
CMB-S4 Science Book

CMB-S4 Collaboration
Working Draft

June 30, 2016

Executive Summary

We describe the scientific case for the next generation ground-based cosmic microwave background experiment, CMB-S4, consisting of dedicated telescopes at the South Pole, the high Chilean Atacama plateau and possibly a northern hemisphere site, all equipped with highly sensitive superconducting cameras that will provide a dramatic leap forward in cosmological studies, crossing critical thresholds in probing inflation, the number and masses of the neutrinos or the existence of other ‘dark radiation’, providing precise constraints on the nature of dark energy, and testing general relativity on large scales.

CMB-S4 will be the definitive ground-based CMB project. It will deliver the data set with which any model for the origin of the primordial fluctuations, be it inflationary or an alternative theory, must be consistent with to be viable. CMB-S4 will allow an investigation of inflation by making ultra-sensitive polarization measurements of the recombination bump at degree angular scales and the requisite measurements at arc minutes scales for de-lensing to search for the B-mode signature imprinted by primordial gravitational waves. It will allow the determination of the neutrino masses by making unprecedentedly precise reconstruction of the matter power spectrum by exploiting the correlations induced on the CMB from the gravitational lensing of large scale structure, and it will allow the determination of the effective number of neutrino-like species to high accuracy through detailed constraints on the energy density in the early Universe as revealed in the temperature and polarization angular power spectra. It will be used to investigate dark energy through growth of structure tests, i.e., CMB lensing and SZ clusters, and through testing gravity on large scales, i.e., though exploiting the kinematic SZ effect to measure the momentum field and large scale flows. The power of these probes is amplified by combining CMB-S4 data with galaxy surveys and Lyman alpha surveys, such as DESI, LSST, Euclid and WFIRST.

There is only one CMB sky. It holds a wealth of information on fundamental physics and the origin and evolution of the Universe. While we have learned a great deal from CMB measurements, including discoveries that have pointed the way to new physics, we have only begun to tap the information contained in CMB polarization, CMB lensing and secondary effects. CMB-S4 will maximize discovery space by producing high fidelity maps of enormous legacy value.

Through the efforts of the CMB experimental groups over the last decade, the technologies needed for CMB-S4 are now in place. There are, however, considerable technical challenges presented by the required scaling up of the instrumentation as well as by the scope and complexity of the data analysis and interpretation. CMB-S4 will require: scaled up superconducting detector arrays with well understood and robust material properties and processing techniques; high throughput mm-wave telescopes and optics with unprecedented precision and rejection of systematic contamination; full internal characterization of astronomical foreground emission; large cosmological simulations and theoretical modeling with accuracies yet to be achieved; and computational methods for extracting minute correlations in massive, multi-frequency data sets contaminated by noise and a host of known and unknown signals.

CMB-S4 is the logical successor to the ongoing Stage 3 CMB projects and should be implemented on a time scale that allows a transition from Stage 3 to Stage 4 that is as seamless and timely as possible, thereby preserving the expertise and ensuring a continued stream of CMB science results and the maximum synergistic enhancement of the science return from other contemporaneous cosmic surveys (e.g., LSST, DESI). This timing is possible as CMB-S4 will use existing technology optimized for scaling up, but with no major downselects expected. Information learned from ongoing Stage 3 experiments (e.g., foregrounds, level of the primordial tensor to scalar ratio) can be easily incorporated into the CMB-S4 survey strategy with little or no impact on its design. In fact, the sensitivity and fidelity of the foreground measurements needed to realize the goals of CMB-S4 will only be provided by CMB-S4 itself, at frequencies just below and above the main CMB channels.

Acknowledgements

Contents

1	Exhortations	1
1.1	Brief History and Current Status of CMB measurements	2
1.2	Science reach of CMB-S4	3
1.3	From science goals to CMB-S4 design	5
1.3.1	Conceptual design of CMB-S4	5
1.3.1.1	Sensitivity and detector count	5
1.3.1.2	Inflationary B-modes: low ℓ sensitivity, foregrounds and atmospheric noise mitigation	6
1.3.1.3	Neutrinos and dark energy: high ℓ sensitivity	6
1.3.2	Refining the CMB-S4 science case and key performance parameters	7
1.4	The Road from Stage 3 to Stage 4	7
2	Inflation Physics from the Cosmic Microwave Background	11
2.1	Introduction	11
2.2	Basics of cosmological inflation	13
2.2.1	Inflation basics I: A heuristic picture	13
2.2.2	Inflation basics II: Quantifying the predictions	15
2.3	Forecasted sensitivity to r	18
2.4	Implications of a detection of primordial gravitational waves	20
2.4.1	The energy scale of inflation	21
2.4.2	Planckian field ranges and symmetries	22
2.5	Implications of an improved upper limit on r	25
2.6	Beyond r : characterizing the B-mode spectrum, testing inflationary expectations and distinguishing models	27
2.6.1	Distinguishing vacuum fluctuations from other particle physics sources of B-modes	28
2.6.2	Probing matter and gravitational interactions at the inflationary scale	29

2.6.3	Constraining alternatives to inflation	32
2.6.4	Constraints on the graviton mass	33
2.7	Improved constraints on primordial density perturbations	34
2.7.1	The power spectrum of primordial density perturbations and mean curvature . . .	34
2.7.2	Higher order correlations of scalar modes	35
2.7.3	Forecasts	37
2.8	Isocurvature	39
2.9	Spatial Curvature	42
2.10	Microwave Background Anomalies	43
2.11	Summary	45
3	Neutrinos and Light Species from the Cosmic Microwave Background	47
3.1	Introduction	47
3.2	Neutrino Mass	48
3.2.1	Theory Review	48
3.2.2	Observational Signatures and Target Sensitivity	50
3.2.3	CMB Lensing	51
3.2.4	Other Cosmological Probes	52
3.2.4.1	Galaxy Cluster Abundance	52
3.2.4.2	CMB Measurements in Context with Other Datasets	53
3.2.5	Forecasts	54
3.2.6	Relation to Lab Experiments	55
3.2.6.1	Determining the neutrino mass scale	56
3.2.6.2	Lepton number violation: Majorana vs. Dirac neutrinos	57
3.2.6.3	Neutrino mass ordering and CP violation	58
3.3	Relativistic Species at Recombination	59
3.3.1	Thermal History of the Early Universe	60
3.3.2	Natural Target	62
3.3.3	Observational Signatures	64
3.3.4	Forecasts	65

3.4	Neutrino constraints on BSM physics	66
3.4.1	Axions	67
3.4.2	Sterile Neutrinos	68
3.4.3	Light Fermions and Vectors	69
3.4.4	Gravitinos	70
3.4.5	Gravitational Waves	70
3.5	Complementarity of CMB and BBN	72
3.5.1	Standard Big Bang Nucleosynthesis	72
3.5.2	Beyond the Standard Model	73
3.5.3	Complementarity with the CMB	73
3.5.4	Forecasts	74
3.6	Detection Scenarios for Labs and Cosmology	74
3.6.1	Neutrino Physics	75
3.6.2	Dark Sectors and Particle Physics	75
3.6.3	Dark Sectors and Neutrino Mass	77
4	Dark Energy and Dark Matter	81
4.1	Dark Energy and Modified Gravity	81
4.1.1	Models and parameters	81
4.1.2	CMB Dark Energy Observables	83
4.1.2.1	Cluster abundance and mass	83
4.1.2.2	Lensing	85
4.1.2.3	Kinematic SZ	86
4.1.3	Forecasts	86
4.1.3.1	Trigger	87
4.1.3.2	Equation of motion parametrization	88
4.1.3.3	Theory parametrization	89
4.2	Dark Matter	94
4.2.1	Dark Matter Annihilation	94
4.2.2	Non-standard Dark Matter Interactions	94

4.2.2.1	Dark Matter-Baryon Scattering	95
4.2.2.2	Dark Matter-Dark Radiation Interaction	95
5	Physics Beyond The Standard Model	99
5.1	Constraints on Axion-like particles	99
5.1.1	Constraints on the axion energy density	99
5.1.2	Axion Isocurvature	100
5.1.3	Constraints on Axion couplings	102
5.2	Cosmic Birefringence	104
5.3	Primordial Magnetic Fields	105
5.4	Cosmic Strings	106
6	CMB Lensing	109
6.1	Introduction to CMB Lensing	109
6.2	Measuring CMB Lensing	110
6.2.1	Constructing a Lensing Map	110
6.2.2	Lensing Power Spectrum	112
6.3	Cross Correlations with CMB Lensing	113
6.3.1	CMB Lensing Cross Galaxy Density	115
6.3.2	CMB Lensing Cross Galaxy Shear	116
6.3.3	CMB Halo Lensing	116
6.4	Delensing	117
6.5	Systematic Effects and Mitigation	119
6.5.1	Astrophysical Systematics	119
6.5.2	Instrumental and Modeling Systematics	120
6.6	Parameter Forecasts	121
7	Simulations and Data Analysis	123
7.1	Introduction	123
7.2	Data Analysis Overview	123
7.3	Time-Ordered Data Processing	126

7.3.1	Pre-Processing and Mission Characterization	126
7.3.2	Map-Making	127
7.4	Component Separation	129
7.4.1	Introduction	129
7.4.1.1	Motivations	129
7.4.1.2	Definition of component separation	130
7.4.2	Description of methods	131
7.4.3	Questions to be addressed during follow-up studies	133
7.5	Statistics and Parameters	135
7.5.1	Current practice	135
7.5.2	Challenges	137
7.5.2.1	Combining different data sets	137
7.5.2.2	Foreground-related uncertainty on cosmological parameters	137
7.5.2.3	CMB lensing covariances for CMB S4	138
7.5.2.4	Delensing	140
7.6	Simulation Overview	141
7.7	Sky Modeling	142
7.7.1	The Galactic interstellar medium	143
7.7.1.1	Synchrotron	144
7.7.1.2	Thermal dust	144
7.7.1.3	Spinning dust	144
7.7.1.4	Other components	144
7.7.2	CMB Secondary Anisotropies and Extragalactic Sources	145
7.8	Data Simulation	147
7.8.1	Time Domain	147
7.8.2	Map Domain	148
7.8.3	Spectral Domain	149
7.9	The Simulation and Data Analysis Pipeline	150
7.10	Forecasting	151
7.10.1	Forecasting CMB-S4 constraints on the tensor-to-scalar ratio	151

7.10.1.1	Spectrum-based domain forecasting	151
7.10.1.2	Map-based domain forecasting	152
7.10.2	Forecasting CMB-S4 constraints on parameters from TT/TE/EE/ $\kappa\kappa$	153
7.10.2.1	S4 specifications	154
7.10.2.2	Non-S4 data specifications	154
7.10.2.3	Fisher code validation	155
7.10.3	Forecasting CMB-S4 constraints on parameters from tSZ/kSZ	155
7.11	Implementation Issues	156
7.11.1	Time-Ordered Data Volume & High Performance Computing	156
8	Experimental Approach	159
8.1	Introduction	159
8.2	Technical Development	160
8.2.1	Detector Arrays	160
8.2.1.1	Development of mass production capability at the national labs	162
8.2.2	Multiplexed Detector Readout	163
8.2.3	Telescope Architecture and Reimaging Optics	163
8.2.4	Polarization Modulation	164

1

Exhortations

(send feedback on this chapter to jc@kicp.chicago.edu)

Fourteen billion years ago, in the first fraction of a second of our universe’s existence, the most extreme high-energy physics experiment took place. Our ability to use the cosmic microwave background (CMB) to investigate this fantastic event, at energy scales a trillion times higher than can be obtained at the CERN, is at the very core of our quest to understand the fundamental nature of space and time and the physics that drive the evolution of the universe.

The CMB allows direct tests of models of the quantum mechanical origin of all we see in the universe. Subtle correlations in its anisotropy imparted by the interplay of gravitational and quantum physics at high energies contain information on the unification of gravity and quantum physics. Separately, correlations induced on the background at later times encode details about the distribution of all the mass, ordinary and dark, in the universe, as well as the properties of the neutrinos, including the number of neutrino species and types, and their still unknown masses.

Here we describe the scientific case for the next generation ground-based cosmic microwave background experiment, CMB-S4, consisting of dedicated telescopes at the South Pole, the high Chilean Atacama plateau and possibly a northern hemisphere site, all equipped with new superconducting cameras that will provide a dramatic leap forward in cosmological studies, crossing critical thresholds in testing inflation, the number and masses of the neutrinos or the existence of other ‘dark radiation’, providing precise constraints on the nature of dark energy, and testing general relativity on large scales.

Through the efforts of the CMB experimental groups over the last decade, the technologies needed for CMB-S4 are now in place. There are, however, considerable technical challenges presented by the required scaling up of the instrumentation as well as by the scope and complexity of the data analysis and interpretation. CMB-S4 will require: scaled up superconducting detector arrays with well understood and robust material properties and processing techniques; high throughput mm-wave telescopes and optics with unprecedented precision and rejection of systematic contamination; full characterization of astronomical foreground emission; large cosmological simulations and theoretical modeling with accuracies yet to be achieved; and computational methods for extracting minute correlations in massive, multi-frequency data sets contaminated by noise and a host of known and unknown signals.

The purpose of this document is to set the scientific goals for CMB-S4 and (eventually) the instrumental configuration required to achieve them. This is of course an iterative process, involving detailed simulations as well as cost considerations. So, at this time the Science Book is a working document with this first iteration focused primarily on defining the possible science reach in several areas, along with the simulations needed to refine science case and set the specifications of the needed measurements. This will set the stage for defining the instrument in the next iteration of the Science Book.

In this chapter we set out the overarching goals for CMB-S4, which are then refined in later chapters. We start with a brief history and the current status of CMB measurements.

1.1 Brief History and Current Status of CMB measurements

From its discovery 50 years ago, measurements of the cosmic microwave background (CMB) have led to spectacular scientific insights into the fundamental workings of space and time, from the quantum mechanical origin of the Universe at extremely high energies in the first moments of the Universe, through the growth of structure and the emergence of the dark energy that now dominates the energy density of the Universe. Studies of the CMB connect physics at the smallest scales and highest energies with the largest scales in the Universe, roughly 68 orders of magnitude in length scale. They connect physics at the earliest times to the structure that surrounds us now, over 52 magnitudes in time scale.

