectrum between these two dark matter probes. Modelling the AGN population with a redshift-dependent bias, the cross-spectrum is well fit by the *Planck* best-fit ΛCDM cosmological model. This measurement characterises the typical halo mass of radio-loud AGN, which we found to be particularly large ($\approx 4 \times 10^{13} M_{\odot}$); these are highly biased tracers of the underlying matter field.

We have shown in detail the prospects for CMB and BAO data to determine the sum of the neutrinos masses to high precision, with a detection possible with next generation instruments. Much of the direct constraining power of these data comes from the sensitivity of the CMB lensing signal to the neutrino mass; a higher neutrino mass will lead to a suppression of clustering on small scales relative to large scales. This propagates into the shape of the lensing angular power spectrum.

Finally, we presented a forecasting and analysis pipeline for weak-lensing cross-correlations. We introduced and motivated the joint analysis of CMB lensing and galaxy lensing, and argued that simultaneous analysis with CMB lensing data provides a robust approach to calibration of systematic effects in weak-lensing analyses. We showed forecasts of the ability of future experimental combinations to detect the lensing cross-correlation.

The work presented in this thesis naturally suggests several future lines of research. We finish

with a partial list of such topics, including a brief description of the problem to be investigated and the expected scientific outcomes.

Cosmological analysis with cross-correlations. To date, lensing cross-correlations have been presented in 'detection' papers, and have yet to be utilised in a full cosmological analysis (e.g., Sherwin et al., 2012; *Planck* Collaboration, 2013b, 2014b; Allison et al., 2015; Kirk et al., 2015a). This will become possible with ongoing and upcoming surveys, allowing for the calibration of multiplicative bias and data-consistency tests. This will be particularly interesting given the current tension between *Planck*- and CFHTLens-derived growth parameters, which itself hints at residual systematic effects or, perhaps, modified gravity. A rigorous combined analysis of CMB and galaxy lensing would clarify this issue.

CMB lensing: delensing and consistency. The lensing potential inferred from temperature and polarisation measurements, using the methods described in this thesis, can be used to *delens* the CMB. This would reduce the lensing contribution to the B-mode power spectrum, allowing for a more precise determination of the tensor-to-scalar ratio r. An important cross-check to make in the near future would be to quantify the consistency between lensing measurements derived from different experiments, e.g., *Planck* and ACTPol. Measurements can be compared at the map and power spectrum level; any discrepancy might suggest the presence of systematic effects or foreground contamination in the lensing estimation. The ground-based experiments are expected to improve the small-scale lensing measurements beyond *Planck*. Furthermore, there is a hint of a 'low' bandpower around $\ell = 700$ in the *Planck*-derived lensing power spectrum; if this persists in ground-based determinations, it leaves open the possibility for detection of the sum of the neutrino masses (*Planck* Collaboration, 2015e).

Shape estimation. Current analyses of galaxy shear surveys derive a point estimate (and corresponding weight) for the ellipticity parameters of individual galaxies identified photometrically in the survey. The methodological approach can lead to a *noise bias* for low signal-to-noise galaxies that will propagate into the estimation of the shear field and, ultimately, cosmological parameters. The development of new techniques to mitigate this bias (e.g., sampling from the full posterior of the shape parameters when inferring the shear field)

will be necessary for upcoming surveys.

Comprehensive forecasting. Going beyond just CMB and BAO data, the sum of the neutrino masses and other ΛCDM extensions such as dark energy and curvature can be more tightly constrained by combination with low-redshift probes of structure formation such as galaxy lensing and clustering. An interesting research project would be to develop a forecasting pipeline which combines all these probes for upcoming surveys (as will be done with the real data when in hand), fully accounting for their covariance.

Model uncertainty. Within the Bayesian framework one can rigorously answer the question of whether the data prefer one cosmological model or another (e.g., ΛCDM versus wCDM.) With suitably chosen priors, one can use Bayesian *model averaging* to fully account for *a priori* model uncertainty. This may be useful in the detection of neutrino mass, where curvature or dark energy can partially mimic its effects on the observables, or in averaging constraints over the normal and inverted hierarchies.

Covariance of multiple probes. With the proliferation of high-quality, wide-field and overlapping data from multiple dark matter probes, a robust combination of datasets in a joint likelihood must necessarily model the *covariance* of the different probes and marginalise over systematic effects for robust cosmological inference. Essentially one must account for the fact that common dark matter modes lens the CMB, lens galaxies, and also host luminous matter (determining clustering); this covariance will be cosmology dependent. The inclusion of cross-correlations, e.g., to calibrate tracer bias and multiplicative bias and to bring additional cosmological information, would be an obvious step in this process.

Inference of the matter field of the universe. A long-term, ambitious goal would be to use joint information from multiple probes to (probabilistically) infer the underlying matter field, essentially creating a three-dimensional map (or set of *plausible* maps) for the large-scale structure in the late-time observable universe (e.g., Leclercq, 2015). This approach would naturally require the assumption of a given cosmological model, such as ΛCDM, and a way to evolve the fluctuations in time, e.g., perturbation theory or – preferably – rapid forward-modelling through numerical simulations to fully capture non-linear evolution on small scales.

Bibliography

Abazajian, K. N., Arnold, K., Austermann, J., Benson, B. A., Bischoff, C. et al., 2013. Neutrino Physics from the Cosmic Microwave Background and Large Scale Structure. *ArXiv e-prints*.

Abbott, B. P., Abbott, R., Abbott, T. D., Abernathy, M. R., Acernese, F. et al., 2016. Observation of Gravitational Waves from a Binary Black Hole Merger. *Phys. Rev. Lett.*, **116**, 061102. URL `http://link.aps.org/doi/10.1103/PhysRevLett.116.061102`.

Abdalla, F. B., Bull, P., Camera, S., Benoit-Lévy, A., Joachimi, B., Kirk, D., Klöckner, H.-R., Maartens, R., Raccanelli, A., Santos, M. G., Zhao, G.-B. & Cosmology SWG, f. t., 2015. Cosmology from HI galaxy surveys with the SKA. *ArXiv e-prints, arXiv:1501.04035*.

Abdalla, F. B. & Rawlings, S., 2007. Determining neutrino properties using future galaxy redshift surveys. *Monthly Notices of the Royal Astronomical Society*, **381**, 1313–1328.

Akeret, J., Seehars, S., Amara, A., Refregier, A. & Csillaghy, A., 2013. CosmoHammer: Cosmological parameter estimation with the MCMC Hammer. Astrophysics Source Code Library.

Allison, R., Caucal, P., Calabrese, E., Dunkley, J. & Louis, T., 2015. Towards a cosmological neutrino mass detection. *Phys. Rev. D*, **92**, 123535. URL `http://link.aps.org/doi/10.1103/PhysRevD.92.123535`.

Allison, R. & Dunkley, J., 2014. Comparison of sampling techniques for Bayesian parameter estimation. *MNRAS*, **437**, 3918–3928.

Allison, R., Lindsay, S. N., Sherwin, B. D., de Bernardis, F., Bond, J. R. et al., 2015. The Atacama Cosmology Telescope: measuring radio galaxy bias through cross-correlation with lensing. *MNRAS*, **451**, 849–858.

Alsing, J., Heavens, A., Jaffe, A. H., Kiessling, A., Wandelt, B. & Hoffmann, T., 2016. Hierarchical cosmic shear power spectrum inference. *Monthly Notices of the Royal Astronomical Society*, **455**(4), 4452–4466.

Amendola, L., Appleby, S., Bacon, D., Baker, T., Baldi, M. et al., 2013. Cosmology and Fundamental Physics with the Euclid Satellite. *Living Reviews in Relativity*, **16**, 6.

Anderes, E., Wandelt, B. D. & Lavaux, G., 2015. Bayesian Inference of CMB Gravitational Lensing. *ApJ*, **808**, 152.

Anderson, L., Aubourg, É., Bailey, S., Beutler, F., Bhardwaj, V. et al., 2014. The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: baryon acoustic oscillations in the Data Releases 10 and 11 Galaxy samples. *Monthly Notices of the Royal Astronomical Society*, **441**, 24–62.

Anderson, L., Aubourg, E., Bailey, S., Bizyaev, D., Blanton, M. et al., 2012. The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: baryon acoustic oscillations in the Data Release 9 spectroscopic galaxy sample. *MNRAS*, **427**, 3435–3467.

Array, K., BICEP2 Collaborations, :, Ade, P. A. R., Ahmed, Z. et al., 2015. BICEP2 / Keck Array VI: Improved Constraints On Cosmology and Foregrounds When Adding 95 GHz Data From Keck Array. *ArXiv e-prints*.

Baldi, P. & Marinucci, D., 2006. Some characterizations of the spherical harmonics coefficients for isotropic random fields. *ArXiv Mathematics e-prints*.

Bardeen, J. M., Bond, J. R., Kaiser, N. & Szalay, A. S., 1986. The statistics of peaks of Gaussian random fields. *Ap. J.*, **304**, 15–61.

Barrow, J. D. & Liddle, A. R., 1997. Can Inflation be Falsified? *General Relativity and Gravitation*, **29**, 1503–1510.

Bartelmann, M. & Schneider, P., 2001. Weak gravitational lensing. *Physics Reports*, **340**, 291–472.

Bassett, B. A., Corasaniti, P. S. & Kunz, M., 2004. The Essence of Quintessence and the Cost of Compression. *The Astrophysical Journal*, **617**, L1–L4.

Baumann, D., 2009. TASI Lectures on Inflation. *ArXiv e-prints*.

Baumann, D., 2015. Cosmology. URL `http://damtp.cam.ac.uk/user/db275/Cosmology/Lectures.pdf`.

Baxter, E. J., Keisler, R., Dodelson, S., Aird, K. A., Allen, S. W. et al., 2015. A Measurement of Gravitational Lensing of the Cosmic Microwave Background by Galaxy Clusters Using Data from the South Pole Telescope. *The Astrophysical Journal*, **806**, 247.

Bayes, T., 1958. Studies in the History of Probability and Statistics: IX. Thomas Bayes's Essay Towards Solving a Problem in the Doctrine of Chances. *Biometrika*, **45**, 296–315.

Becker, M. R., Troxel, M. A., MacCrann, N., Krause, E., Eifler, T. F. et al., 2015. Cosmic Shear Measurements with DES Science Verification Data. *ArXiv e-prints*.

Becker, R. H., White, R. L. & Helfand, D. J., 1995. The FIRST Survey: Faint Images of the Radio Sky at Twenty Centimeters. *The Astrophysical Journal*, **450**, 559.

Benoit-Lévy, A., Smith, K. M. & Hu, W., 2012. Non-Gaussian structure of the lensed CMB power spectra covariance matrix. *Physical Review D*, **86**(12), 123008.

Benson, B. A., Ade, P. A. R., Ahmed, Z., Allen, S. W., Arnold, K. et al., 2014. SPT-3G: a next-generation cosmic microwave background polarization experiment on the South Pole telescope. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, vol. 9153 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, 1.

Beutler, F., Blake, C., Colless, M., Jones, D. H., Staveley-Smith, L., Campbell, L., Parker, Q., Saunders, W. & Watson, F., 2011. The 6dF Galaxy Survey: baryon acoustic oscillations and the local Hubble constant. *MNRAS*, **416**, 3017–3032.

Bianchini, F., Bielewicz, P., Lapi, A., Gonzalez-Nuevo, J., Baccigalupi, C. et al., 2014. Cross-correlation between the CMB lensing potential measured by Planck and high-z sub-mm galaxies detected by the Herschel-ATLAS survey. *ArXiv e-prints, arXiv:1410.4502*.

BICEP2 Collaboration, Ade, P. A. R., Aikin, R. W., Amiri, M., Barkats, D. et al., 2014. BICEP2. II. Experiment and three-year Data Set. *ApJ*, **792**, 62.

BICEP2/Keck and *Planck* Collaborations, 2015. Joint Analysis of BICEP2/Keck Array and *Planck* Data. *Physical Review Letters*, **114**(10), 101301.

Bird, S., Viel, M. & Haehnelt, M. G., 2012. Massive neutrinos and the non-linear matter power spectrum. *MNRAS*, **420**, 2551–2561.

Blake, C., Mauch, T. & Sadler, E. M., 2004a. Angular clustering in the Sydney University Molonglo Sky Survey. *MNRAS*, **347**, 787–794.

Blake, C. & Wall, J., 2002. Measurement of the angular correlation function of radio galaxies from the NRAO VLA Sky Survey. *MNRAS*, **329**, L37–L41.

Blake, C. A., Abdalla, F. B., Bridle, S. L. & Rawlings, S., 2004b. Cosmology with the SKA. *NAR*, **48**, 1063–1077.

Bleem, L. E., van Engelen, A., Holder, G. P., Aird, K. A., Armstrong, R. et al., 2012. A Measurement of the Correlation of Galaxy Surveys with CMB Lensing Convergence Maps from the South Pole Telescope. *The Astrophysical Journal Letters*, **753**, L9.

Bock, D. C.-J., Large, M. I. & Sadler, E. M., 1999. SUMSS: A Wide-Field Radio Imaging Survey of the Southern Sky. I. Science Goals, Survey Design, and Instrumentation. *AJ*, **117**, 1578–1593.

Bond, J. R., Efstathiou, G. & Tegmark, M., 1997. Forecasting cosmic parameter errors from microwave background anisotropy experiments. *MNRAS*, **291**, L33–L41.

Bonnett, C., Troxel, M. A., Hartley, W., Amara, A., Leistedt, B. et al., 2015. Redshift distributions of galaxies in the DES Science Verification shear catalogue and implications for weak lensing. *ArXiv e-prints*.

Bower, R. G., Benson, A. J., Malbon, R., Helly, J. C., Frenk, C. S., Baugh, C. M., Cole, S. & Lacey, C. G., 2006. Breaking the hierarchy of galaxy formation. *Monthly Notices of the Royal Astronomical Society*, **370**, 645–655.

Brandenberger, R. H., 2011. Alternatives to the Inflationary Paradigm of Structure Formation. *International Journal of Modern Physics Conference Series*, **1**, 67–79.

Burgess, C. P., 2013. The Cosmological Constant Problem: Why it's hard to get Dark Energy from Micro-physics. *ArXiv e-prints*.

Calabrese, E., de Putter, R., Huterer, D., Linder, E. V. & Melchiorri, A., 2011. Future CMB constraints on early, cold, or stressed dark energy. *Physical Review D*, **83**(2), 023011.

Calabrese, E., Hložek, R., Battaglia, N., Bond, J. R., de Bernardis, F. et al., 2014. Precision epoch of reionization studies with next-generation CMB experiments. *JCAP*, **8**, 10.

Calabrese, E., Slosar, A., Melchiorri, A., Smoot, G. F. & Zahn, O., 2008. Cosmic microwave weak lensing data as a test for the dark universe. *Physical Review D*, **77**(12), 123531.

Camera, S., Santos, M. G., Bacon, D. J., Jarvis, M. J., McAlpine, K., Norris, R. P., Raccanelli, A. & Röttgering, H., 2012. Impact of redshift information on cosmological applications with next-generation radio surveys. *MNRAS*, **427**, 2079–2088.

Carbone, C., Fedeli, C., Moscardini, L. & Cimatti, A., 2012. Measuring the neutrino mass from future wide galaxy cluster catalogues. *JCAP*, **3**, 23.

Carilli, C. L. & Rawlings, S., 2004. Motivation, key science projects, standards and assumptions. *New Astron. Rev.*, **48**, 979–984.

Challinor, A., 2009. Part III - Cosmology. URL `http://cosmologist.info/teaching/EU/ADC_Structure_formation2.pdf`.

Chevallier, M. & Polarski, D., 2001. Accelerating Universes with Scaling Dark Matter. *International Journal of Modern Physics D*, **10**, 213–223.

Chisari, N. E., Dunkley, J., Miller, L. & Allison, R., 2015. Contamination of early-type galaxy alignments to galaxy lensing-CMB lensing cross-correlation. *MNRAS*, **453**, 682–689.

Chopin, N. & Robert, C., 2010. Properties of Nested Sampling. *Biometrika*.

Clifton, T., Ferreira, P. G., Padilla, A. & Skordis, C., 2012. Modified gravity and cosmology. *Phys. Rep.*, **513**, 1–189.

Comparat, J., Delubac, T., Jouvel, S., Raichoor, A., Kneib, J. et al., 2015. The SDSS-IV eBOSS emission-line galaxy pilot survey. *ArXiv e-prints*.

Condon, J. J., Cotton, W. D. & Broderick, J. J., 2002. Radio Sources and Star Formation in the Local Universe. *The Astronomical Journal*, **124**, 675–689.

Condon, J. J., Cotton, W. D., Greisen, E. W., Yin, Q. F., Perley, R. A., Taylor, G. B. & Broderick, J. J., 1998. The NRAO VLA Sky Survey. *AJ*, **115**, 1693–1716.

Costanzi Alunno Cerbolini, M., Sartoris, B., Xia, J.-Q., Biviano, A., Borgani, S. & Viel, M., 2013. Constraining neutrino properties with a Euclid-like galaxy cluster survey. *JCAP*, **6**, 20.

Cox, R. T., 1946. Probability, Frequency and Reasonable Expectation. *American Journal of Physics*, **14**(1).

Cress, C. M., Helfand, D. J., Becker, R. H., Gregg, M. D. & White, R. L., 1996. The Angular Two-Point Correlation Function for the FIRST Radio Survey. *Ap J*, **473**, 7.