The deep connections of CMB studies and particle physics predate the discovery of the background, going back to the 1940s when Alpher and Gamow were considering a hot, dense, early Universe as a possible site for nucleosynthesis. To produce the amount of helium observed in the local Universe, they concluded there had to be about 10^{10} thermal photons for every nucleon. Alpher and Herman subsequently predicted that this background of photons would persist to the present day as a thermal bath at a few degrees Kelvin.

The continuing, remarkably successful, story of CMB studies is one driven by the close interplay of theory and phenomenology with increasingly sensitive and sophisticated experiments. The high degree of isotropy of the CMB across the sky, to a part of one in a hundred thousandth, led to the theory of inflation and cold dark matter in the 1980's. It was not until 1992 that COBE discovered the anisotropy, and pinned the level of anisotropy for the following higher angular resolution measurements to characterize. In 2006 the COBE measurements of the background anisotropy and its black-body spectrum were recognized with the second Nobel Prize in physics; the first was awarded in 1978 to Penzias and Wilson for the discovery of the CMB. In the decade after the COBE results, measurements with ground and balloon-based instruments revealed the acoustic peaks in the CMB angular power spectrum, which showed that the Universe was geometrically flat in accord with predictions of inflation and provided strong support for contemporary Type 1a SN based claims for an accelerating Universe, which were recognized with the 2011 Nobel Prize in physics. The early anisotropy measurements also provided an estimate of the universal baryon density and found it to be in excellent agreement with the level estimated at $t \sim 1$ second by BBN calculations constrained to match the observed elemental abundances. The CMB measurements also clearly showed that dark matter was non-baryonic. The polarization anisotropy was discovered ten years after COBE at the level predicted from temperature anisotropy measurements. The now standard Λ CDM cosmological model was firmly established.

Two CMB satellites have mapped the entire sky over the last 15 years, first WMAP with moderate angular resolution up to 12 arcminutes, followed by Planck with resolution up to 5 arcminutes. Higher resolution maps of smaller regions of the sky have been provided by ground-based experiments, most notably by the 10m South Pole Telescope (SPT) and the 6m Atacama Cosmology Telescope. The primary CMB temperature anisotropy is now well characterized through the damping tail, i.e., to multipoles $\ell \sim 3000$, and secondary anisotropies have been measured to multipoles up to 10,000. The Λ CDM model continues to hold up stunningly well, even as the precision of the CMB determined parameters has increased substantially. Inflationary constraints include limits on curvature constrained to be less than 3% of the energy density, non-Gaussian fluctuations limited to $f_{NL} < 10$, and the predicted small departure from pure scale invariance of the primordial fluctuations detected at 5 sigma confidence. Also of interest to particle physics, the effective number of light relativistic species (i.e., neutrinos and any yet identified “dark radiation”) is shown to be within one sigma of $N_{\text{eff}} = 3.046$, the number predicted by BBN. The sum of the masses of the neutrinos is found to be less than 0.6 eV. Dark matter is shown to be non-baryonic matter at > 40 sigma. Early dark energy models are highly constrained as are models of decaying dark matter.

There remains much science to extract from the CMB, including: 1) using CMB B-mode polarization to search for primordial gravitational waves to constrain the energy scale of inflation and to test alternative

models, and to provide insights into quantum gravity; 2) obtaining sufficiently accurate and precise determinations of the effective number of light relativistic species (dark radiation) to allow independent and rigorous tests of BBN as well as our understanding of the evolution of the Universe at $t = 1$ sec; 3) a detection of the sum of the neutrino masses, even if at the minimum mass allowed by oscillation experiments and in the normal hierarchy; 4) using secondary CMB anisotropy measurements to provide precision tests of dark energy through its impact on the growth of structure; and 5) testing general relativity and constraining alternate theories of gravity on large scales.

Currently the best cosmological constraints come from analyzing the combination of primary and secondary CMB anisotropy measurements with other cosmological probes, such as baryon acoustic oscillations (BAO) and redshift distortions, weak lensing, galaxy and galaxy cluster surveys, Lyman-alpha forest, Hubble constant, Type 1a SN, and others. The CMB primary anisotropy measurements provide highly complementary data for the combined analysis, in particular by providing a precision measurement of the Universe at $z = 1100$, which will provide a precise prediction for measurements of the late time Universe for any cosmological model and set of parameters – the Hubble constant, BAO scale, and the normalization of the present day matter fluctuation spectrum being excellent examples. Secondary CMB measurements provide late time probes directly from the CMB measurement, e.g., CMB lensing, the SZ effects and SZ cluster catalogs, which will provide critical constraints on the standard cosmological models and extensions to it. The cosmological reach of future cosmological surveys at all wavelengths will be greatly extended by their joint analyses with secondary CMB anisotropy measurements.

1.2 Science reach of CMB-S4

CMB-S4 should be the definitive ground-based CMB project. The key science it should cover, and cover well, are

1. Inflation: CMB-S4 should make the definitive B-mode measurements of the recombination bump at degree angular scales. This includes multiple bands to untangle the foregrounds and degree through arcminute angular scales to obtain the required CMB lensing and E-mode measurements for de-lensing inflationary gravitational wave B-modes. If it can be demonstrated that foregrounds and atmospheric noise can be mitigated at very low multipoles, CMB-S4 should also target the re-ionization bump. At the lowest multipoles, CMB-S4, balloon and satellite mission would be highly complementary.

CMB-S4 should answer whether or not large scale slow-roll-single-field inflation models are viable ($r \gtrsim 0.01$) with high significance. If no detection at $r \gtrsim 0.01$, then CMB-S4 should be able to test the currently popular Starobinski model and others by achieving $\sigma(r) < 10^{-4}$ with an ultra deep survey.

If r is detected before or by CMB-S4, then CMB-S4 should provide a robust cosmic variance limited measure of its value (requiring a large area survey), and set the best possible constraints on n_t (requiring an ultra deep survey).

CMB-S4 should provide the polarization data to test predictions of models that attempt to explain the low-ell temperature power spectrum “anomalies”, that may offer clues to inflation. It will be particularly important to achieve accurate $20 < \ell < 100$ E-mode measurements.

CMB-S4 will also extend the leverage arm for n_s , particularly in the EE spectrum. It may be possible to extend the primary E-mode spectrum to multipoles exceeding 10,000 because of the very low level of polarized foregrounds at high ℓ .

The CMB-S4 data set should be the definitive data set with which any model for the origin of the primordial fluctuations, be it inflationary or an alternative theory, must be consistent with to be viable.

2. Neutrinos and light relativistic species:

There are two primary areas in which CMB-S4 will provide interesting neutrino constraints.

- a) The first is the effective number of light relativistic species, N_{eff} . This is uniquely probed by the CMB and provides a critical constraint on any model for the neutrinos and their interactions. It is a highly complementary probe to BBN and to sterile neutrino models. Finding consistency with $N_{\text{eff}} = 3.046$ at a precision of 0.020 would be an exciting and fundamental achievement linking particle physics and our understanding of the evolution of the first seconds of the Universe. Finding a departure from 3.046 would be even more exciting.
- b) The second is the constraint on the sum of the masses of the neutrinos, Σm_ν . Here CMB-S4 will achieve $\sigma(\Sigma m_\nu) = 16$ meV (with DESI BAO prior), with the CMB sensitivity coming primarily through CMB lensing. This will lead to a definite detection of neutrino mass, even at the minimum mass and the normal hierarchy. The sensitivity of cosmic probes to the sum of the masses is unique and complementary to terrestrial neutrino experiments.

3. Dark Energy and Gravity:

The CMB can be used to investigate dark energy through growth of structure tests, i.e., CMB lensing and SZ clusters, and through testing Gravity on large scales, i.e., though exploiting the kinematic SZ effect to measure the momentum field and large scale flows. The power of these probes is amplified by combining CMB-S4 data with galaxy surveys and Lyman alpha surveys, such as DESI, LSST, Euclid andWFIRST.

- a) CMB lensing maps from CMB-S4 will provide hi-fidelity projected mass maps that will be cross-correlated with optical survey maps. This will increase the reach and precision of the dark energy constraints, as well as provide independent checks. Papers in the literature have quantified the dark energy figure of merit (FOM) improvement of various projects with the addition of CMB lensing. Simulations need to be done to quantify the projected improvements with CMB-S4.
- b) The dark energy task force (DETF) pointed out that galaxy cluster evolution had the highest sensitivity of the DE probes considered. However, it also had the largest systematic. The issue is the uncertainty in understanding the mass scaling of the cluster observable. The thermal SZ effect has now been demonstrated to be a low scatter observable with the extraordinary feature of its brightness being redshift independent; an SZ survey probes all redshifts to a limiting mass. However, there still remain large uncertainties in the SZ observable mass scaling. CMB-S4 will be revolutionary in that it is expected to be able to calibrate the mass scaling to better than 1% through CMB lensing. This coupled with a low mass threshold will enable CMB-S4 to identify of order 100,000 clusters, probe the growth of structure to redshifts beyond $z \sim 2.5$, and will allow CMB-S4 to realize the full potential of galaxy clusters as a probe of dark energy. In combination with other Stage-IV baryon acoustic oscillation, supernova, and weak lensing surveys, a Stage-IV cluster survey similar to CMB-S4 should improve the overall dark energy figure of merit to approximately 1250, nearly a factor of two improvement than achieved without clusters.
- c) Testing GR on large scales is important for our understanding of dark energy and the underlying workings of space and matter in general. The kinematic SZ effect allows measurement on the peculiar velocity (departure from Hubble flow) of structures. By measuring the differences in kSZ between pairs of clusters with known redshifts (a synergy of CMB-S4 and optical surveys), gravity can be tested on scales of 100 Mpc and larger. In this way, CMB-S4 paired with a Stage-IV spectroscopic survey would improve constraints on the growth rate predicted by general relativity by a factor of two.

Lastly it would be an oversight not to point out the obvious: there is only one CMB sky. It holds a wealth of information on fundamental physics and the origin and evolution of the Universe. While we have learned

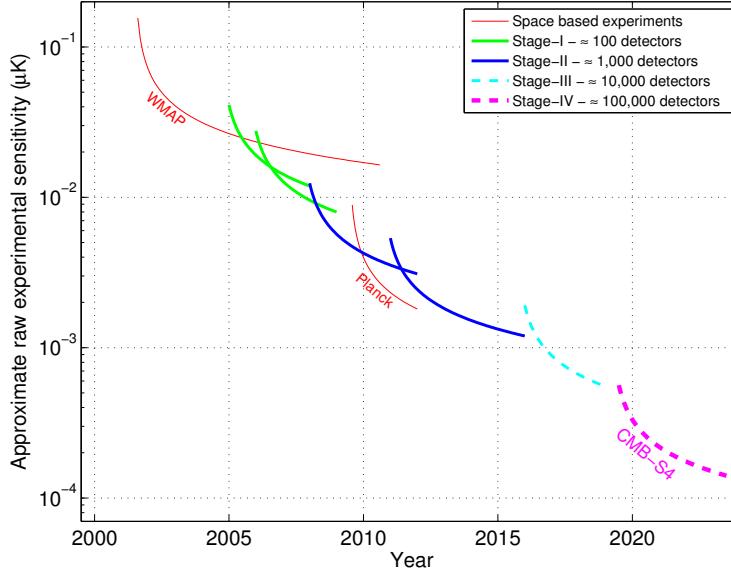


Figure 1. Plot illustrating the evolution of the raw sensitivity of CMB experiments, which scales as the total number of bolometers. Ground-based CMB experiments are classified into Stages with Stage II experiments having $O(1000)$ detectors, Stage III experiments having $O(10,000)$ detectors, and a Stage IV experiment (such as CMB-S4) having $O(100,000)$ detectors. Figure from Snowmass CF5 Neutrino planning document.

a great deal from CMB measurements, including discoveries that have pointed the way to new physics, we have only begun to tap the information contained in CMB polarization, CMB lensing and secondary effects. CMB-S4 should be designed to maximize discovery space by producing high fidelity maps.

1.3 From science goals to CMB-S4 design

1.3.1 Conceptual design of CMB-S4

The science goals discussed above leads to a rough conceptual design of CMB-S4.

1.3.1.1 Sensitivity and detector count

The sensitivity of CMB measurements has increased enormously since Penzias and Wilson's discovery in 1965, following a Moore's Law like scaling, doubling every roughly 2.3 years. Fig. 1 shows the sensitivity of recent experiments as well as expectations for upcoming Stage 3 experiments, characterized by order 10,000 detectors on the sky, as well as the projection for a Stage 4 experiment with order 100,000 detectors. To obtain many of the CMB-S4 science goals requires of order 1 μK arcminute sensitivity over roughly 70% of the sky, which for a four year survey requires of order 500,000 CMB-sensitive detectors.

To maintain the Moore's Law-like scaling requires a major leap forward, it requires a phase change in the mode of operation of the ground based CMB program. Two constraints drive the change: 1) CMB detectors

are background limited, so more pixels are needed on the sky to increase sensitivity; and 2) the pixel count for CMB cameras are nearing saturation. Even using multichroic pixels and wide field of view optics, CMB telescopes are expected to field only tens of thousands of polarization detectors, far fewer than needed to meet the CMB-S4 science goals.

CMB-S4 thus requires multiple telescopes, each with a maximally outfitted focal plane of pixels utilizing superconducting, background limited, CMB detectors. To achieve the large sky coverage and to take advantage of the best atmospheric conditions, the South Pole and the Chilean Atacama sites are baselined, with the possibility of adding a new northern site to increase sky coverage to 100%.

1.3.1.2 Inflationary B-modes: low ℓ sensitivity, foregrounds and atmospheric noise mitigation

At the largest angular scales (low ℓ)—the angular scales that must be measured well to pursue inflationary B-modes as well as critical tests of the E-mode polarization—the CMB polarization anisotropy is highly contaminated by foregrounds. Galactic synchrotron dominates at low frequencies and galactic dust at high frequencies, as recently shown by the Planck and Planck/BICEP/KECK polarization results. Multi-band polarization measurements are required to distinguish the primordial polarized signals from the foregrounds.

Adding to the complexity of low multipole CMB observations is the need to reject the considerable atmospheric noise contributions over the large scans needed to extract the low ℓ polarization. While the spatial and temporal fluctuations of the atmosphere are not expected to be polarized, any mismatches in the polarized beams or detector gains will lead to T-P leakage. These issues can be mitigated by including additional modulations into the instrument design, such as bore-sight rotation or modulation of the entire optics with a polarization modulation scheme in front of the telescope. Implementing such modulations is easier for small telescopes, although they could in principle be implemented on large telescopes as well. The cost of a small aperture telescope is dominated by the detector array, making it feasible to deploy multiple telescopes each optimized for a single band, or perhaps multiple bands within the relatively narrow atmosphere windows.

It is therefore an attractive option for CMB-S4 to include dedicated small aperture telescopes for pursuing low- ℓ polarization. The default plan for CMB-S4 is to target the recombination bump, with E-mode and B-mode polarization down to $\ell \sim 20$. If Stage 3 experiments demonstrate that it is feasible to target the reionization bump from the ground, those techniques may be incorporated into CMB-S4. More likely, however, this is the ℓ range for which CMB-S4 will be designed to be complementary to balloon-based and satellite based measurements.

1.3.1.3 Neutrinos and dark energy: high ℓ sensitivity

At the highest angular resolution (high ℓ)—the angular scales needed for de-lensing the inflationary B-modes, constraining N_{eff} and Σm_ν , investigating dark energy and performing gravity tests with secondary CMB anisotropy—the CMB polarization anisotropy is much less affected by both foregrounds and atmospheric noise. In fact, it should be possible to measure the primary CMB anisotropy in E-mode polarization to multipoles a factor of a few times higher than possible in TT, thereby extending the lever arm to measure the spectral index and running of the primordial scalar (density) fluctuations. CMB-lensing benefits from ℓ_{max} of order 5000 and secondary CMB measurements are greatly improved with ℓ_{max} of order 10,000 and higher, requiring large aperture telescopes with diameters of several meters. Owing to the steep scaling of telescope cost with aperture diameter, it is likely not cost-effective to consider separate large aperture telescopes each optimized for a single frequency band.

CMB-S4 is therefore envisioned to include dedicated large aperture, wide field of view telescopes equipped with multiple band detector arrays.

1.3.2 Refining the CMB-S4 science case and key performance parameters

The rough conceptual design outlined above clearly needs to be refined. The first priorities are to determine the instrumental specifications to meet each of the science goals. We need to determine: the required resolution and sensitivity; the number of bands to mitigate foreground contamination, which is likely to be function of angular scale; the required sky coverage; the beam specifications (can we tolerate segmented primary reflectors?); the scanning strategy and instrument stability; etc.

Determining these specifications requires simulations, informed by the best available data and phenomenological models. Only when we have these specifications in hand can we design the instrument and answer such basic questions as the number and sizes of the telescopes.

1.4 The Road from Stage 3 to Stage 4

The Stage 2 and 3 experiments are logical technical and scientific stepping stones to CMB-S4. Fig. 2 shows the timeline of the CMB sensitivity and the expected improvement in a few of the key cosmological parameters. The enormous jump in sensitivity with the corresponding improvement in science reach is clear.

Finally, in Fig. 3 we show how the scientific findings (yellow), the technical advances (blue) and satellite selections (green) would effect the science goals, survey strategy and possibly the design of CMB-S4. [Figure to be updated]

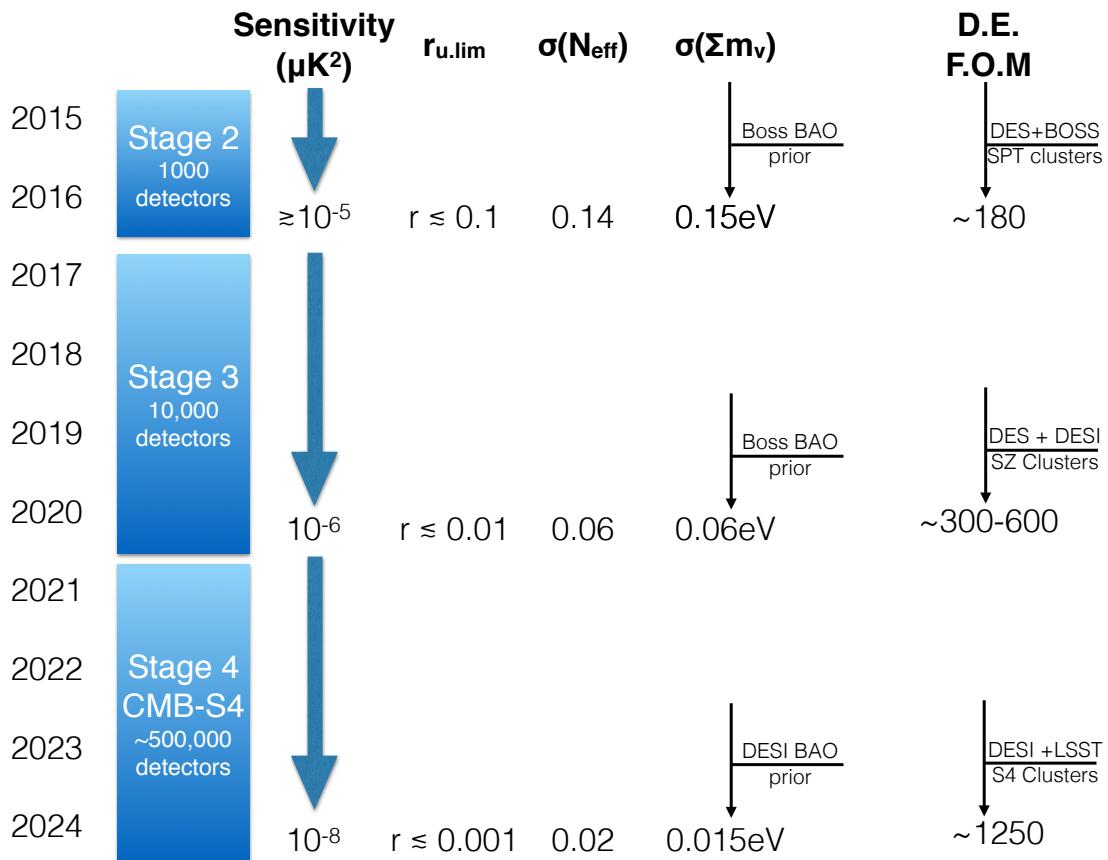


Figure 2. Schematic timeline of evolution of Stage 3 and CMB-S4 sensitivity in μK^2 and the expected improvement in a few of the key cosmological parameters.

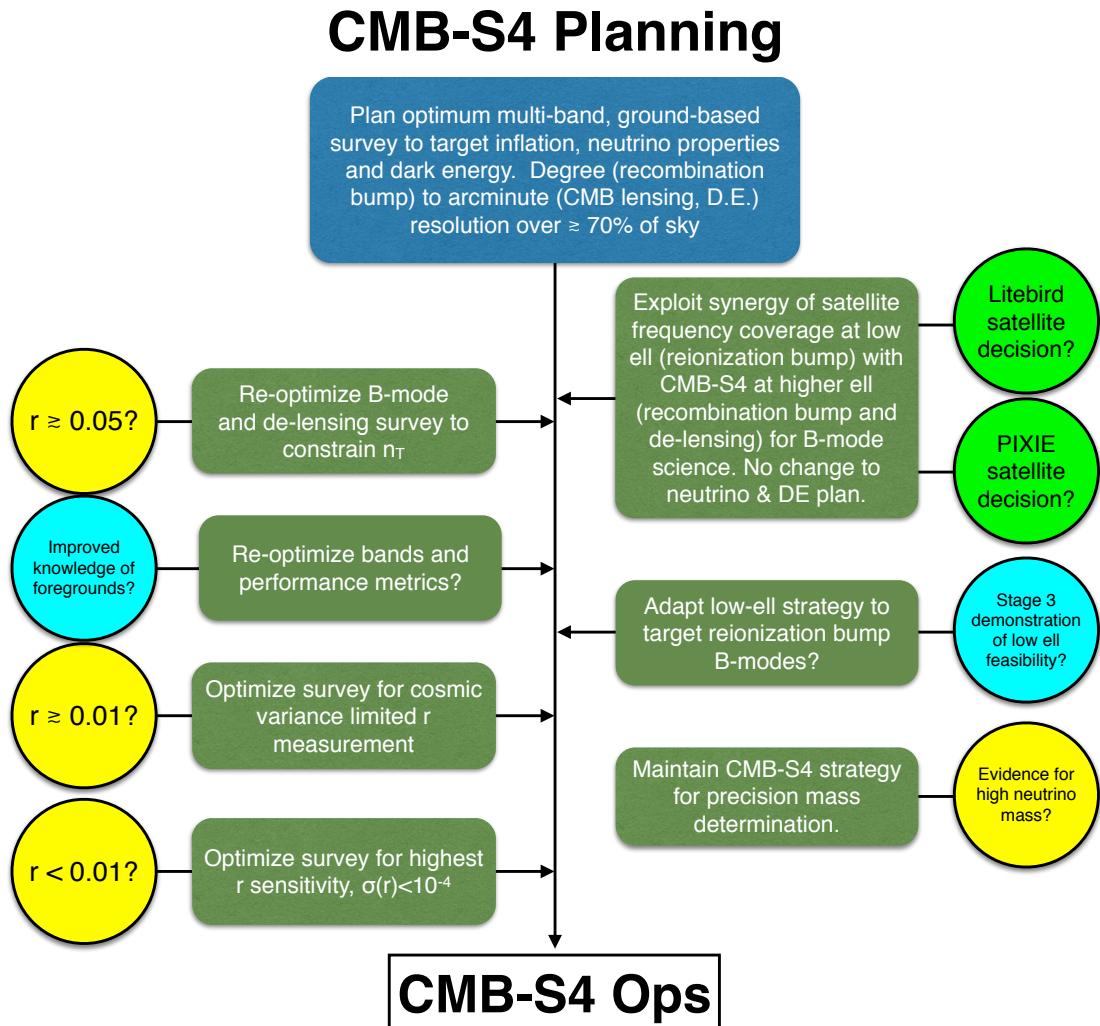


Figure 3. Schematic flow chart showing how the scientific findings (yellow), the technical advances (blue) and satellite selections (green) would effect the science goals, survey strategy and possibly the design of CMB-S4 (green boxes) [Figure to be updated]

Inflation Physics from the Cosmic Microwave Background

(send feedback on this chapter to s4_inflation@cosmo.uchicago.edu)

2.1 Introduction

The theory of cosmic inflation, in which the universe is proposed to have undergone a period of rapid, accelerating expansion in the first fraction of a second of its existence, is the current leading model for the universe's very early history. Proposed at first to solve certain conceptual puzzles of the hot big bang, inflation was soon recognized as a highly compelling mechanism for the generation of primordial perturbations as well. The inflationary paradigm has met with enormous empirical success. The simplest models of inflation predict that our Hubble patch should have nearly zero mean curvature and an adiabatic and nearly (but not quite) scale-invariant spectrum of highly Gaussian primordial scalar perturbations. To date, we have fairly tight upper limits on mean curvature, primordial non-Gaussianity, and the amplitudes of any non-adiabatic (isocurvature) contributions to the perturbations. We also have a detection of a small departure from scale invariance, consistent with the expectations of simple models.

These successes, and the difficulties in the construction of alternative scenarios that are also consistent with the data, have led most cosmologists to see inflation as our best bet for the creation of the primordial perturbations. But many questions remain. If inflation did occur, how did it occur? Was there a single effective field dominating the dynamics of both the background expansion and the perturbations, or were multiple fields involved? Are ground-state fluctuations truly the source of density perturbations? What is the connection of inflation physics to the rest of physics? What observations could falsify inflation, if any?

CMB-S4 will provide answers to some of these questions by opening a new window on the study of inflation, and on the generation of perturbations to the background in general. Thus far, observed anisotropies can all be interpreted as resulting from scalar perturbations to the spacetime metric tensor. With CMB-S4 we have an opportunity to open up an exciting new window on the mechanism of the creation of these primordial perturbations: the *tensor* sector. The sensitivity of CMB measurements to tensor perturbations arises from the now well-known phenomenon of the generation of polarization in the CMB: scalar perturbations produce only curl-free E-mode polarization to first order, while tensor perturbations produce divergence-free B-mode polarization as well. Thus a measurement of B-mode polarization in the CMB (with the standard caveats relating to foregrounds and gravitational lensing) is a direct measurement of the amplitude of tensor perturbations.

While the value of this new window can be understood more broadly than from an inflationary context; here we use that context as a useful concrete framework for exploring the potential impact of these measurements. The tensor sector offers a more direct probe of the dynamics of the inflationary expansion because the tensor perturbations are an inevitable consequence of the degrees of freedom of the spacetime metric obeying the uncertainty principle. In other words, the existence of an inflationary epoch in the universe's past directly implies the existence of a background of tensor perturbations. Furthermore, the amplitude of the tensor

perturbations depends only on the rate of expansion during inflation. In contrast, the amplitude of the scalar perturbations depends on both the amplitude and slope of the effective potential of the field responsible for inflation, and more generally on the sound speed of the inflaton field as well.

In addition to probing the origin of all structure in the universe, opening the tensor sector also opens up a probe of physics at length scales $\sim 10^9$ times smaller than those probed at the LHC. This small length scale is the size of the future horizon during inflation if it takes place at sufficiently high energies for us to observe the resulting tensor fluctuations. It is accessible because the immense amount of expansion during and since the inflationary epoch magnifies these small length scales to ones of astrophysical size. If the tensor perturbations are detectable, we are already probing physics at these length scales via the scalar perturbations, but we cannot know this until the tensor perturbations are actually detected.