Croton, D. J., Springel, V., White, S. D. M., De Lucia, G., Frenk, C. S., Gao, L., Jenkins, A., Kauffmann, G., Navarro, J. F. & Yoshida, N., 2006. The many lives of active galactic nuclei: cooling flows, black holes and the luminosities and colours of galaxies. *Monthly Notices of the Royal Astronomical Society*, **365**, 11–28.

Das, S., Louis, T., Nolta, M. R., Addison, G. E., Battistelli, E. S. et al., 2013. The Atacama Cosmology Telescope: Temperature and Gravitational Lensing Power Spectrum Measurements from Three Seasons of Data. *ArXiv e-prints*.

Das, S., Louis, T., Nolta, M. R., Addison, G. E., Battistelli, E. S. et al., 2014. The Atacama Cosmology Telescope: temperature and gravitational lensing power spectrum measurements from three seasons of data. *JCAP*, **4**, 14.

Das, S., Sherwin, B. D., Aguirre, P., Appel, J. W., Bond, J. R. et al., 2011. Detection of the Power Spectrum of Cosmic Microwave Background Lensing by the Atacama Cosmology Telescope. *Physical Review Letters*, **107**(2), 021301.

de Putter, R., Zahn, O. & Linder, E. V., 2009. CMB lensing constraints on neutrinos and dark energy. *Physical Review D*, **79**(6), 065033.

de Zotti, G., Massardi, M., Negrello, M. & Wall, J., 2010. Radio and millimeter continuum surveys and their astrophysical implications. *aapr*, **18**, 1–65.

Diehl, H. T., Abbott, T. M. C., Annis, J., Armstrong, R., Baruah, L. et al., 2014. The Dark Energy Survey and operations: Year 1. In *Observatory Operations: Strategies, Processes, and Systems V*, vol. 9149 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, 91490V.

Dodelson, S., 2003. *Modern cosmology*.

Dolgov, A. D., 2002. Neutrinos in cosmology. *Physics Reports*, **370**, 333–535.

Doran, M. & Robbers, G., 2006. Early dark energy cosmologies. *JCAP*, **6**, 26.

Doroshkevich, A. G., Khlopov, M. I., Sunyaev, R. A., Szalay, A. S. & Zeldovich, I. B., 1981. Cosmological impact of the neutrino rest mass. *Annals of the New York Academy of Sciences*, **375**, 32–42.

Doroshkevich, A. G., Zeldovich, Y. B., Syunyaev, R. A. & Khlopov, M. Y., 1980. Astrophysical implications of the neutrino rest mass. II - The density-perturbation spectrum and small-scale fluctuations in the microwave background. III - Nonlinear growth of perturbations and the missing mass. *Pisma v Astronomicheskii Zhurnal*, **6**, 457–469.

Dubois, Y., Gavazzi, R., Peirani, S. & Silk, J., 2013. AGN-driven quenching of star formation: morphological and dynamical implications for early-type galaxies. *Monthly Notices of the Royal Astronomical Society*, **433**, 3297–3313.

Dunkley, J., Bucher, M., Ferreira, P. G., Moodley, K. & Skordis, C., 2005. Fast and reliable Markov chain Monte Carlo technique for cosmological parameter estimation. *MNRAS*, **356**, 925–936.

Dunkley, J., Calabrese, E., Sievers, J., Addison, G. E., Battaglia, N. et al., 2013. The Atacama Cosmology Telescope: likelihood for small-scale CMB data. *JCAP*, **7**, 25.

Eisenstein, D. J., Weinberg, D. H., Agol, E., Aihara, H., Allende Prieto, C., Anderson, S. F., Arns, J. A., Aubourg, É., Bailey, S., Balbinot, E. & et al., 2011. SDSS-III: Massive Spectroscopic Surveys of the Distant Universe, the Milky Way, and Extra-Solar Planetary Systems. *The Astrophysical Journal*, **142**, 72.

Eisenstein, D. J. et al., 2005. Detection of the baryon acoustic peak in the large-scale correlation function of SDSS luminous red galaxies. *Astrophys. J.*, **633**, 560–574.

Elgarøy, Ø. & Lahav, O., 2005. Neutrino masses from cosmological probes. *New Journal of Physics*, **7**, 61.

Erben, T., Hildebrandt, H., Miller, L., van Waerbeke, L., Heymans, C. et al., 2013. CFHTLenS: the Canada-France-Hawaii Telescope Lensing Survey - imaging data and catalogue products. *MNRAS*, **433**, 2545–2563.

Errard, J., Feeney, S. M., Peiris, H. V. & Jaffe, A. H., 2015. Robust forecasts on fundamental physics from the foreground-obscured, gravitationally-lensed CMB polarization. *ArXiv e-prints*.

Essinger-Hileman, T., Ali, A., Amiri, M., Appel, J. W., Araujo, D. et al., 2014. CLASS: the cosmology large angular scale surveyor. In *Millimeter, Submillimeter, and Far-Infrared Detectors and Instrumentation for Astronomy VII*, vol. 9153 of *Proc. SPIE*, 91531I.

Fang, W., Hu, W. & Lewis, A., 2008. Crossing the phantom divide with parametrized post-Friedmann dark energy. *Physical Review D*, **78**(8), 087303.

Feldman, G. J., Hartnell, J. & Kobayashi, T., 2012. A Review of Long-baseline Neutrino Oscillation Experiments. *ArXiv e-prints*.

Feng, C., Aslanyan, G., Manohar, A. V., Keating, B., Paar, H. P. & Zahn, O., 2012. Measuring gravitational lensing of the cosmic microwave background using cross correlation with large scale structure. *Physical Review D*, **86**(6), 063519.

Fernandes, C. A. C., Jarvis, M. J., Martínez-Sansigre, A., Rawlings, S., Afonso, J., Hardcastle, M. J., Lacy, M., Stevens, J. A. & Vardoulaki, E., 2015. Black hole masses, accretion rates and hot- and cold-mode accretion in radio galaxies at z $\sim$ 1. *Monthly Notices of the Royal Astronomical Society*, **447**, 1184–1203.

Feroz, F. & Hobson, M. P., 2008. Multimodal nested sampling: an efficient and robust alternative to Markov Chain Monte Carlo methods for astronomical data analyses. *MNRAS*, **384**, 449–463.

Feroz, F., Hobson, M. P. & Bridges, M., 2009. MULTINEST: an efficient and robust Bayesian inference tool for cosmology and particle physics. *MNRAS*, **398**, 1601–1614.

Feroz, F., Hobson, M. P., Cameron, E. & Pettitt, A. N., 2013. Importance Nested Sampling and the MultiNest Algorithm. *ArXiv e-prints*.

Ferramacho, L. D., Santos, M. G., Jarvis, M. J. & Camera, S., 2014. Radio galaxy populations and the multitracer technique: pushing the limits on primordial non-Gaussianity. *Monthly Notices of the Royal Astronomical Society*, **442**, 2511–2518.

Ferroni, F., Vissani, F., Brofferio, C. & italiana di fisica, S., 2009. *Measurements of Neutrino Mass*. International School of Physics Enrico Fermi. IOS Press. ISBN 9781607500384. URL `https://books.google.co.uk/books?id=5kdxFAOX0i0C`.

Font-Ribera, A., McDonald, P., Mostek, N., Reid, B. A., Seo, H.-J. & Slosar, A., 2014. DESI and other Dark Energy experiments in the era of neutrino mass measurements. *JCAP*, **5**, 23.

Foreman-Mackey, D., Hogg, D. W., Lang, D. & Goodman, J., 2013. emcee: The MCMC Hammer. *Publications of the Astronomical Society of the Pacific*, **125**, 306–312.

Fornengo, N., Perotto, L., Regis, M. & Camera, S., 2014. Evidence of cross-correlation between the CMB lensing and the gamma-ray sky. *ArXiv e-prints, arXiv:1410.4997*.

Fraisse, A. A., Ade, P. A. R., Amiri, M., Benton, S. J., Bock, J. J. et al., 2013. SPIDER: probing the early Universe with a suborbital polarimeter. *JCAP*, **4**, 047.

Fukugita, M., 2006. Massive neutrinos in cosmology. *Nucl. Phys. Proc. Suppl.*, **155**, 10–17.

Geach, J. E., Hickox, R. C., Bleem, L. E., Brodwin, M., Holder, G. P. et al., 2013. A Direct Measurement of the Linear Bias of Mid-infrared-selected Quasars at z $\sim$ 1 Using Cosmic Microwave Background Lensing. *The Astrophysical Journal Letters*, **776**, L41.

Gelman, A., Roberts, G. & Gilks, W., 1996. Efficient Metropolis jumping rules. In J. M. Bernado et al., eds., *Bayesian Statistics*, vol. 5, 599. OUP.

Gelman, A. & Rubin, D., 1992. Inference from iterative simulation using multiple sequences. *Statistical Science*, **7**, 457–511. http://www.stat.columbia.edu/ gelman/research/published/itsim.pdf.

Giannantonio, T., Fosalba, P., Cawthon, R., Omori, Y., Crocce, M. et al., 2015. CMB lensing tomography with the DES Science Verification galaxies. *ArXiv e-prints*.

Giannantonio, T., Fosalba, P., Cawthon, R., Omori, Y., Crocce, M. et al., 2016. CMB lensing tomography with the DES Science Verification galaxies. *MNRAS*, **456**, 3213–3244.

Gilks, W. R., Richardson, S. & Spiegelhalter, D. J., 1995. *Markov Chain Monte Carlo in Practice*. Chapman & Hall, London.

Gill, R. D. & Levit, B. Y., 1995. Applications of the van Trees inequality: a Bayesian CramŐr-Rao bound. *Bernoulli*, **1**(1-2), 59–79. URL `http://projecteuclid.org/euclid.bj/1186078362`.

Giunti, C. & Kim, C. W., 2007. *Fundamentals of Neutrino Physics and Astrophysics*. Oxford University Press, Oxford, UK.

Gonzalez-Garcia, M. C. & Nir, Y., 2003. Neutrino masses and mixing: evidence and implications. *Reviews of Modern Physics*, **75**, 345–402.

Goodman, J. & Weare, J., 2010. Ensemble samplers with affine invariance. *Comm. App. Math. and Comp. Sci.*, **5**, 65–80.

Griffiths, L. M., Barbosa, D. & Liddle, A. R., 1999. Cosmic microwave background constraints on the epoch of reionization. *MNRAS*, **308**, 854–862.

Gubitosi, G., Lagos, M., Magueijo, J. & Allison, R., 2015. Inflation, evidence and falsifiability. *ArXiv e-prints*.

Hall, A. & Taylor, A., 2014. Intrinsic alignments in the cross-correlation of cosmic shear and cosmic microwave background weak lensing. *MNRAS*, **443**, L119–L123.

Hall, A. C. & Challinor, A., 2012. Probing the neutrino mass hierarchy with cosmic microwave background weak lensing. *MNRAS*, **425**, 1170–1184.

Hamann, J., Hannestad, S. & Wong, Y. Y. Y., 2012. Measuring neutrino masses with a future galaxy survey. *JCAP*, **11**, 52.

Hamilton, A. J. S., 2001. Formulae for growth factors in expanding universes containing matter and a cosmological constant. *MNRAS*, **322**, 419–425.

Hand, N., Leauthaud, A., Das, S., Sherwin, B. D., Addison, G. E. et al., 2015. First measurement of the cross-correlation of CMB lensing and galaxy lensing. *Phys. Rev. D*, **91**, 062001. URL `http://link.aps.org/doi/10.1103/PhysRevD.91.062001`.

Handley, W. J., Hobson, M. P. & Lasenby, A. N., 2015. POLYCHORD: nested sampling for cosmology. *MNRAS*, **450**, L61–L65.

Hannestad, S., 2005. Neutrino Masses and the Dark Energy Equation of State:Relaxing the Cosmological Neutrino Mass Bound. *Physical Review Letters*, **95**(22), 221301.

Hannestad, S., 2006. Primordial Neutrinos. *Annual Review of Nuclear and Particle Science*, **56**, 137–161.

Hannestad, S., 2010. Neutrino physics from precision cosmology. *Progress in Particle and Nuclear Physics*, **65**, 185–208.

Hatch, N. A., Wylezalek, D., Kurk, J. D., Stern, D., De Breuck, C., Jarvis, M. J., Galametz, A., Gonzalez, A. H., Hartley, W. G., Mortlock, A., Seymour, N. & Stevens, J. A., 2014. Why z $>$ 1 radio-loud galaxies are commonly located in protoclusters. *Monthly Notices of the Royal Astronomical Society*, **445**, 280–289.

Heath, D. J., 1977. The growth of density perturbations in zero pressure Friedmann-Lemaitre universes. *MNRAS*, **179**, 351–358.

Heavens, A., Alsing, J., Jaffe, A., Hoffmann, T., Kiessling, A. & Wandelt, B., 2016. Bayesian hierarchical modelling of weak lensing - the golden goal. *ArXiv e-prints*.

Helfand, D. J., White, R. L. & Becker, R. H., 2015. The Last of FIRST: The Final Catalog and Source Identifications. *ArXiv e-prints, arXiv:1501.01555*.

Henderson, S. W., Allison, R., Austermann, J., Baildon, T., Battaglia, N. et al., 2015. Advanced ACTPol Cryogenic Detector Arrays and Readout. *ArXiv e-prints*.

Hennig, P., Osborne, M. A. & Girolami, M., 2015. Probabilistic numerics and uncertainty in computations. **471**(2179).

Heymans, C., Grocutt, E., Heavens, A., Kilbinger, M., Kitching, T. D. et al., 2013. CFHTLenS tomographic weak lensing cosmological parameter constraints: Mitigating the impact of intrinsic galaxy alignments. *MNRAS*, **432**, 2433–2453.

Heymans, C., Van Waerbeke, L., Bacon, D., Berge, J., Bernstein, G. et al., 2006. The Shear Testing Programme - I. Weak lensing analysis of simulated ground-based observations. *MNRAS*, **368**, 1323–1339.

Heymans, C., Van Waerbeke, L., Miller, L., Erben, T., Hildebrandt, H. et al., 2012. CFHTLenS: the Canada-France-Hawaii Telescope Lensing Survey. *MNRAS*, **427**, 146–166.

Hinshaw, G., Larson, D., Komatsu, E., Spergel, D. N., Bennett, C. L. et al., 2013. Nine-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Parameter Results. *ApJs*, **208**, 19.

Hirata, C. M., Ho, S., Padmanabhan, N., Seljak, U. & Bahcall, N. A., 2008. Correlation of CMB with large-scale structure. II. Weak lensing. *Physical Review D*, **78**(4), 043520.

Hirata, C. M. & Seljak, U., 2003. Reconstruction of lensing from the cosmic microwave background polarization. *Physical Review D*, **68**(8), 083002.

Hobson, M. P., Bridle, S. L. & Lahav, O., 2002. Combining cosmological data sets: hyperparameters and Bayesian evidence. *MNRAS*, **335**, 377–388.

Hobson, M. P. & Magueijo, J., 1996. Observability of secondary Doppler peaks in the cosmic microwave background radiation power spectrum by experiments with small fields. *MNRAS*, **283**, 1133–1146.

Hopkins, P. F., Somerville, R. S., Hernquist, L., Cox, T. J., Robertson, B. & Li, Y., 2006. The Relation between Quasar and Merging Galaxy Luminosity Functions and the Merger-driven Star Formation History of the Universe. *The Astrophysical Journal*, **652**, 864–888.

Hou, Z., Reichardt, C. L., Story, K. T., Follin, B., Keisler, R. et al., 2014. Constraints on Cosmology from the Cosmic Microwave Background Power Spectrum of the 2500 deg^2 SPT-SZ Survey. *ApJ*, **782**, 74.

Hu, W. & Dodelson, S., 2002. Cosmic Microwave Background Anisotropies. *ARAA*, **40**, 171–216.

Hu, W., Eisenstein, D. J. & Tegmark, M., 1998. Weighing Neutrinos with Galaxy Surveys. *Physical Review Letters*, **80**, 5255–5258.

Hu, W. & Okamoto, T., 2002. Mass Reconstruction with Cosmic Microwave Background Polarization. *The Astrophysical Journal*, **574**, 566–574.

Hu, W. & Sugiyama, N., 1995. Anisotropies in the cosmic microwave background: an analytic approach. *ApJ*, **444**, 489–506.

Hu, W., Sugiyama, N. & Silk, J., 1997. The physics of microwave background anisotropies. *Nature*, **386**, 37–43.

Ichikawa, K., Fukugita, M. & Kawasaki, M., 2005. Constraining neutrino masses by CMB experiments alone. *Physical Review D*, **71**(4), 043001.

Jarosik, N., Bennett, C. L., Dunkley, J., Gold, B., Greason, M. R. et al., 2011. Seven-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Sky Maps, Systematic Errors, and Basic Results. *The Astrophysical Journal Supplement*, **192**, 14.

Jarvis, M. J., Bacon, D., Blake, C., Brown, M. L., Lindsay, S. N., Raccanelli, A., Santos, M. & Schwarz, D., 2015. Cosmology with SKA Radio Continuum Surveys. *ArXiv e-prints, arXiv:1501.03825*.

Jarvis, M. J., Bonfield, D. G., Bruce, V. A., Geach, J. E., McAlpine, K. et al., 2013. The VISTA Deep Extragalactic Observations (VIDEO) survey. *Monthly Notices of the Royal Astronomical Society*, **428**, 1281–1295.

Jarvis, M. J. & Rawlings, S., 2000. On the redshift cut-off for flat-spectrum radio sources. *Monthly Notices of the Royal Astronomical Society*, **319**, 121–136.