To date we only have upper limits on the amplitude of tensor perturbations, upper limits that are as strong as they can be from measuring temperature anisotropies. To detect tensor perturbations we need to improve measurements of CMB polarization. In the tensor sector, CMB-S4 will improve current constraints by almost two orders of magnitude. This is especially interesting because it allows this next-generation instrument to reach theoretically well-motivated thresholds for the tensor-to-scalar ratio (the ratio of power in tensor modes to power in scalar modes), which consequently serves as the primary inflationarFinn:2001qiy science driver for the design.

It is worth pointing out explicitly that these tensor perturbations are by definition gravitational waves. With the recent LIGO detections [1, 1], we have entered the era of gravitational wave astronomy, and with CMB-S4 we will be targeting the ultra long-wavelength end of nature’s gravitational wave spectrum. We expect that the background of inflation-produced gravitational waves to be nearly scale-invariant, and that waves with frequencies 10^8 (TBD) times higher than those detectable with CMB-S4 may one day be detectable with a space-based observatory, greatly enhancing the value of any CMB-S4 detection.

Inflation predicts B-mode fluctuations sourced by primordial gravitational waves. But more generally, the B-mode signal carries information about both the spectrum of primordial perturbations in the tensor (and vector) components of the metric and any physics that affected the evolution of those modes once they re-entered the horizon. Furthermore, the inflationary sector is not isolated from the rest of particle physics at high energies. In the context of specific proposals for physics beyond the standard model, including dark matter models, a detection of B-modes can have consequences for predictions for the post-inflation spectrum of particles and their thermodynamics. These models may also provide observables other than the amplitude of B-modes that constrain inflationary physics. The rich interplay of inflation models and other physics beyond the standard model is discussed in detail in Section 5.

A detection of primordial gravitational waves would open a completely new window on the physical processes of the early universe and reveal a new length scale of particle physics far below those accessible with terrestrial particle colliders.

If the overall amplitude of the B-mode signal is large enough to be detected at high significance by the CMB-S4 instrument, we will be able to begin further characterizing the statistics of the perturbations. Investigating the scale-dependence of the amplitude of tensor fluctuations and their Gaussianity will allow us to determine if the signal is consistent with the amplification of quantum vacuum fluctuations of the metric during inflation. If CMB-S4 measurements are consistent with a nearly scale-invariant and a weakly non-Gaussian spectrum, a detection would

- Identify the energy scale of inflation.
- Provide strong evidence that gravity is quantized, at least at the linear level.

- Provide strong evidence that the complete theory of quantum gravity must accommodate a Planckian field range for the inflaton.

Departures from a nearly scale-invariant, Gaussian spectrum would reveal new physics beyond the simplest inflationary models. Existing models propose some examples of predictions from a richer inflationary or post-inflationary sector, and these predictions would be tested. However, given the lack of observational constraints on physics at such high energy scales there is also enormous discovery potential. Polarization data also provides consistency checks on the current dominant theoretical framework, including model-independent constraints on the graviton mass and constraints on alternatives to inflation.

In the absence of a detection, CMB-S4 would put some of the most significant constraints on inflation models to date, ruling out large classes models. **What else should be stated here?**

Insert paragraph emphasizing inflation-related gains from the scalar sector with CMB-S4.

In Section 2.2 we provide a basic introduction of the inflationary paradigm in its simplest form. In Section 2.4 we review in detail what a detection of primordial gravitational waves would mean and what follow-up measurements should or could be done to further characterize any signal. Section 2.5 explains the implications of a robust upper limit of $r < 0.001$. Section 2.3 lays out what is required to achieve that goal. The final two sections describe the significant gains CMB-S4 will allow in constraining other aspects of the primordial universe, both standard and more speculative. These include characterizing the scalar power spectrum, constraining curvature, non-Gaussianity, isocurvature modes, further probes of CMB ‘anomalies’ and tests/constraints of cosmic strings.

2.2 Basics of cosmological inflation

In this section, we introduce the basic concepts of cosmological inflation. We do this in two stages, first giving a broad outline of the important concepts then proceeding to a more technical definition.

2.2.1 Inflation basics I: A heuristic picture

Inflation is, by definition, a period of accelerating expansion. As explained in Fig. 4, an accelerating universe has a causal structure very different from that of a decelerating universe. In a decelerating universe, a pair of separated comoving particles evolves from being causally disconnected – in which case the particles, unable to influence each other, are said to be superhorizon – to being causally connected, or subhorizon. In an accelerating universe, the opposite occurs. In the inflationary scenario, the universe undergoes an accelerating stage, which is followed by a long period of deceleration.

In view of the early period of accelerating expansion, two separated regions in the universe that are now causally disconnected could have been able to interact with each other during the inflationary epoch. Causally connected perturbations in those two regions – for example, an underdensity in one and an overdensity in the other – could thus have been created at very early times. Quantum mechanics provides a mechanism for generating such perturbations, and in fact makes them unavoidable. Quantum mechanical fluctuations initially created with subnuclear wavelengths are stretched by the cosmic expansion to millimeter length scales within a tiny fraction of a second; at present they are astrophysically large. Thus observations of cosmic structure give us an opportunity to probe physics on extremely small length scales.

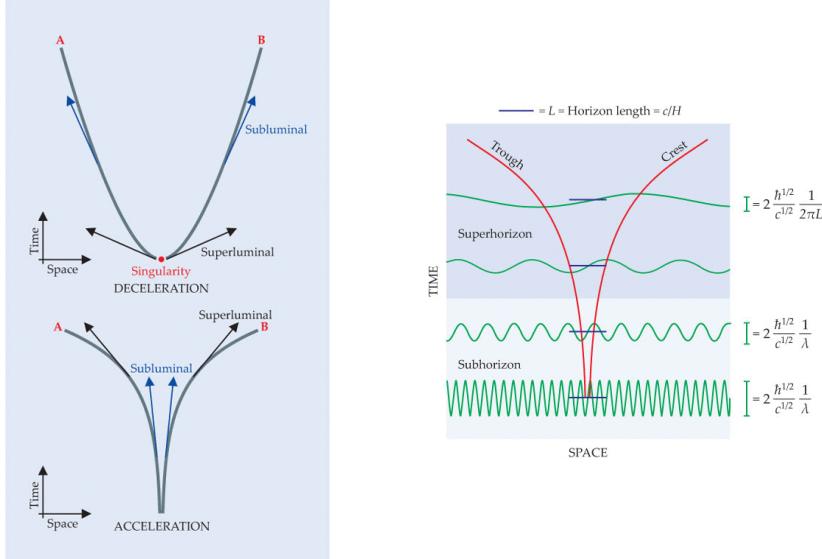


Figure 4. **Left panel:** In an expanding universe, the distance between two separated points increases over time, simply due to the expansion of the space between them. The two panels here show the spacetime trajectories of two comoving points, A and B. For the decelerating expansion illustrated in the top panel, the separation rate is greater in the past and even exceeds the speed of light at sufficiently early time. Thus A and B go from being out of causal contact – unable to influence each other – to being in causal contact. In an accelerating universe, the separation rate is smaller in the past; the two points go from being in causal contact to being out of causal contact. In the inflationary universe scenario, an early epoch of acceleration – the inflationary era – smoothly maps onto a long period of deceleration. Thus two points can go from being in causal contact to out of causal contact and, much later, back into causal contact. **Right panel:** Fluctuations in the value of the inflaton field, which is responsible for the accelerating expansion of the cosmos, evolve differently, depending on whether their wavelength λ is less than or greater than the horizon length $L = c/H$. When $\lambda \ll L$, the uncertainty principle limits how smooth the field can be. As a result, the amplitude of the fluctuation is inversely proportional to λ and thus decreases as the universe expands. (The influence of the uncertainty principle is reflected by the appearance of Planck's constant \hbar in the expression for the amplitude.) As λ becomes larger than the horizon, the crest and trough of the wave cease to be in causal contact, so the amplitude stops evolving. For superhorizon evolution, its asymptotic value corresponds to replacing the wavelength in the subhorizon case with $2\pi L$. Eventually, cosmic expansion stretches the fluctuations to astrophysically large length scales.

Accelerating expansion requires the universe to have an energy density that dilutes relatively slowly with expansion. In inflationary models, such an energy density is usually obtained via the introduction of a new field ϕ , called the inflaton field with Lagrangian density, in the simplest cases, given by

$$\mathcal{L} = \frac{1}{2} \partial_\mu \phi \partial^\mu \phi - V(\phi) \quad (2.1)$$

where $V(\phi)$ is a potential energy density.

A generic inflaton field configuration will not lead to inflation. But if there is a large enough patch of space in which ϕ takes values for which the potential is sufficiently flat, ϕ will rapidly evolve to satisfy the “slow-roll condition” $\frac{1}{2} (d\phi/dt)^2 \ll V(\phi)$. When both the spatial and temporal derivatives of the inflaton field are small, $V(\phi)$ is nearly constant in time and makes the dominant contribution to the energy density. Under such conditions, and given the Friedmann equation $\dot{a}/a \propto \sqrt{\rho}$, the patch inflates. In the limit that the energy

density is completely constant in time, the scale factor grows as e^{Ht} , and points separated by more than c/H are causally disconnected.

A standard assumption in the calculation of inflationary perturbation spectra is that the field is as smooth as it possibly can be, and still be consistent with the uncertainty principle. As Fig. 4 shows, these fluctuations will be stretched to astrophysically large length scales by cosmic expansion. In an inflationary scenario quantum fluctuations provide the initial seeds of all structure in the universe.

As ϕ rolls toward the potential minimum, $V(\phi)$ eventually becomes smaller than $\frac{1}{2}(d\phi/dt)^2$; the slow-roll condition is no longer met, and inflation ends. Decays of the inflaton to other particles – irrelevant during inflation because the decay products were quickly diluted by expansion – then become important. The remaining energy in the ϕ field converts to a thermal bath of the particles of the standard model, and perhaps other particles as well.

The small but nonzero spatial fluctuations in ϕ cause inflation to end at different times in different locations. In those regions where inflation ends relatively early, the mass density is lower due to the extra expansion that the region has undergone since the end of inflation. Thus the slightly different expansion histories of different locations result in density differences; those small density perturbations eventually grow under the influence of gravity to create all the structures we observe in the universe today.

The spacetime metric itself, at least in a linearized treatment, presumably obeys the uncertainty principle as well. As a result, we expect a nearly scale-invariant spectrum of gravitational waves to be produced during inflation as well. Just as with fluctuations of the inflaton field, they obey an uncertainty principle and, in the course of superluminal expansion, have their amplitude set to a value proportional to the Hubble parameter H during inflation. Detecting the influence of that gravitational-wave background on the CMB would allow cosmologists to infer H and hence the energy scale of the inflationary potential; observations of density perturbations, by contrast, provide a relatively indirect look at the inflationary era. As emphasized in the previous section, CMB-S4 is poised to detect, or place interesting upper limits, on the amplitude and spectrum of inflation-produced gravitational waves via their signature in B-mode polarization.

2.2.2 Inflation basics II: Quantifying the predictions

The overall evolution of the universe is well modeled by a Friedmann-Lemaître-Robertson-Walker line element

$$ds^2 = -dt^2 + a^2(t) \left[\frac{dr^2}{1 - kr^2} + r^2 d\Omega^2 \right], \quad (2.2)$$

where $k = \pm 1$ allows for spatial curvature and the time evolution is specified by the scale factor, $a(t)$. The Hubble parameter, $H = \dot{a}/a$, gives the rate of expansion of the universe. A period of inflation will drive the spatial curvature close to zero, in good agreement with current observations. We will assume spatial flatness and set $k = 0$ for most of the discussion, but see Section 2.9 for a discussion of CMB-S4 constraints on the curvature.

Since the period of cosmic inflation must end, there must exist a clock, or scalar degree of freedom. According to the uncertainty principle this clock must fluctuate, generating density perturbations that are adiabatic. In the most economic scenarios, these density perturbations are the seeds that grow into the anisotropies observed in the cosmic microwave background radiation and the stars and galaxies around us. Other degrees of freedom could, of course, also be present during this phase and might even be responsible for the generation of density perturbations we observe.

LK: I think this paragraph should probably get moved elsewhere – perhaps to a discussion of inflation falsifiability that we need to include somewhere. We mention an alternative to inflation here. Observations of the cosmic microwave background provide strong evidence for a period preceding the hot big bang during which the co-moving Hubble radius, $(a|H|)^{-1}$, was decreasing with time: the measured average CMB temperature and the statistics of the measured anisotropies are the same over regions that otherwise share no causal history. Inflation is one way of arranging for a period of decreasing co-moving Hubble radius, as an expanding and accelerating universe will have such a Hubble radius. Alternatively, the phase of decreasing co-moving Hubble radius could have occurred during a period of decelerating contraction which must then be followed by a bounce as in the ekpyrotic or matter bounce scenarios [2, 3, 4, 5, 6, 7, 8].

For these early times, the ADM formalism provides a convenient parametrization of the line element

$$\begin{aligned} ds^2 &= -N^2 dt^2 + h_{ij}(dx^i + N^i dt)(dx^j + N^j dt) \\ h_{ij} &= a^2(t)[e^{2\zeta}\delta_{ij} + \gamma_{ij}]. \end{aligned} \quad (2.3)$$

The equations of motion for N (the lapse) and N^i (the shift) are the Hamiltonian and momentum constraints, while ζ (equivalent to $-\mathcal{R}$ in the *Planck* collaboration papers) and γ_{ij} contain the dynamical scalar and tensor degrees of freedom. In scenarios with matter sources other than a scalar field there may also be vector perturbations. These rapidly decay and can be neglected unless they are actively sourced in the post-inflationary universe, e.g. by cosmic strings.