Jarvis, M. J. & Rawlings, S., 2004. The accretion history of the universe with the SKA. *New Astronomy Reviews*, **48**, 1173–1185.

Jarvis, M. J., Rawlings, S., Eales, S., Blundell, K. M., Bunker, A. J., Croft, S., McLure, R. J. & Willott, C. J., 2001a. A sample of 6C radio sources designed to find objects at redshift z>4 - III. Imaging and the radio galaxy K-z relation. *Monthly Notices of the Royal Astronomical Society*, **326**, 1585–1600.

Jarvis, M. J., Rawlings, S., Willott, C. J., Blundell, K. M., Eales, S. & Lacy, M., 2001b. On the redshift cut-off for steep-spectrum radio sources. *Monthly Notices of the Royal Astronomical Society*, **327**, 907–917.

Jaynes, E. T., 1957. Information Theory and Statistical Mechanics. *Physical Review*, **106**, 620–630.

Jaynes, E. T. & Bretthorst, G. L., 2003. *Probability Theory*.

Joudaki, S. & Kaplinghat, M., 2012. Dark energy and neutrino masses from future measurements of the expansion history and growth of structure. *Physical Review D*, **86**(2), 023526.

Kaiser, N., 1984. On the spatial correlations of Abell clusters. *Ap. J. Lett.*, **284**, L9–L12.

Kaiser, N. & Squires, G., 1993. Mapping the dark matter with weak gravitational lensing. *ApJ*, **404**, 441–450.

Kamionkowski, M., Kosowsky, A. & Stebbins, A., 1997. Statistics of cosmic microwave background polarization. *Physical Review D*, **55**, 7368–7388.

Keisler, R., Hoover, S., Harrington, N., Henning, J. W., Ade, P. A. R. et al., 2015. Measurements of Sub-degree B-mode Polarization in the Cosmic Microwave Background from 100 Square Degrees of SPTpol Data. *ApJ*, **807**, 151.

Kesden, M., Cooray, A. & Kamionkowski, M., 2003. Lensing reconstruction with CMB temperature and polarization. *PRD*, **67**(12), 123507.

Khoury, J., 2013. Les Houches Lectures on Physics Beyond the Standard Model of Cosmology. *ArXiv e-prints*.

Kirk, D., Benoit-Lévy, A., Abdalla, F. B., Bull, P. & Joachimi, B., 2015a. Cross correlation surveys with the Square Kilometre Array. *ArXiv e-prints, arXiv:1501.03848*.

Kirk, D., Omori, Y., Benoit-Lévy, A., Cawthon, R., Chang, C. et al., 2015b. Cross-correlation of gravitational lensing from DES Science Verification data with SPT and Planck lensing. *ArXiv e-prints*.

Kitching, T. D., Heavens, A. F., Verde, L., Serra, P. & Melchiorri, A., 2008. Finding evidence for massive neutrinos using 3D weak lensing. *Physical Review D*, **77**(10), 103008.

Kitching, T. D., Verde, L., Heavens, A. F. & Jimenez, R., 2016. Discrepancies between CFHTLenS cosmic shear and Planck: new physics or systematic effects? *ArXiv e-prints*.

Knox, L., 1995. Determination of inflationary observables by cosmic microwave background anisotropy experiments. *Physical Review D*, **52**, 4307–4318.

Kogut, A., Fixsen, D. J., Chuss, D. T., Dotson, J., Dwek, E., Halpern, M., Hinshaw, G. F., Meyer, S. M., Moseley, S. H., Seiffert, M. D., Spergel, D. N. & Wollack, E. J., 2011. The Primordial Inflation Explorer (PIXIE): a nulling polarimeter for cosmic microwave background observations. *JCAP*, **7**, 25.

Komatsu, E., Smith, K. M., Dunkley, J., Bennett, C. L., Gold, B. et al., 2011. Seven-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation. *The Astrophysical Journal Supplement*, **192**, 18.

Komatsu, E. et al., 2009. Five-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation. *Astrophys. J. Suppl.*, **180**, 330–376.

Koopmans, L., Pritchard, J., Mellema, G., Aguirre, J., Ahn, K. et al., 2015. The Cosmic Dawn and Epoch of Reionisation with SKA. *Advancing Astrophysics with the Square Kilometre Array (AASKA14)*, 1.

Kuijken, K., Heymans, C., Hildebrandt, H., Nakajima, R., Erben, T. et al., 2015. Gravitational lensing analysis of the Kilo-Degree Survey. *MNRAS*, **454**, 3500–3532.

Lahav, O., Kiakotou, A., Abdalla, F. B. & Blake, C., 2010. Forecasting neutrino masses from galaxy clustering in the Dark Energy Survey combined with the Planck measurements. *MNRAS*, **405**, 168–176.

Larsen, P. & Challinor, A., 2015. Intrinsic alignment contamination to CMB lensing-galaxy weak lensing correlations from tidal torquing. *ArXiv e-prints*.

Larson, D., Dunkley, J., Hinshaw, G., Komatsu, E., Nolta, M. R. et al., 2011. Seven-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Power Spectra and WMAP-derived Parameters. *Astrophysical Journal Supplement*, **192**, 16.

Laureijs, R., Amiaux, J., Arduini, S., Augueres, J. ., Brinchmann, J., Cole, R., Cropper, M., Dabin, C., Duvet, L., Ealet, A. & et al., 2011. Euclid Definition Study Report. *ArXiv e-prints, arXiv:1110.3193*.

Leclercq, F., 2015. Bayesian large-scale structure inference and cosmic web analysis. *ArXiv e-prints*.

Leonard, C. D., Baker, T. & Ferreira, P. G., 2015. Exploring degeneracies in modified gravity with weak lensing. *PRD*, **91**(8), 083504.

Lesgourgues, J., 2011. The Cosmic Linear Anisotropy Solving System (CLASS) I: Overview. *ArXiv e-prints*.

Lesgourgues, J. & Pastor, S., 2006. Massive neutrinos and cosmology. *Physics Reports*, **429**, 307–379.

Lesgourgues, J. & Pastor, S., 2012. Neutrino mass from Cosmology. *ArXiv e-prints*.

Levi, M., Bebek, C., Beers, T., Blum, R., Cahn, R., Eisenstein, D., Flaugher, B., Honscheid, K., Kron, R., Lahav, O., McDonald, P., Roe, N., Schlegel, D. & representing the DESI collaboration, 2013. The DESI Experiment, a whitepaper for Snowmass 2013. *ArXiv e-prints*.

Lewis, A., 2013. Efficient sampling of fast and slow cosmological parameters. *Physical Review D*, **87**(10), 103529.

Lewis, A. & Bridle, S., 2002. Cosmological parameters from CMB and other data: a Monte-Carlo approach. *Phys. Rev.*, **D66**, 103511.

Lewis, A. & Challinor, A., 2006. Weak gravitational lensing of the CMB. *Physics Reports*, **429**, 1–65.

Lewis, A., Challinor, A. & Lasenby, A., 2000. Efficient Computation of CMB anisotropies in closed FRW models. *Astrophys. J.*, **538**, 473–476.

Liddle, A. R., 1994. The inflationary energy scale. *Phys. Rev. D*, **49**, 739–747.

Limber, D. N., 1953. The Analysis of Counts of the Extragalactic Nebulae in Terms of a Fluctuating Density Field. *ApJ*, **117**, 134.

Linder, E. V., 2003. Exploring the Expansion History of the Universe. *Physical Review Letters*, **90**(9), 091301.

Lindsay, S. N., Jarvis, M. J. & McAlpine, K., 2014a. Evolution in the bias of faint radio sources to z $\sim$ 2.2. *MNRAS*, **440**, 2322–2332.

Lindsay, S. N., Jarvis, M. J., Santos, M. G., Brown, M. J. I., Croom, S. M., Driver, S. P., Hopkins, A. M., Liske, J., Loveday, J., Norberg, P. & Robotham, A. S. G., 2014b. Galaxy and Mass Assembly: the evolution of bias in the radio source population to z$\sim$1.5. *MNRAS*, **440**, 1527–1541.

Liu, A., Pritchard, J. R., Allison, R., Parsons, A. R., Seljak, U. & Sherwin, B. D., 2015. Eliminating the optical depth nuisance from the CMB with 21 cm cosmology. *ArXiv e-prints*.

Liu, J. & Hill, J. C., 2015. Cross-correlation of Planck CMB lensing and CFHTLenS galaxy weak lensing maps. *PRD*, **92**(6), 063517.

Liu, J., Ortiz-Vazquez, A. & Hill, J. C., 2016. Constraining Multiplicative Bias in CFHTLenS Weak Lensing Shear Data. *ArXiv e-prints*.

López-Caniego, M., Rebolo, R., Aguiar, M., Génova-Santos, R., Gómez-Reñasco, F. et al., 2014. The QUIJOTE CMB Experiment: status and first results with the multi-frequency instrument. *ArXiv e-prints*.

Loredo, T. J., 2004. Bayesian Adaptive Exploration. In G. J. Erickson & Y. Zhai, eds., *Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, vol. 707 of *American Institute of Physics Conference Series*, 330–346.

LoVerde, M. & Afshordi, N., 2008. Extended Limber approximation. *Phys. Rev. D*, **78**, 123506. URL `http://link.aps.org/doi/10.1103/PhysRevD.78.123506`.

LSST Dark Energy Science Collaboration, 2012. Large Synoptic Survey Telescope: Dark Energy Science Collaboration. *ArXiv e-prints*.

LSST Science Collaboration, Abell, P. A., Allison, J., Anderson, S. F., Andrew, J. R., Angel, J. R. P., Armus, L., Arnett, D., Asztalos, S. J., Axelrod, T. S. & et al., 2009. LSST Science Book, Version 2.0. *ArXiv e-prints*.

Ma, C.-P. & Bertschinger, E., 1995. Cosmological Perturbation Theory in the Synchronous and Conformal Newtonian Gauges. *ApJ*, **455**, 7.

Maartens, R., Abdalla, F. B., Jarvis, M., Santos, M. G. & SKA Cosmology SWG, f. t., 2015. Cosmology with the SKA – overview. *ArXiv e-prints, arXiv:1501.04076*.

Maartens, R., Zhao, G.-B., Bacon, D., Koyama, K. & Raccanelli, A., 2013. Relativistic corrections and non-Gaussianity in radio continuum surveys. *JCAP*, **2**, 44.

MacCrann, N., Zuntz, J., Bridle, S., Jain, B. & Becker, M. R., 2015. Cosmic discordance: are Planck CMB and CFHTLenS weak lensing measurements out of tune? *MNRAS*, **451**, 2877–2888.

Mackay, D., 2003. *Information Theory, Inference, and Learning Algorithms*. Cambridge University Press.

Madhavacheril, M., Sehgal, N., Allison, R., Battaglia, N., Bond, J. et al., 2014. The Atacama Cosmology Telescope: Detection of Lensing of the Cosmic Microwave Background by Dark Matter Halos. *ArXiv e-prints, arXiv:1411.7999*.

Magliocchetti, M., Maddox, S. J., Lahav, O. & Wall, J. V., 1998. Variance and skewness in the FIRST survey. *MNRAS*, **300**, 257–268.

Maltoni, M., Schwetz, T., Tórtola, M. & Valle, J. W. F., 2004. Status of global fits to neutrino oscillations. *New Journal of Physics*, **6**, 122.

Mana, G., Giuliano Albo, P. A. & Lago, S., 2013. Bayesian estimate of the degree of a polynomial given a noisy data sample. *ArXiv e-prints*.

Mandelbaum, R., 2015. Instrumental systematics and weak gravitational lensing. *Journal of Instrumentation*, **10**, C5017.

Marsh, D. J. E., Macaulay, E., Trebitsch, M. & Ferreira, P. G., 2012. Ultralight axions: Degeneracies with massive neutrinos and forecasts for future cosmological observations. *Physical Review D*, **85**(10), 103514.

Martin, J., Ringeval, C. & Vennin, V., 2014. Encyclopædia Inflationaris. *Physics of the Dark Universe*, **5**, 75–235.

Mather, J. C., Cheng, E. S., Cottingham, D. A., Eplee, Jr., R. E., Fixsen, D. J. et al., 1994. Measurement of the cosmic microwave background spectrum by the COBE FIRAS instrument. *Ap. J.*, **420**, 439–444.

Matsumura, T., Akiba, Y., Borrill, J., Chinone, Y., Dobbs, M. et al., 2014. Mission Design of LiteBIRD. *Journal of Low Temperature Physics*, **176**, 733–740.

Mauch, T. & Sadler, E. M., 2007. Radio sources in the 6dFGS: local luminosity functions at 1.4GHz for star-forming galaxies and radio-loud AGN. *Monthly Notices of the Royal Astronomical Society*, **375**, 931–950.

Mead, A., Heymans, C., Lombriser, L., Peacock, J., Steele, O. & Winther, H., 2016. Accurate halo-model matter power spectra with dark energy, massive neutrinos and modified gravitational forces. *ArXiv e-prints*.

Mead, A. J., Peacock, J. A., Heymans, C., Joudaki, S. & Heavens, A. F., 2015. An accurate halo model for fitting non-linear cosmological power spectra and baryonic feedback models. *MNRAS*, **454**, 1958–1975.

Mészáros, P., 1974. The behaviour of point masses in an expanding cosmological substratum. *A&A*, **37**, 225–228.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. & Teller, E., 1953. Equation of State Calculations by Fast Computing Machines. *The Journal of Chemical Physics*, **21**, 1087–1092.

Miller, L., Heymans, C., Kitching, T. D., van Waerbeke, L., Erben, T. et al., 2013. Bayesian galaxy shape measurement for weak lensing surveys - III. Application to the Canada-France-Hawaii Telescope Lensing Survey. *MNRAS*, **429**, 2858–2880.

Mo, H. J. & White, S. D. M., 1996. An analytic model for the spatial clustering of dark matter haloes. *Monthly Notices of the Royal Astronomical Society*, **282**, 347–361.

Mocz, P., Fabian, A. C. & Blundell, K. M., 2013. Cosmological growth and feedback from supermassive black holes. *Monthly Notices of the Royal Astronomical Society*, **432**, 3381–3390.

Mueller, E.-M., de Bernardis, F., Bean, R. & Niemack, M. D., 2015. Constraints on massive neutrinos from the pairwise kinematic Sunyaev-Zel'dovich effect. *Physical Review D*, **92**(6), 063501.

Mukherjee, P. & Parkinson, D., 2008. Cosmological Model Selection. *International Journal of Modern Physics A*, **23**, 787–802.

Mukherjee, P., Parkinson, D. & Liddle, A. R., 2006. A Nested Sampling Algorithm for Cosmological Model Selection. *The Astrophysical Journal*, **638**, L51–L54.

Naess, S., Hasselfield, M., McMahon, J., Niemack, M. D., Addison, G. E. et al., 2014. The Atacama Cosmology Telescope: CMB polarization at $200 < l < 9000$. *JCAP*, **10**, 7.

Namikawa, T., Saito, S. & Taruya, A., 2010. Probing dark energy and neutrino mass from upcoming lensing experiments of CMB and galaxies. *JCAP*, **12**, 27.

Namikawa, T. & Takahashi, R., 2014. Bias-hardened CMB lensing with polarization. *MNRAS*, **438**, 1507–1517.

Niemack, M. D., Ade, P. A. R., Aguirre, J., Barrientos, F., Beall, J. A. et al., 2010. ACTPol: a polarization-sensitive receiver for the Atacama Cosmology Telescope. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, vol. 7741 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*.

Norris, R. P., Afonso, J., Bacon, D., Beck, R., Bell, M. et al., 2013. Radio Continuum Surveys with Square Kilometre Array Pathfinders. *PASA*, **30**, 20.

Okamoto, T. & Hu, W., 2003. Cosmic microwave background lensing reconstruction on the full sky. *Phys. Rev. D*, **67**(8), 083002.

Overzier, R. A., Röttgering, H. J. A., Rengelink, R. B. & Wilman, R. J., 2003. The spatial clustering of radio sources in NVSS and FIRST; implications for galaxy clustering evolution. *AAP*, **405**, 53–72.

Palanque-Delabrouille, N., Yeche, C., Baur, J., Magneville, C., Rossi, G., Lesgourgues, J., Borde, A., Burtin, E., LeGoff, J.-M., Rich, J., Viel, M. & Weinberg, D., 2015. Cosmology with Lyman-alpha forest power spectrum. *ArXiv e-prints*.

Palanque-Delabrouille, N., Yèche, C., Borde, A., Le Goff, J.-M., Rossi, G. et al., 2013. The one-dimensional Lyα forest power spectrum from BOSS. *AAP*, **559**, A85.

Pan, Z. & Knox, L., 2015. Constraints on neutrino mass from Cosmic Microwave Background and Large Scale Structure. *ArXiv e-prints*.

Pan, Z., Knox, L. & White, M., 2014. Dependence of the cosmic microwave background lensing power spectrum on the matter density. *MNRAS*, **445**, 2941–2945.

Parkinson, D. & Liddle, A. R., 2013. Bayesian model averaging in astrophysics: a review. *Statistical Analysis and Data Mining*, **6**(1), 3–14. ISSN 1932-1872. URL `http://dx.doi.org/10.1002/sam.11179`.

Peacock, J. A., 1998. The isotropic universe. In *Cosmological Physics*, 65–100. Cambridge University Press. ISBN 9780511804533. URL `http://dx.doi.org/10.1017/CBO9780511804533.004`. Cambridge Books Online.