Because the equations of motion are invariant under translations and the perturbations are linear or nearly so, it is convenient to work with the Fourier transforms

$$\zeta(t, \vec{x}) = \int \frac{d^3 k}{(2\pi)^3} \zeta(t, \vec{k}) e^{i\vec{k} \cdot \vec{x}} + h.c. \quad \text{and} \quad \gamma_{ij}(t, \vec{x}) = \sum_s \int \frac{d^3 k}{(2\pi)^3} \gamma_s(t, \vec{k}) e_{ij}(\vec{k}, s) e^{i\vec{k} \cdot \vec{x}} + h.c., \quad (2.4)$$

where $e_{ij}(\vec{k}, s)$ is the transverse traceless polarization tensor for the graviton. The solutions oscillate when the modes are deep inside the horizon, $k \gg aH$. By definition, the modes exit the horizon when $k = aH$ and in single-field models approach a constant outside the horizon when $k \ll aH$.

The statistical properties of the scalar and tensor fluctuations, ζ and γ_s , at times sufficiently late so that they have frozen out provide the link between the primordial era and the observed CMB today as well as other probes of the structure of the late universe. For a universe that is statistically homogeneous and isotropic and in which the primordial fluctuations are Gaussian, the information about the statistical properties is contained in the two-point correlation functions

$$\begin{aligned} \langle \zeta(\vec{k}) \zeta(\vec{k}') \rangle &= (2\pi)^3 \delta^3(\vec{k} + \vec{k}') \frac{2\pi^2}{k^3} \mathcal{P}_\zeta(k) \\ \langle \gamma_s(\vec{k}) \gamma_{s'}(\vec{k}') \rangle &= (2\pi)^3 \delta_{ss'} \delta^3(\vec{k} + \vec{k}') \frac{2\pi^2}{k^3} \frac{1}{2} \mathcal{P}_t(k), \end{aligned} \quad (2.5)$$

where the factor of $1/2$ in the second to last line accounts for the fact that the measured power includes contributions from each of the two graviton polarizations. In single field slow-roll inflation, the gauge invariant combination of metric and scalar field fluctuations that is conserved outside the horizon has the power spectrum

$$\mathcal{P}_\zeta(k) = \frac{1}{2\epsilon M_p^2} \left(\frac{H}{2\pi} \right)^2 \Bigg|_{k=aH}, \quad (2.6)$$

where $\epsilon = -\dot{H}/H^2$ is the first slow-roll parameter, and $M_p = 1/\sqrt{8\pi G}$ is the reduced Planck mass. As indicated, the Hubble parameter and ϵ are to be evaluated at horizon exit when the wavenumber k is equal to the inverse comoving Hubble radius. In the absence of additional sources, the tensor power spectrum generated by inflation is

$$\mathcal{P}_t(k) = \frac{8}{M_p^2} \left(\frac{H}{2\pi} \right)^2 \Big|_{k=aH}. \quad (2.7)$$

It is convenient to introduce the logarithmic derivatives of these power spectra

$$n_s(k) - 1 \equiv \frac{d \ln \mathcal{P}_\zeta}{d \ln k} \quad \text{and} \quad n_t(k) \equiv \frac{d \ln \mathcal{P}_t}{d \ln k}. \quad (2.8)$$

If the Hubble rate and slow-roll parameter only weakly depend on time as in slow-roll inflation, these will be $n_s(k) \approx 1$ and $n_t(k) \approx 0$ and can be expanded around a pivot scale k_* accessible by the CMB

$$n_s(k) - 1 = n_s - 1 + \frac{dn_s(k)}{d \ln k} \Big|_{k_*} \ln(k/k_*) + \dots \quad \text{and} \quad n_t(k) = n_t + \frac{dn_t(k)}{d \ln k} \Big|_{k_*} \ln(k/k_*) + \dots \quad (2.9)$$

In this approximation, the power spectra are

$$\begin{aligned} \mathcal{P}_\zeta(k) &= A_s \left(\frac{k}{k_*} \right)^{n_s - 1 + \frac{1}{2} \frac{dn_s}{d \ln k} \Big|_{k=k_*} \ln(k/k_*) + \dots}, \\ \mathcal{P}_t(k) &= A_t \left(\frac{k}{k_*} \right)^{n_t + \frac{1}{2} \frac{dn_t}{d \ln k} \Big|_{k=k_*} \ln(k/k_*) + \dots}, \end{aligned} \quad (2.10)$$

where A_s , A_t are the scalar and tensor amplitudes, and n_s and n_t , are the scalar and tensor spectral index, respectively, both at the pivot scale. The tensor-to-scalar ratio, r , is the relative power in the two types of fluctuations at a chosen pivot scale k_* accessible by the CMB

$$r = \frac{A_t}{A_s}. \quad (2.11)$$

The power spectra of ζ and γ_s are time-independent as long as the modes are outside the horizon, and only begin to evolve once the modes of interest re-enter the horizon at late times. In particular, they set the initial conditions for the system of equations governing the time evolution of the universe from a temperature of around 10^9 K (when electrons and positrons have just annihilated) to the present. To exhibit the link between the primordial perturbations and late time observables explicitly, note that in a spatially flat universe, the contributions of primordial scalar perturbations to the angular power spectra of temperature or E-mode anisotropies are given by

$$C_{XX,\ell}^{(S)} = \int \frac{dk}{k} \mathcal{P}_\zeta(k) \left| \int_0^{\tau_0} d\tau S_X^{(S)}(k, \tau) j_\ell(k(\tau_0 - \tau)) \right|^2, \quad (2.12)$$

where j_ℓ is a spherical Bessel function that encodes the (spatially flat) geometry of the universe and $S_X^{(S)}(k, \tau)$ with $X = T, E$ are source functions that encode the evolution of the modes in the hot big bang universe (in particular, the physics of recombination is very important for B-modes). At linear order, scalar perturbations only contribute to angular power spectra of temperature and E-mode polarization and the cross-spectrum of temperature and E-mode polarization, while the tensor perturbations in addition generate

B-mode polarization. The primordial contribution of the tensor perturbations to the angular power spectrum of B-modes is

$$C_{BB,\ell} = \int \frac{dk}{k} \mathcal{P}_t(k) \left| \int_0^{\tau_0} d\tau S_B^{(T)}(k, \tau) j_\ell(k(\tau_0 - \tau)) \right|^2. \quad (2.13)$$

where $S_B^{(T)}(k, \tau)$ is the appropriate source function.

The results of calculations using Eqs. (2.12) and (2.13), performed with the Code for Anisotropies in the Microwave Background (CAMB, [9]), are shown in Fig. 5. Results for the total (scalar + tensor) temperature and E-mode spectra are given by the black and red lines, respectively, while the result for the tensor B-mode spectrum is given by the blue lines for two possible values of r . Also shown are predictions for contaminations to the primordial B-mode spectrum by gravitational lensing of E modes (green line) and foregrounds (brown dashed lines). For more discussion of B-mode signal shape and detection prospects, as well as contamination mitigation, see Sec. 2.3. Finally, the black points in Fig. 5 show current constraints on the total (primordial + contamination) B modes from BICEP2/Keck [10] and SPTpol [11].

At present, bounds on the tensor contribution to the temperature and E-mode anisotropies are comparable to constraints on the tensor-to-scalar ratio from B-mode observations. The former constraints are now cosmic variance-limited. There is no limit on the latter from cosmic variance, and improvements and a potential detection with CMB-S4 will rely on measurements of B-mode polarization, most likely targeting the degree-scale feature in the primordial B-mode spectrum (which is easier to access from the ground than the much larger-scale $\ell < 10$ feature).

Constraints on the amplitude of primordial tensor modes already strongly disfavor once popular inflationary models like minimally coupled chaotic inflation with a quadratic potential. In the next section we will discuss in detail what a detection of primordial gravitational waves would imply for theories of the primordial universe.

2.3 Forecasted sensitivity to r

Achieving the CMB-S4 target sensitivity to r ($\sigma_r \sim 10^{-3}$) will require exquisite measurements of the B-mode power spectrum. It is expected that CMB-S4 will target the degree-scale recombination feature rather than the tens-of-degree-scale reionization feature (see Fig. 5), because these largest scales are difficult to access from the ground due to atmosphere and sidelobe pickup (though some Stage-3 ground-based experiments are attempting this measurement, notably CLASS [12]).

As can be seen from Fig. 5, the first requirement for this level of sensitivity to r is a large leap forward in raw instrument sensitivity. As detailed in Sec. 8.2.1, for ground-based bolometric detectors, this means a large increase in detector count. The forecasts in this section use a baseline of 250,000 detectors operating for four years (or 10^6 detector years), dedicated solely to maximizing sensitivity to r . It will be necessary to split this total effort among many electromagnetic frequencies, to separate the CMB from polarized Galactic foregrounds. The forecasts here assume eight frequency bands, ranging from 30 to 270 GHz. Contamination from gravitationally lensed E modes must also be mitigated. While a precise prediction for the cosmological mean of the lensing B-mode power spectrum can be made and subtracted from the observed spectrum, there will be a sample variance residual between this prediction and the real lensing B modes on a particular patch of sky. To suppress this sample variance, it will be necessary to “de-lens” the B-mode maps with a prediction for the lensing signal from that particular patch of sky, constructed from the E-mode map and some tracer of the lensing potential (see Sec. 6.4 for details). Forecasts in this section assign part of the

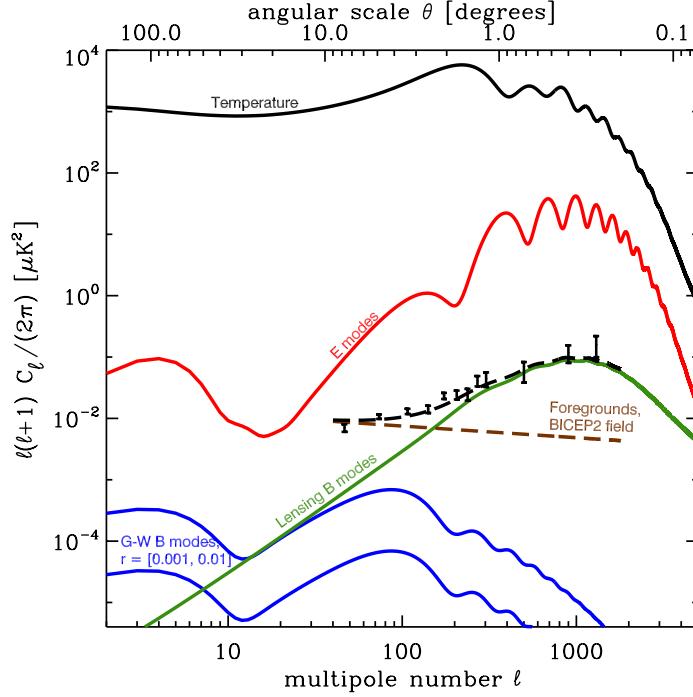


Figure 5. Theoretical predictions for the temperature (black), E-mode (red), and tensor B-mode (blue) power spectra. Primordial B-mode spectra are shown for two representative values of the tensor-to-scalar ratio: $r = 0.001$ and $r = 0.01$. The contribution to tensor B modes from scattering at recombination peaks at $\ell \sim 80$ and from reionization at $\ell < 10$. Also shown are expected values for the contamination to primordial B modes from gravitationally lensed E modes (green) and foregrounds for one small patch of sky (brown dashed). Current constraints on the total (primordial + lensing + foreground) B-mode spectrum are shown as black points, and sum of lensing + foregrounds is shown with a black dashed line. The lensing contribution to the B-mode spectrum can be at least partially removed from measurements by exploiting the non-Gaussian statistics of the lensing, while the foreground contamination can be mitigated by a multi-frequency measurement (for details see Sec. 2.3).

total detector count to a dedicated de-lensing effort, assumed to be a large-aperture (≥ 6 -meter) telescope at a single frequency. (See below for a discussion of assumptions about aperture size.) Finally, from the relative amplitudes of the temperature, E-mode, and B-mode power spectra, it is clear that instrumental systematics that mix temperature or E-mode power into B modes must also be exquisitely controlled. To account for the potential hit in raw sensitivity taken by guarding against systematics, and to generally tie the Fisher-matrix, power-spectrum-based forecasts here with real-world results, the forecasts here use scaled versions of achieved power-spectrum covariance matrices from the BICEP2/Keck experiments. This conservative assumption regarding systematics mitigation and overall efficiency necessarily results in worse constraints on r than a naive, raw-sensitivity calculation would imply. Further details of the forecasting methodology, including assumptions regarding foreground properties and de-lensing efficiency, can be found in Section 7.10.1.1.

The trade-off between raw sensitivity, ability to remove foregrounds, and ability to de-lens results in a complicated optimization with respect to sky coverage. Figure 6 shows the forecasted sensitivity to r for

CMB-S4 as a function of the observed sky fraction under an upper limits scenario ($r = 0$, left) or a detection scenario ($r = 0.01$, right). Focusing on the $r = 0$ case, we see that an effort devoted to an initial detection of r will benefit from a deep survey that targets a small sky area. Concerns about foreground complexity beyond what has been simulated will increase our preference for a small area survey that focuses on the cleanest regions. Likewise, our ability to identify and address instrumental systematics is often limited by the noise level of the maps, so deeper maps can serve as a guard against instrumental problems (Comment: Would be helpful to refine this argument with a simple model.).

The preference for small sky area is in tension with other CMB-S4 science goals that prefer large sky areas but have much lower requirements for foreground cleaning. To balance these goals, we assume that roughly one half of total CMB-S4 experiment is devoted to a deep survey targeting degree-scale B modes while the other half is spent on a broad survey. Figure 7 shows the forecasted sensitivity to r as a function of the total effort spent on the deep survey. With 250,000 detectors operating for four years, CMB-S4 will significantly exceed the $r = 0.001$ benchmark. If r is as large as 0.01, then larger sky area will be needed to improve precision; the CMB-S4 deep survey must be designed with the flexibility to increase sky area in the event of a detection.