Pearson, R. & Zahn, O., 2014. Cosmology from cross correlation of CMB lensing and galaxy surveys. *Physical Review D*, **89**(4), 043516.

Peebles, P., 1980. *The Large-scale Structure of the Universe*. Princeton series in physics. Princeton University Press. ISBN 9780691082400. URL `https://books.google.co.uk/books?id=0_BPaHFtX1YC`.

Qian, S. S., Stow, C. A. & Borsuk, M. E., 2003. On Monte Carlo methods for Bayesian inference. *Ecological Modelling*, **159**(2Ð3), 269 – 277. ISSN 0304-3800. URL `http://www.sciencedirect.com/science/article/pii/S0304380002002995`.

Raccanelli, A., Zhao, G.-B., Bacon, D. J., Jarvis, M. J., Percival, W. J., Norris, R. P., Röttgering, H., Abdalla, F. B., Cress, C. M., Kubwimana, J.-C., Lindsay, S., Nichol, R. C., Santos, M. G. & Schwarz, D. J., 2012. Cosmological measurements with forthcoming radio continuum surveys. *MNRAS*, **424**, 801–819.

Rao, R. C., 1945. Information and the accuracy attainable in the estimation of statistical parameters. *Bull. Calcutta Math. Soc.*, **37**, 81–91. URL `http://www.ams.org/mathscinet-getitem?mr=0015748`.

Rawlings, S. & Jarvis, M. J., 2004. Evidence that powerful radio jets have a profound influence on the evolution of galaxies. *Monthly Notices of the Royal Astronomical Society*, **355**, L9–L12.

Rengelink, R. B., Tang, Y., de Bruyn, A. G., Miley, G. K., Bremer, M. N., Roettgering, H. J. A. & Bremer, M. A. R., 1997. The Westerbork Northern Sky Survey (WENSS), I. A 570 square degree Mini-Survey around the North Ecliptic Pole. *AAPS*, **124**, 259–280.

Rigby, E. E., Best, P. N., Brookes, M. H., Peacock, J. A., Dunlop, J. S., Röttgering, H. J. A., Wall, J. V. & Ker, L., 2011. The luminosity-dependent high-redshift turnover in the steep spectrum radio luminosity function: clear evidence for downsizing in the radio-AGN population. *Monthly Notices of the Royal Astronomical Society*, **416**, 1900–1915.

Riotto, A., 2010. Particle cosmology. *ArXiv e-prints*.

Rocca-Volmerange, B., Le Borgne, D., De Breuck, C., Fioc, M. & Moy, E., 2004. The radio galaxy K-z relation: The 10^{12} $M_?$ mass limit. Masses of galaxies from the L_K luminosity, up to $z > 4$. *Astronomy and Astrophysics*, **415**, 931–940.

Ross, A. J., Samushia, L., Howlett, C., Percival, W. J., Burden, A. & Manera, M., 2015. The clustering of the SDSS DR7 main Galaxy sample - I. A 4 per cent distance measure at $z = 0.15$. *MNRAS*, **449**, 835–847.

Ruiz de Austri, R., Trotta, R. & Roszkowski, L., 2006. A Markov chain Monte Carlo analysis of the CMSSM. *Journal of High Energy Physics*, **5**, 2.

Santos, M. G., Bull, P., Alonso, D., Camera, S., Ferreira, P. G., Bernardi, G., Maartens, R., Viel, M., Villaescusa-Navarro, F., Abdalla, F. B., Jarvis, M., Metcalf, R. B., Pourtsidou, A. & Wolz, L., 2015. Cosmology with a SKA HI intensity mapping survey. *ArXiv e-prints*.

Schmidt, F., 2008. Weak lensing probes of modified gravity. *PRD*, **78**(4), 043002.

Schmittfull, M. M., Challinor, A., Hanson, D. & Lewis, A., 2013. Joint analysis of CMB temperature and lensing-reconstruction power spectra. *Physical Review D*, **88**(6), 063012.

Seljak, U. & Warren, M. S., 2004. Large-scale bias and stochasticity of haloes and dark matter. *MNRAS*, **355**, 129–136.

Seljak, U. & Zaldarriaga, M., 1997. Signature of Gravity Waves in the Polarization of the Microwave Background. *Physical Review Letters*, **78**, 2054–2057.

Seymour, N., Stern, D., De Breuck, C., Vernet, J., Rettura, A. et al., 2007. The Massive Hosts of Radio Galaxies across Cosmic Time. *The Astrophysical Journal Supplement*, **171**, 353–375.

Shaw, J. R., Bridges, M. & Hobson, M. P., 2007. Efficient Bayesian inference for multimodal problems in cosmology. *MNRAS*, **378**, 1365–1370.

Shen, Y., Strauss, M. A., Ross, N. P., Hall, P. B., Lin, Y.-T., Richards, G. T., Schneider, D. P., Weinberg, D. H., Connolly, A. J., Fan, X., Hennawi, J. F., Shankar, F., Vanden Berk, D. E., Bahcall, N. A. & Brunner, R. J., 2009. Quasar Clustering from SDSS DR5: Dependences on Physical Properties. *The Astrophysical Journal*, **697**, 1656–1673.

Sherwin, B. D., Das, S., Hajian, A., Addison, G., Bond, J. R. et al., 2012. The Atacama Cosmology Telescope: Cross-correlation of cosmic microwave background lensing and quasars. *Physical Review D*, **86**(8), 083006.

Sherwin, B. D., Dunkley, J., Das, S., Appel, J. W., Bond, J. R. et al., 2011. Evidence for Dark Energy from the Cosmic Microwave Background Alone Using the Atacama Cosmology Telescope Lensing Measurements. *Physical Review Letters*, **107**(2), 021302.

Sievers, J. L., Hlozek, R. A., Nolta, M. R., Acquaviva, V., Addison, G. E. et al., 2013. The Atacama Cosmology Telescope: cosmological parameters from three seasons of data. *JCAP*, **10**, 60.

Silk, J., 1968. Cosmic Black-Body Radiation and Galaxy Formation. *ApJ*, **151**, 459.

Sivia, D., 1996. *Data Analysis: A Bayesian Tutorial*. Oxford Science Publications. Clarendon Press. ISBN 9780198518891. URL `http://books.google.co.uk/books?id=wR5yljKasLsC`.

Skilling, J., 2004. Nested Sampling. In R. Fischer, R. Preuss & U. V. Toussaint, eds., *American Institute of Physics Conference Series*, vol. 735 of *American Institute of Physics Conference Series*, 395–405.

Planck Collaboration, 2005. *PLANCK - The Scientific Programme, ESA-SCI(2005)1, PLANCK Bluebook*. ESA-SCI(2005)1.

Planck Collaboration, 2013a. *Planck* 2013 results. XVI. Cosmological parameters. *ArXiv e-prints*.

Planck Collaboration, 2013b. *Planck* 2013 results. XVII. Gravitational lensing by large-scale structure. *ArXiv e-prints*.

Planck Collaboration, 2014a. *Planck* 2013 results. XV. CMB power spectra and likelihood. *Astronomy & Astrophysics*, **571**, A15.

Planck Collaboration, 2014b. *Planck* 2013 results. XVIII. The gravitational lensing-infrared background correlation. *Astronomy & Astrophysics*, **571**, A18.

Planck Collaboration, 2014c. *Planck* 2013 results. XXII. Constraints on inflation. *Astronomy & Astrophysics*, **571**, A22.

Planck Collaboration, 2015a. *Planck* 2015 results. I. Overview of products and scientific results. *ArXiv e-prints*.

Planck Collaboration, 2015b. *Planck* 2015 results. XI. CMB power spectra, likelihoods, and robustness of parameters. *ArXiv e-prints*.

Planck Collaboration, 2015c. *Planck* 2015 results. XIII. Cosmological parameters. *ArXiv e-prints*.

Planck Collaboration, 2015d. *Planck* 2015 results. XIV. Dark energy and modified gravity. *ArXiv e-prints*.

Planck Collaboration, 2015e. *Planck* 2015 results. XV. Gravitational lensing. *ArXiv e-prints, arXiv:1502.01591*.

Planck Collaboration, 2015f. *Planck* 2015 results. XX. Constraints on inflation. *ArXiv e-prints*.

Planck Collaboration, 2015g. *Planck* 2015 results. XXIV. Cosmology from Sunyaev-Zeldovich cluster counts. *ArXiv e-prints*.

Planck Collaboration, 2015h. *Planck* intermediate results. XLI. A map of lensing-induced B-modes. *ArXiv e-prints*.

Smirnov, A. Y., 2006. Neutrino mass and New physics. *Journal of Physics Conference Series*, **53**, 44–82.

Smith, K. M., Hanson, D., LoVerde, M., Hirata, C. M. & Zahn, O., 2012. Delensing CMB polarization with external datasets. *JCAP*, **6**, 14.

Smith, K. M., Hu, W. & Kaplinghat, M., 2006. Cosmological information from lensed CMB power spectra. *Physical Review D*, **74**(12), 123002.