In these forecasts, it is assumed that the degree-scale CMB and foregrounds are measured using small aperture telescopes, while the de-lensing is achieved with a separate, large-aperture telescope. In all combinations of sky fraction and total detector effort, these forecasts indicate that at least 25–50% of the total effort must be spent on high-resolution maps that can be used for de-lensing. If the large-aperture, high-resolution data can also be used for degree-scale science, the allocation of resources between bands and telescopes would change slightly, and the overall constraints on r would improve. If, furthermore, cost per detector were independent of telescope aperture size, an argument could be made to do all the science with large-aperture telescopes. Neither of these assumptions is clearly supported by current data, however. Finally, we note that for the deepest option, targeting just 1% of the sky, achieving the forecasted sensitivity to r will require an 80% reduction in the map *r.m.s.* level of the CMB lensing B modes. While this is achievable from a sensitivity standpoint (see Sec. 6.4), systematics and foregrounds will need to be carefully considered.

2.4 Implications of a detection of primordial gravitational waves

The inflationary amplification of vacuum fluctuations of the metric leads to a nearly scale-invariant, very nearly Gaussian tensor power spectrum. This signal is very well characterized by a single parameter defining the (relative) amplitude of tensor fluctuations, r . In this section, we consider the consequences of a detection of primordial gravitational waves consistent with this simplest inflationary expectation. Together with implications of an improved upper limit on r presented in the next section, these expectations motivate the threshold level of sensitivity CMB S4 will aspire to and guide the baseline proposals for the instrument in Section 2.3.

Of course, if a detection is achieved it will be essential to characterize the accuracy with which we can test the standard inflationary prediction. In Section 2.6 we will use the baseline instrument design from Section 2.3 with hypothetical detection levels to forecast constraints on n_T and tensor mode non-Gaussianity. We also use Section 2.6 to discuss alternatives to the inflationary vacuum prediction, including non-vacuum sources during inflation. We will see that non-vacuum scenarios would be distinguishable as long as r is detected at high significance.

The remainder of this section derives the remarkable implications of a detection of primordial gravitational waves with amplitude accessible by CMB S4, and with a nearly scale-invariant, nearly Gaussian spectrum: revealing the energy scale of inflation, providing excellent evidence for linearized quantum gravity, and very

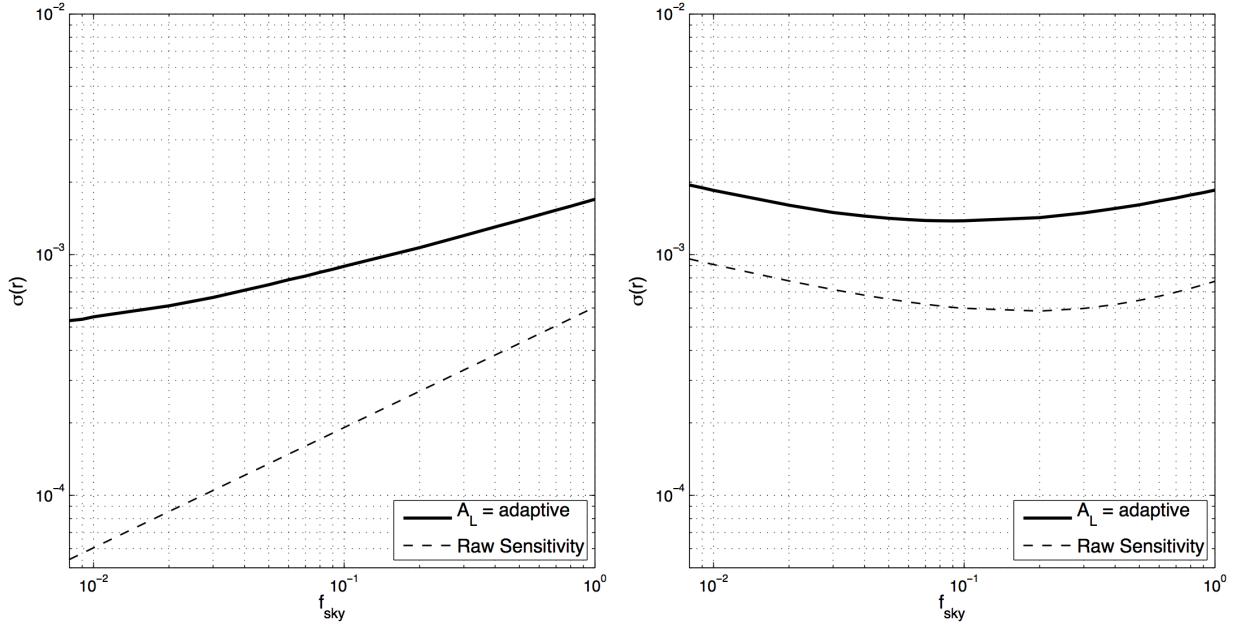


Figure 6. Forecasted uncertainty on r , as a function of f_{sky} , for an effort of 10^6 detector-years (150 equivalent), assuming $r = 0$ (left panel) and $r = 0.01$ (right panel). The forecasting procedure is specifically targeted towards optimizing tensor-to-scalar parameter constraints in the presence of Galactic foregrounds and gravitational lensing of the CMB. The optimization assumes an amount of achieved de-lensing that varies with f_{sky} ; temporary note: for a detailed description of the forecasting, please see: [this posting](#).

strong evidence for structure in non-linear quantum gravity that accommodates a large field range for the inflaton.

2.4.1 The energy scale of inflation

According to the inflationary prediction for the amplitude of primordial gravitational waves, Eq. (2.7), a detection provides a direct measurement of the Hubble scale during inflation. In single field slow-roll models the Friedmann equation relates the Hubble scale to the potential energy V of the inflaton, $3H^2M_p^2 \approx V$. The inflationary prediction for the amplitude of scalar fluctuations, Eq. (2.6) can be used to write H in terms of the measured amplitude and the so far undetermined slow-roll parameter, ϵ . Since ϵ is directly proportional to the tensor-to-scalar ratio this allows us to express the energy scale of inflation in terms of measured numbers, known constants, and r (all at the pivot scale $k_* = 0.05 \text{ Mpc}^{-1}$)

$$V^{1/4} = 1.04 \times 10^{16} \text{ GeV} \left(\frac{r_*}{0.01} \right)^{1/4}. \quad (2.14)$$

A detection of primordial gravitational waves therefore determines the energy scale of inflation to within a few per cent.

A detection of primordial gravitational waves by CMB-S4 would reveal a new scale of particle physics near the GUT scale. In the event the signal is reasonably scale-invariant and at most weakly non-Gaussian, this scale corresponds to the energy scale of inflation.

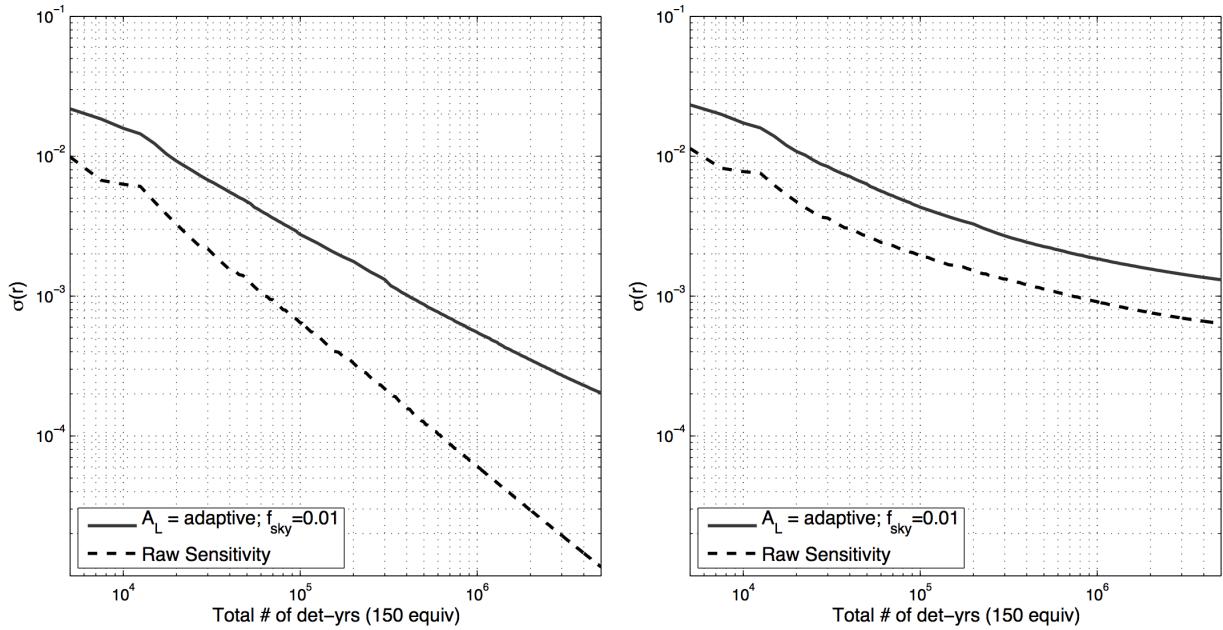


Figure 7. Forecasted uncertainty on r , as a function of effort, for a fixed $f_{\text{sky}} = 0.01$, assuming $r = 0$ (left panel) and $r = 0.01$ (right panel). The forecasting procedure is specifically targeted towards optimizing tensor-to-scalar parameter constraints in the presence of Galactic foregrounds and gravitational lensing of the CMB. The optimization assumes an amount of achieved de-lensing that varies with f_{sky} ; temporary note: for a detailed description of the forecasting, please see: this posting.

2.4.2 Planckian field ranges and symmetries

The spectrum of tensor fluctuations depends only on the Hubble parameter H during inflation, while the scalar power depends on both H and the evolution of the homogeneous field sourcing inflation. As a consequence, the tensor-to-scalar ratio r determines the inflaton field range in Planck units [13]

$$\frac{\Delta\phi}{M_p} = \int_0^{\mathcal{N}_*} d\mathcal{N} \left(\frac{r}{8}\right)^{1/2}, \quad (2.15)$$

where (applying the general equation to the observationally accessible regime) \mathcal{N}_* is the number of e-folds between the end of inflation and the moment when the mode with $k_* = 0.05 \text{ Mpc}^{-1}$ corresponding to the CMB pivot scale exits the horizon. In many common inflationary models r is a monotonic function of \mathcal{N} so that

$$\frac{\Delta\phi}{M_p} \gtrsim \left(\frac{r_*}{8}\right)^{1/2} \mathcal{N}_* \gtrsim \left(\frac{r}{0.01}\right)^{1/2}. \quad (2.16)$$

The value of \mathcal{N}_* is not well constrained and depends on unknown details of reheating, but $\mathcal{N}_* \gtrsim 30$ provides a conservative lower limit, justifying the second inequality in equation (2.16). Thus, a tensor-to-scalar ratio $r > 10^{-2}$ typically corresponds to a trans-Planckian excursion in field space between the end of inflation and the epoch when the modes we observe in the CMB exit the horizon.

While it is a familiar (if still extraordinary) feature of inflation that the fluctuations are quantum mechanical in origin, the relation in Eq.(2.16) is significant because the inflationary framework also uses quantum field theory to describe the field sourcing the *background* accelerated expansion. The action for the inflaton should

be under good quantum control over the entire field range. The Lyth bound, Eq.(2.16), naturally puts the field range in units of the highest known, fundamental energy scale. If the field range is close to Planckian, the inflationary model requires a special feature in the quantum theory at and above the scale of quantum gravity.

To understand why, recall that unless we work in a UV complete theory such as string theory, we rely on an effective field theory description of the inflationary epoch. General relativity viewed as an effective field theory breaks down as energies approach the Planck scale because interactions between gravitons become strongly coupled. The same is true for matter coupled to general relativity, so that the effective field theory governing the inflationary period will generically have a sub-Planckian cut-off $\Lambda_{\text{UV}} < M_p$. In fact, in any weakly coupled UV completion of general relativity the new degrees of freedom must enter well below the Planck scale to ensure weak coupling so that $\Lambda_{\text{UV}} \ll M_p$. Although we do not know the complete theory of quantum gravity, our understanding of the implications of the Lyth bound are based on experience with effective field theories at much lower scales, borne out in relation to quantum gravity through models of inflation in string theory.

According to the bound in Eq.(2.16), a tensor-to-scalar ratio $r > 10^{-2}$ (and even somewhat smaller) requires a displacement in field space that is larger than the cut-off of the effective field theory. While this does not invalidate an effective field theory description, it has important consequences. Assuming the UV complete theory is known, the effective field theory is obtained by integrating out all modes parametrically heavier than the cut-off Λ_{UV} of the single-field model. In the absence of symmetries, we expect the inflaton ϕ to couple to heavy degrees of freedom χ that, once integrated out, will introduce significant structure in the potential for the inflation on scales $\Delta\phi \ll \Lambda_{\text{UV}}$. For example, consider the action

$$S = \int d^4x \sqrt{-g} \left[-\frac{1}{2}g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - \frac{1}{2}g^{\mu\nu}\partial_\mu\chi\partial_\nu\chi - \frac{1}{2}m^2\phi^2 - \frac{1}{2}M^2\chi^2 - \frac{1}{2}\mu\phi\chi^2 + \dots \right]. \quad (2.17)$$

By assumption, the mass of the heavy degrees of freedom to be integrated out is $M \gtrsim \Lambda_{\text{UV}}$, and the dots represent various other interaction terms. Generically the dimensionful coupling μ is also expected to be of order the cut-off, $\mu \sim \Lambda_{\text{UV}}$. From the last two terms in equation (2.17), we see that displacements of ϕ by a distance comparable to the cut-off may lead to cancellations in the effective mass of the heavy degrees of freedom, and heavy states, in this case χ , may become light if ϕ is displaced by a distance large compared to the cut-off. In particular, since $\Lambda_{\text{UV}} < M_p$ we should not expect potentials that are smooth over super-Planckian distances in a generic low energy effective field theory with cut-off $\Lambda_{\text{UV}} < M_p$.