Smith, K. M., Zahn, O. & Doré, O., 2007. Detection of gravitational lensing in the cosmic microwave background. *Physical Review D*, **76**(4), 043510.

Smith, R. E., Peacock, J. A., Jenkins, A., White, S. D. M., Frenk, C. S., Pearce, F. R., Thomas, P. A., Efstathiou, G. & Couchman, H. M. P., 2003. Stable clustering, the halo model and non-linear cosmological power spectra. *Monthly Notices of the Royal Astronomical Society*, **341**, 1311–1332.

Spergel, D. N., Verde, L., Peiris, H. V., Komatsu, E., Nolta, M. R., Bennett, C. L., Halpern, M., Hinshaw, G., Jarosik, N., Kogut, A., Limon, M., Meyer, S. S., Page, L., Tucker, G. S., Weiland, J. L., Wollack, E. & Wright, E. L., 2003. First-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Determination of Cosmological Parameters. *The Astrophysical Journal Supplement*, **148**, 175–194.

Spergel, D. N. & Zaldarriaga, M., 1997. Cosmic Microwave Background Polarization as a Direct Test of Inflation. *Physical Review Letters*, **79**, 2180–2183.

Story, K. T., Hanson, D., Ade, P. A. R., Aird, K. A., Austermann, J. E. et al., 2015. A Measurement of the Cosmic Microwave Background Gravitational Lensing Potential from 100 Square Degrees of SPTpol Data. *ApJ*, **810**, 50.

Swetz, D. S., Ade, P. A. R., Allen, C., Amiri, M., Appel, J. W. et al., 2008. Instrument design and characterization of the Millimeter Bolometer Array Camera on the Atacama Cosmology Telescope. In *Millimeter and Submillimeter Detectors and Instrumentation for Astronomy IV*, vol. 7020 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, 702008.

Takahashi, R., Sato, M., Nishimichi, T., Taruya, A. & Oguri, M., 2012. Revising the Halofit Model for the Nonlinear Matter Power Spectrum. *The Astrophysical Journal*, **761**, 152.

Tegmark, M., 2005. Cosmological neutrino bounds for non-cosmologists. *Phys. Scripta*, **T121**, 153–155.

The Dark Energy Survey Collaboration, 2015. Cosmology from Cosmic Shear with DES Science Verification Data. *ArXiv e-prints*.

The POLARBEAR Collaboration, 2010. The new generation CMB B-mode polarization experiment: POLARBEAR. *ArXiv e-prints*.

The POLARBEAR Collaboration, 2013. Measurement of the Cosmic Microwave Background Polarization Lensing Power Spectrum with the POLARBEAR experiment. *ArXiv e-prints*.

The POLARBEAR Collaboration, 2014a. A Measurement of the Cosmic Microwave Background B-mode Polarization Power Spectrum at Sub-degree Scales with POLARBEAR. *ApJ*, **794**, 171.

The POLARBEAR Collaboration, 2014b. Evidence for Gravitational Lensing of the Cosmic Microwave Background Polarization from Cross-Correlation with the Cosmic Infrared Background. *Physical Review Letters*, **112**(13), 131302.

Tinker, J. L., Robertson, B. E., Kravtsov, A. V., Klypin, A., Warren, M. S., Yepes, G. & Gottlöber, S., 2010. The Large-scale Bias of Dark Matter Halos: Numerical Calibration and Model Tests. *The Astrophysical Journal*, **724**, 878–886.

Tomaru, T., Hazumi, M., Lee, A. T., Ade, P., Arnold, K. et al., 2012. The POLARBEAR-2 experiment. URL `http://dx.doi.org/10.1117/12.926158`.

Trotta, R., 2007. Applications of Bayesian model selection to cosmological parameters. *MNRAS*, **378**, 72–82.

Troxel, M. A. & Ishak, M., 2014. Cross-correlation between cosmic microwave background lensing and galaxy intrinsic alignment as a contaminant to gravitational lensing cross-correlated probes of the Universe. *PRD*, **89**(6), 063528.

Vallinotto, A., 2012. Using Cosmic Microwave Background Lensing to Constrain the Multiplicative Bias of Cosmic Shear. *Ap. J.*, **759**, 32.

van Engelen, A., Keisler, R., Zahn, O., Aird, K. A., Benson, B. A. et al., 2012. A Measurement of Gravitational Lensing of the Microwave Background Using South Pole Telescope Data. *The Astrophysical Journal*, **756**, 142.

van Engelen, A., Sherwin, B. D., Sehgal, N., Addison, G. E., Allison, R. et al., 2014. The Atacama Cosmology Telescope: Lensing of CMB Temperature and Polarization Derived from Cosmic Infrared Background Cross-Correlation. *ArXiv e-prints, arXiv:1412.0626*.

Van Trees, H. L., 1968. *Detection, estimation, and modulation theory. 1. , detection, estimation, and linear modulation theory*. J. Wiley and sons, New York, Chichester. ISBN 0-471-89955-0. URL `http://opac.inria.fr/record=b1105063`.

Van Waerbeke, L., Benjamin, J., Erben, T., Heymans, C., Hildebrandt, H. et al., 2013. CFHTLenS: mapping the large-scale structure with gravitational lensing. *MNRAS*, **433**, 3373–3388.

Villaescusa-Navarro, F., Bull, P. & Viel, M., 2015. Weighing neutrinos with cosmic neutral hydrogen. *ArXiv e-prints*.

von Toussaint, U., 2011. Bayesian inference in physics. *Rev. Mod. Phys.*, **83**, 943–999. URL `http://link.aps.org/doi/10.1103/RevModPhys.83.943`.

Wall, J. V., Jackson, C. A., Shaver, P. A., Hook, I. M. & Kellermann, K. I., 2005. The Parkes quarter-Jansky flat-spectrum sample. III. Space density and evolution of QSOs. *Astronomy and Astrophysics*, **434**, 133–148.

Weinberg, S., 1972. *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity*. Wiley, New York, NY. URL `https://cds.cern.ch/record/100595`.

White, M., Myers, A. D., Ross, N. P., Schlegel, D. J., Hennawi, J. F. et al., 2012. The clustering of intermediate-redshift quasars as measured by the Baryon Oscillation Spectroscopic Survey. *Monthly Notices of the Royal Astronomical Society*, **424**, 933–950.

White, R. L., Becker, R. H., Helfand, D. J. & Gregg, M. D., 1997. A Catalog of 1.4 GHz Radio Sources from the FIRST Survey. *The Astrophysical Journal*, **475**, 479–493.

Wilman, R. J., Jarvis, M. J., Mauch, T., Rawlings, S. & Hickey, S., 2010. An infrared-radio simulation of the extragalactic sky: from the Square Kilometre Array to Herschel. *Monthly Notices of the Royal Astronomical Society*, **405**, 447–461.

Wilman, R. J., Miller, L., Jarvis, M. J., Mauch, T., Levrier, F., Abdalla, F. B., Rawlings, S., Klöckner, H.-R., Obreschkow, D., Olteanu, D. & Young, S., 2008. A semi-empirical simulation of the extragalactic radio continuum sky for next generation radio telescopes. *MNRAS*, **388**, 1335–1348.

Wu, W. L. K., Errard, J., Dvorkin, C., Kuo, C. L., Lee, A. T., McDonald, P., Slosar, A. & Zahn, O., 2014. A Guide to Designing Future Ground-based Cosmic Microwave Background Experiments. *ApJ*, **788**, 138.

Wylezalek, D., Galametz, A., Stern, D., Vernet, J., De Breuck, C., Seymour, N., Brodwin, M., Eisenhardt, P. R. M., Gonzalez, A. H., Hatch, N., Jarvis, M., Rettura, A., Stanford, S. A. & Stevens, J. A., 2013. Galaxy Clusters around Radio-loud Active Galactic Nuclei at 1.3 z 3.2 as Seen by Spitzer. *The Astrophysical Journal*, **769**, 79.

York, D. G., Adelman, J., Anderson, Jr., J. E., Anderson, S. F., Annis, J. et al., 2000. The Sloan Digital Sky Survey: Technical Summary. *The Astronomical Journal*, **120**, 1579–1587.

Zaldarriaga, M., 1997. Polarization of the microwave background in reionized models. *Physical Review D*, **55**, 1822–1829.

Zaldarriaga, M. & Seljak, U., 1997. All-sky analysis of polarization in the microwave background. *Physical Review D*, **55**, 1830–1840.

Zaldarriaga, M. & Seljak, U., 1998. Gravitational lensing effect on cosmic microwave background polarization. *Physical Review D*, **58**(2), 023003.

Zhao, G.-B., Wang, Y., Ross, A. J., Shandera, S., Percival, W. J. et al., 2015. The extended Baryon Oscillation Spectroscopic Survey (eBOSS): a cosmological forecast. *ArXiv e-prints*.