We can only expect potentials suitable for large-field inflation if some mass scales, in the example m and μ are well below the cut-off, or if dimensionless couplings are small. This occurs naturally if the UV theory respects a weakly broken shift symmetry $\phi \rightarrow \phi + c$ that ensures that quantum corrections from the inflaton and graviton will not introduce large corrections to the inflationary Lagrangian [14, 15, 16, 17, 18]. At the level of an effective field theory we can simply postulate such an approximate shift symmetry, but one should keep in mind that we ultimately require the existence of such a symmetry in quantum gravity.

As the best developed theory of quantum gravity, string theory is a useful framework for exploring mechanisms that allow large-field inflation to be realized even in the presence of heavy degrees of freedom. Axions are ubiquitous in string theory and provide natural candidates for the inflaton because they enjoy a shift symmetry that is weakly broken by instanton effects as well as the presence of branes or fluxes [19]. Early field theory models relied on the familiar periodic contributions to the potential generated by instantons to drive inflation [20, 21]. In string theory the periods are expected to be sub-Planckian [22, 23], while constraints on the scalar spectral index require super-Planckian axion periods so that a UV completion of these models does not currently exist. The excitement over the BICEP2 results led to renewed interest in models in which the inflaton is an axion with a potential that is entirely due to instanton effects and

intensified the discussion to what extent some means to achieve large field inflation via multiple axions may be incompatible with basic principles of quantum gravity [24, 25, 26, 27, 28, 29, 30, 31, 32, 33].

In addition to the familiar non-perturbative contributions that break the continuous shift symmetry to a discrete one, the presence of fluxes and branes lead to contributions to the axion potentials that break the discrete shift symmetry as well. As the axion is displaced by one period, one unit of charge is induced, so that the axion field space becomes non-compact. As a consequence, super-Planckian decay constants are not required for super-Planckian excursions in these monodromy models [34, 35, 36, 37, 38, 39, 40, 41, 42]. Generically both contributions to the potential are present and these models predict periodic effects at some level, either directly from the periodic features in the potential or from periodic bursts of string or particle production. Unfortunately, the strength of the signal is very model dependent, and a detection of these effects with CMB-S4 is not guaranteed.

In writing (2.16), we have assumed that r is monotonic, or at least of the same order of magnitude throughout the inflationary period. One can easily construct models in which r is non-monotonic to weaken the bound [43, 44, 45]. In the case of a detection with CMB-S4 of a spectrum that is at least approximately scale-invariant, we can write the weaker bound

$$\frac{\Delta\phi}{M_p} \gtrsim \left(\frac{r}{0.3}\right)^{1/2}, \quad (2.18)$$

which bounds the distance in field space traveled during the time the modes we observe in the CMB exited the horizon. This inequality implies that even if the distance in field space traveled during this period is sub-Planckian, it is not parametrically smaller than M_p . Because general relativity is not UV complete and becomes strongly coupled at M_p , any weakly coupled UV completion will come with a scale of new physics M , e.g. the string scale, that must be parametrically smaller than the Planck scale to ensure weak coupling. This implies that we cannot avoid the question of the embedding of the inflation model into quantum gravity for $r = 0.01$ or even for $r = 0.005$ unless we assume the UV completion of general relativity is strongly coupled.

In deriving the primordial power spectra and Eq.(2.15), we have assumed the Bunch-Davies state. The relation between r and the scale of inflation is modified if we assume that the tensor modes (and the scalar modes) either do not start in the Bunch-Davies state [46, 47], or that the evolution during inflation will lead to departures from it. The first option generically introduces a stronger scale-dependence into the tensor spectrum [48, 49] (and additional non-Gaussianity). In addition, this way of achieving observable primordial B -modes from a low-scale model has a similar feature to large-field models: one should show that the initial state is not only acceptable from the point of view of low energy considerations, but can be generated by pre-inflationary physics. The second option, discussed in section 2.4.1, leads to non-trivial higher n -point functions that are in principle measurable.

In summary, a conclusive detection of primordial B -modes with CMB-S4 would provide evidence that the theory of quantum gravity must accommodate a Planckian field range for the inflaton. Conversely, the absence of a detection of B -modes with CMB-S4 will mean that a large field range is not required.

A detection of r , together with high confidence that the gravitational waves are predominantly due to vacuum fluctuations, would provide the only observational window into non-linear quantum gravity for the foreseeable future.

2.5 Implications of an improved upper limit on r

As detailed in previous sections, a detection of primordial gravitational waves would have profound implications; however, even excluding the presence of gravitational waves at a level observable by CMB-S4 would have important consequences for the theory of inflation. Current constraints already strongly disfavor models that were plausible candidates such as chaotic inflation with a quadratic potential [50]. Upper limits from CMB-S4 would rule out entire classes of inflationary models.

We first present a version of an argument developed in [51, 52, 53] that does not rely on microscopic details of inflationary models. In the limit $\epsilon \ll 1$, equations (2.6) and (2.8) lead to a differential equation

$$\frac{d \ln \epsilon}{d \mathcal{N}} - (n_s(\mathcal{N}) - 1) - 2\epsilon = 0, \quad (2.19)$$

where \mathcal{N} is the number of e-folds until the end of inflation, and $n_s(\mathcal{N}) - 1$ denotes the spectral index evaluated at the wavenumber of the mode which exits the horizon \mathcal{N} e-folds before the end of inflation. Note that ϵ is small (but positive) during inflation and $\epsilon \sim 1$ when inflation ends. If ϵ is a monotonic function of \mathcal{N} this implies $n_s(\mathcal{N}) - 1 \leq 0$ in agreement with observations.

Denoting the number of e-folds before the end of inflation at which the CMB pivot scale exits the horizon as \mathcal{N}_* , the departure from a scale-invariant spectrum observed by the *Planck* satellite is $\mathcal{O}(1/\mathcal{N}_*)$. While this could be a coincidence, it would find a natural explanation if

$$n_s(\mathcal{N}) - 1 = -\frac{p+1}{\mathcal{N}}, \quad (2.20)$$

up to subleading corrections in an expansion in large \mathcal{N} for some real p . Under this assumption, the general solution to equation (2.19) is

$$\epsilon(\mathcal{N}) = \frac{p}{2\mathcal{N}} \frac{1}{1 \pm (\mathcal{N}/\mathcal{N}_{\text{eq}})^p}, \quad (2.21)$$

where we have chosen to parameterize the integration constant by \mathcal{N}_{eq} so that the magnitudes of the first and second term in the denominator become equal when $\mathcal{N} = \mathcal{N}_{\text{eq}}$. We take $\mathcal{N}_{\text{eq}} > 0$ and indicate the choice of sign for the integration constant by ‘±’.

Assuming the epoch during which the modes we observe in the CMB exit is not special so that $\mathcal{N}_* \gg \mathcal{N}_{\text{eq}}$ or $\mathcal{N}_* \ll \mathcal{N}_{\text{eq}}$, equation (2.19) leads to four classes of solutions

$$\text{I. } \epsilon(\mathcal{N}) = \frac{p}{2\mathcal{N}}, \quad (2.22)$$

$$\text{II. } \epsilon(\mathcal{N}) = \frac{p}{2\mathcal{N}} \left(\frac{\mathcal{N}_{\text{eq}}}{\mathcal{N}} \right)^p \quad \text{with} \quad p > 0 \quad \text{and} \quad \mathcal{N}_{\text{eq}} \ll \mathcal{N}_*, \quad (2.23)$$

$$\text{III. } \epsilon(\mathcal{N}) = \frac{|p|}{2\mathcal{N}} \left(\frac{\mathcal{N}}{\mathcal{N}_{\text{eq}}} \right)^{|p|} \quad \text{with} \quad p < 0 \quad \text{and} \quad \mathcal{N}_{\text{eq}} \gg \mathcal{N}_*, \quad (2.24)$$

$$\text{IV. } \epsilon(\mathcal{N}) = \frac{1}{2\mathcal{N} \ln \mathcal{N}_{\text{eq}}/\mathcal{N}} + \frac{p}{4\mathcal{N}} + \dots \quad \text{with} \quad |p| \ll \frac{1}{\ln \mathcal{N}_{\text{eq}}/\mathcal{N}_*} \quad \text{and} \quad \mathcal{N}_{\text{eq}} \gg \mathcal{N}_*. \quad (2.25)$$

As we explain in what follows, if CMB-S4 does not detect primordial B -modes, only class II with $\mathcal{N}_{\text{eq}} \lesssim 1$ will remain viable, the rest will be disfavored or excluded.

The value of \mathcal{N}_* depends on the post-inflationary history of the universe. Equation (2.20) implies that a measurement of the spectral index and its running would determine p and hence \mathcal{N}_* , but unfortunately a

measurement of the running at a level of $(n_s - 1)^2$ is out of reach for CMB-S4. A given reheating scenario predicts \mathcal{N}_* , but the space of reheating scenarios is large. Instantaneous reheating leads to $\mathcal{N}_* \approx 57$ for $k_* = 0.05\text{Mpc}^{-1}$, smaller values correspond to less efficient reheating. We will assume $47 < \mathcal{N}_* < 57$ for the following discussion.

Current constraints on n_s and r from [50] disfavor class III at just over 2σ relative to class II. Furthermore, the best-fit of class III occurs for $p \approx 0$ where classes I, II, and III degenerate so that class III need not be discussed separately. Class IV is disfavored at $2 - 3\sigma$ relative to class II. As a consequence we focus on classes I and II in what follows.

For class I, constraints from the *Planck* satellite and the *BICEP2* and *Keck Array* experiments [50] translate into $p = 0.32 \pm 0.16$ at 1σ , and favor models with inefficient reheating. At the best-fit point in this class, $r = 0.044$ and $n_s = 0.973$, which is currently disfavored relative to class II at $1 - 2\sigma$. Upper limits on r directly translate into constraints on p . A 1σ upper limit on the amount of primordial gravitational waves from CMB-S4 at a level of $r < 0.001$ would imply $p < 0.013$ and effectively rule out this class as it degenerates into class II in this limit.

For class II the tensor-to-scalar ratio is naturally smaller than in class I as long as p is of order unity because $\mathcal{N}_* \gg \mathcal{N}_{\text{eq}}$. Under the additional assumption that the scaling (2.23) should be valid until the end of inflation we have $\mathcal{N}_{\text{eq}} \simeq 1$. In this case, current data from [50] imply $p = 0.67 \pm 0.24$ after marginalization over \mathcal{N}_* . The best-fit occurs for $p = 0.83$ and for instantaneous reheating so that in this class the data favors models with efficient reheating. At the best-fit point, $r = 0.004$ and $n_s = 0.968$. An upper limit of $r < 0.001$ would disfavor this scenario relative to scenarios with $\mathcal{N}_{\text{eq}} \ll 1$ at approximately 2σ . The precise significance depends slightly on the true value of the spectral index. Similarly, for an upper limit of $r < 0.001$, the regime with $p \ll 1$ and equivalently class I would be disfavored relative to class II with $\mathcal{N}_{\text{eq}} \ll 1$ at 3σ . To disfavor the scenario with $\mathcal{N}_{\text{eq}} \simeq 1$ at approximately 3σ relative to $\mathcal{N}_{\text{eq}} \ll 1$ would require an upper limit of $r \lesssim 5 \times 10^{-4}$.

In summary, in the absence of a detection of primordial gravitational waves, CMB-S4 would place constraints on n_s and r that are strong enough to rule out or disfavor all models that naturally explain the observed value of the scalar spectral index in the sense that $n_s(\mathcal{N}) - 1 \propto 1/\mathcal{N}$ and in which the behavior (2.22)-(2.25) provides a good approximation until the end of inflation.

To understand the implications better, let us discuss the models that underlie the classes favored by current data, classes I and II. The potentials can be obtained from

$$\frac{d\phi}{d\mathcal{N}} = M_p^2 \frac{V'}{V} \quad \text{and} \quad \left(\frac{d\phi}{d\mathcal{N}} \right)^2 = 2\epsilon M_p^2, \quad (2.26)$$

where M_p is the reduced Planck mass.

Class I corresponds to models of chaotic inflation with monomial potentials $V(\phi) = \mu^{4-2p}\phi^{2p}$ already considered in [54]. The most commonly studied examples were $p = 1, 2$, both of which are now ruled out or strongly constrained [50]. Models with fractional powers $1/3 < p < 1$ that are still viable candidates have naturally appeared in the study of large-field models of inflation in string theory [34, 35, 55]. If gravitational waves are not observed with CMB-S4, these would be ruled out.

Provided $p \neq 1$, class II corresponds to potentials of the form

$$V(\phi) = V_0 \exp \left[- \left(\frac{\phi}{M} \right)^{\frac{2p}{p-1}} \right], \quad (2.27)$$

with $M = \sqrt{\alpha(p)\mathcal{N}_{\text{eq}}}M_p$ where $\alpha(p)$ is of order unity for the range of p of interest. For $p > 1$ inflation occurs when $\phi \ll M$. In this regime, the potential behaves like a hilltop model $V(\phi) \approx V_0(1 - (\phi/M)^n)$ with $n = 2p/(p-1)$. For $0 < p < 1$ inflation occurs for $\phi \gg M$ and $V(\phi) \approx V_0(1 - (M/\phi)^n)$ with $n = 2p/(1-p)$. In the limit $p \rightarrow 0$ in which classes I, II, III become degenerate, the ϕ -dependence becomes logarithmic.

For the special case $p = 1$ the dependence on the inflaton in (2.27) becomes exponential and in the inflationary regime the potential is well approximated by $V(\phi) \approx V_0(1 - \exp(-\phi/M))$ with $M = \sqrt{\mathcal{N}_{\text{eq}}}M_p$. There are many examples of models with a potential with this asymptotic behavior for $\phi \gg M$. Some of them are the Starobinsky model [56], Higgs inflation [57, 58], an early example of chaotic inflation [59], and the T-model [60].

If only the asymptotic forms of the potentials agree with (2.27), equation (2.20) will not be exact and the departures from (2.27) will be encoded in the subleading terms that vanish more rapidly than $1/\mathcal{N}$ in the limit $\mathcal{N} \rightarrow \infty$. Unfortunately, just like the running of the scalar spectral index, the subleading contributions are typically too small to be detected.

Note that \mathcal{N}_{eq} sets the characteristic scale in field space. For \mathcal{N}_{eq} of order unity, the variation of the inflaton is naturally given in units of the reduced Planck mass while for $\mathcal{N}_{\text{eq}} \ll 1$ the characteristic scale in field space is sub-Planckian.

This allows us to rephrase the lesson we can draw from an upper limit on r from CMB-S4.

In the absence of a detection, CMB-S4 would rule out or disfavor all models that naturally explain the observed value of the scalar spectral index and in which the characteristic scale in field space equals or exceeds the Planck scale.

Unfortunately, because of the scaling $M \propto \sqrt{\mathcal{N}_{\text{eq}}}$ it will only be possible to obtain constraints $M \lesssim M_p$ but not $M \ll M_p$. It should also be kept in mind that a natural explanation of the value of the scalar spectral index is not guaranteed and its value could be an accident. That a natural explanation is possible is, however, encouraging.

2.6 Beyond r : characterizing the B-mode spectrum, testing inflationary expectations and distinguishing models

If a detection of primordial gravitational waves is made with CMB S4, the next step is to understand the possible sources of the signal. The spectrum of B-modes from vacuum fluctuations of the metric, amplified by inflation driven by a scalar field, is nearly scale-invariant and very nearly Gaussian. Any deviation from these features would be an indication of a richer inflationary sector, or possibly even of non-inflationary primordial physics. In this section we discuss how well CMB S4, given a detection, could characterize the shape of the B-mode spectrum, test for significant higher order correlations involving tensor modes, and test for parity violation. We frame this discussion in terms of current candidate models that produce a B-mode spectrum significantly different from the simplest inflationary signal.

Do we need to talk about this somewhere? maybe not here? Also, it is important to note that a detection of primordial B-modes with a Stage 3 instrument *would* change the survey strategy for Stage 4. This is further outlined in Section ?? and particularly in Figures ??.

2.6.1 Distinguishing vacuum fluctuations from other particle physics sources of B-modes

In non-minimal models with additional sectors coupled to the inflaton, excitations and particle production associated with other fields during inflation can source additional primordial gravitational waves [61, 62, 63]. The new fields and interactions that generate additional tensor fluctuations also generically contribute to the scalar fluctuations, so the non-vacuum B-mode signal is significantly constrained by the observed scalar power spectrum and its high degree of Gaussianity [63, 64, 65, 66, 67]. It is a rather generic feature that scalar non-Gaussianity constraints are relevant for models that produce B-modes from a mechanism other than the inflationary amplification of vacuum fluctuations, so in Figure XXX we present the expected constraints from CMB S4 in the $r - f_{\text{NL}}^{\text{equil}}$ plane. (Scalar non-Gaussianity is discussed in more detail in Section 2.7.2.)

It has been found that in cases where additional sectors are directly coupled to the inflaton (with stronger than gravitational strength couplings) *Planck* satellite constraints do not allow for the secondary source signal to have an amplitude competitive with the vacuum signal [67, 66]. However, a model with a significant non-vacuum signal can be constructed if the inflationary sector is only gravitationally coupled to a hidden sector containing a light pseudo-scalar and a gauge field during inflation [63, 68]. Fluctuations of the light scalar excite fluctuations of the gauge field, which in turn leads to gravitational wave production. To evade constraints from scalar non-Gaussianity, the source field's potential must be adjusted so that the production of gauge field quanta occurs only around the time the modes contributing to the multipoles relevant for the B-mode search leave the horizon [69]. Then, at this expense of fine-tuning the scales on which production occurs, there exists a range of values for other parameters which can lead to a gravitational wave signal competitive with the vacuum fluctuations while remaining consistent with existing Planck data [69, 68]. For example, the gravitational waves from gauge field production could be measured at a level of $r = 10^{-1}$ with a vacuum contribution of only $r = 10^{-4}$. While in that case the determination of the scale of inflation is affected by less than one order of magnitude, adjusting the parameters of the scenario may allow for more dramatic modifications of the relationship between r and the inflationary energy scale given by Eq. (2.14).

When secondary production of this sort is large enough to dominate the signal, the predicted gravitational wave spectrum differs from that of the vacuum fluctuations in several ways. First, since the production mechanism is not continuous (to avoid non-Gaussianity constraints), and so sources a B-mode spectrum that is far from scale-invariant. Second, the tensor spectrum is strongly non-Gaussian. Finally, the gravitational waves resulting from the gauge field come with a definite handedness [70, 71] and the angular bispectrum of B-modes would be dominated by $\ell_1 + \ell_2 + \ell_3 = \text{even}$, which would vanish in any theory that respects parity.

Figure XXX below shows the projected contours in the $r-n_T$ plane for CMB S4, assuming a detection of $r = XXX$. Figure XXX shows the expected constraint on non-Gaussianity in the tensor sector. Figure XXX shows constraints on parity violation. The predictions from the secondary source model are shown on the figures, and we see that this example is clearly distinguished if the non-vacuum contribution is so large that r is not a good indicator of the scale of inflation.

Tests for parity violation and the shape of the spectrum are also of broader interest for inflation models whose particle content departs from the minimal scalar field for reasons other than specifically generating non-vacuum gravitational waves. In particular, scenarios in which non-Abelian gauge fields play a significant role in sourcing inflation are closely related to the models discussed above. In chromo-natural inflation and gauge-flation scenarios [72, 73, 74, 75, 76, 77, 78], the central piece is a homogeneous and isotropic, flavor-space locked gauge field that helps slow the roll of the inflaton or else is the inflaton itself. For a non-Abelian field with SU(2) symmetry, this means the three flavor gauge vector potentials are mutually orthogonal in space. The stress-energy of this configuration could leave a unique imprint on a spectrum of primordial gravitational waves, which would be transferred to the B-mode spectrum in the CMB. The non-Abelian nature of the field

introduces a preferred handedness onto this medium leading to an enhancement of left (or right) circularly polarized gravitational waves. Again this would lead to parity-violating EB and TB correlations [79, 80] or parity violating higher n -pt functions. If this process takes place in the post-inflationary environment, the gauge field could further impress a periodic modulation on the gravitational wave spectrum [81, 82]. Although the basic chromo-natural and gauge-flation models have been ruled out [83], these unique features are expected to be generic to any viable variations on these scenarios and would be constrained by CMB S4 as shown in Figure XXX above.

Multi-field inflationary scenarios that end with phase transitions [84, 85, 86, 87, 88, 89, 90] and models of brane-inflation in string theory [91, 92, 93] generically predict some level of vector and tensor modes actively sourced by topological defects. In particular, either a breaking of a $U(1)$ symmetry or the production of fundamental strings at the end of inflation can lead to “cosmic strings” whose B-mode spectrum is primarily generated by vector modes and peaks on small scales ($\ell \sim 600 - 1000$) and is more similar in shape to the E to B lensing signal than to the vacuum spectrum. CMB-S4 should be able to distinguish even a small contribution from such sources [94], but the precise bounds from non-detection are related to the precision with which the lensing signal can be removed. Estimates made in reference [95, 96] indicate that CMB-S4 should be able to improve the limit on cosmic string tension by at least an order of magnitude beyond the current bounds from the CMB ($G\mu \sim 10^{-7}$ [97, 98]) and may be competitive with direct detection limits from the stochastic gravitational wave background ($G\mu \sim 10^{-11}$ or 10^{-8} depending on the model assumed for string loops [99]). In addition, the spectra of different types of defects have different shapes, and should be distinguishable [100, 96]. Measuring the location of the main peak would provide valuable insights into fundamental physics. For example, in the case of cosmic superstrings the position of the peak of the B-mode spectrum constrains the value of the fundamental string coupling g_s in string theory [96].

Post-inflationary phase transitions themselves have also been proposed as a source of nearly scale-invariant gravitational waves detectable through CMB polarization (and direct detection) [101, 102]. Even for a spectrum that matches the inflationary result on small scales, any such signal can in principle be distinguished from the inflationary signal by the absence of super-horizon correlations at the time of recombination. A framework to extract specifically this bit of the signal was proposed in [103] and could be applied to robustly extract the part of any signal that must come from physics outside of the hot big bang paradigm. Existing forecasts in the literature [104] indicate that a ground based survey alone will not be able to detect super-horizon correlations at high significance if r is much below 0.1. But, if CMB S4 does make a detection, this physics could be in reach of an eventual satellite mission. **This seems to be a really important problem for interpretation of B modes, as observed by S4, as due to vacuum fluctuations. How should this problem be handled? Do we highlight this in the introduction to this chapter? Do we get it into Chapter 1 as well?**

2.6.2 Probing matter and gravitational interactions at the inflationary scale

The tensor tilt as a probe of the potential and non-minimal coupling: If the amplitude of primordial B-modes is large enough to be measured, we can begin to constrain the shape of the spectrum. The simplest inflation scenarios all predict a red spectrum for gravitational waves, and the canonical single field consistency relation fixes $n_t = -r/8$. For a single field with a sound speed less than one, or multiple fields, $n_t/r < -1/8$ instead [105]. However, allowing the inflaton to couple to higher curvature terms can produce a blue tilt [106]. A detection of primordial gravitational waves on CMB scales would allow predictions, especially relevant for a blue index, for the amplitude expected on the much smaller scales accessible to direct detection. The recent detection of gravitational waves by LIGO, as well as the beginning of operation of the LISA pathfinder instrument, open an exciting new era of gravitational wave science. If CMB-S4 also sees a signal, LIGO and

future instruments may be particle physics detectors as well as astrophysical observatories. Recent analysis [107, 108] shows the complementarity between observations over a wide range of scales in constraining the spectrum (although one must assume a constant tilt n_t over many orders of magnitude). In principle, the number of relativistic species N_{eff} also puts constraints on n_t [107] (see Sec. 3.4.5).

The tensor amplitude and field content that modifies the scalar power: Since non-minimal inflation models with multiple fields or a small sound speed for a single degree of freedom predict a tensor-to-scalar ratio that is suppressed, a detection of gravitational waves can be used to constrain the physics that produces the suppression in these scenarios. For single clock scenarios, this link is relatively straightforward [109] and a detection of r can provide an upper limit on the speed of sound (and so a limit on non-Gaussianities). For multi-field scenarios many more details of the model must be specified [110], but r together with bounds on isocurvature and local type non-Gaussianities may aid in model discrimination.

Other signatures of a modified gravitational sector: Coupling the inflaton to higher curvature terms can also introduce parity violation in the spectrum of primordial gravitational waves [79, 111, 112, 113]. Reference [113] contains some example amplitudes of the coupling that would be detectable for a detection of $r = 0.05$; reference [80] discusses distinguishability of chiral gravity waves from other possible sources of parity violation, such as uniform cosmic birefringence. In addition, the momentum structure of the three-point function of gravitational waves would also be a sensitive probe of possible extensions of Einstein gravity [114]. However, the amplitude of the three point correlations between tensors alone in both standard inflation and extensions is small, at most $f_{\text{NL}}^{\text{tensor}} \lesssim 1$ [115, 114]. So while any constraint on the gravitational three-point function would be a useful data point for secondary sources, it is unlikely to be significant for vacuum fluctuations. Effectively then, there are 3 different possibilities to produce larger effects; different symmetry pattern (e.g. solid inflation [116], gauge-flation [72]); GWs not produced as vacuum fluctuations or; multiple tensors (e.g. bigravity) [117]. A detection would therefore constitute a clear signature of new physics. Any non-zero detection could also be sourced by higher order massive spin field that couples to two scalars and one graviton [118]. From an observational point of view, such interactions are interesting to constrain using correlators that couple one B field with a combination of two E and T modes. No constraint on this type of non-Gaussianity exists and CMB-S4 is projected to improve constraints on graviton interactions by two orders of magnitude [119] compared to Planck. Finally, it is worth noting that, if primordial gravitational waves were indeed chiral, they may present themselves first through a non-vanishing cross-correlation of B -modes with temperature, as demonstrated in Ref. [112].

The below is currently just a dump from what was in NG section Correlators including at least one B mode will benefit significantly from the improved sensitivity of B modes as we will show next. We define [119]

$$\langle \zeta(\vec{k}_1)\zeta(\vec{k}_2)\gamma_s(\vec{k}_3) \rangle = (2\pi)^3 \delta^{(3)} \left(\sum_{n=1}^3 \vec{k}_n \right) \mathcal{B}(k_1, k_2, k_3) e_{ij}^s(\vec{k}_3) \hat{k}_1^i \hat{k}_2^j, \quad (2.28)$$

with

$$\mathcal{B}(k_1, k_2, k_3) = 16\pi^4 A_s^2 \sqrt{r} f_{\text{NL}}^{\gamma\zeta\zeta} F(k_1, k_2, k_3) \quad (2.29)$$

and e_{ij}^s the transverse traceless polarization tensor as before. In the simplest model of inflation $f_{\text{NL}}^{\gamma\zeta\zeta} = \sqrt{r}/16$ [115, 114] which is undetectable with CMB measurements. As mentioned in Sec. 2.6.2, a measurement of this correlation would be an immediate indication of some deviation from the simple inflationary paradigm and as such would provide valuable information [117, 118]. The coupling above can be constrained using any type of correlation that contains one B mode, e.g. $\langle BTT \rangle$ or $\langle BEE \rangle$. For F we consider local and equilateral triangles and forecast the constraint on the amplitude in Fig. 12 for $\langle BTT \rangle$. We anticipate similar constraints for $\langle BTE \rangle$ and $\langle BEE \rangle$.

We consider two scenarios; one in which the B modes are noise dominated and one in which they are cosmic variance limited by a future detection of r . For a local shape primordial bispectrum the error decreases

Type	Planck	CMB-S4	rel. improvement
local	$\sigma(\sqrt{r}f_{\text{NL}}) = 15.2$	$\sigma(\sqrt{r}f_{\text{NL}}) = 0.3$	50.7
equilateral	$\sigma(\sqrt{r}f_{\text{NL}}) = 200.5$	$\sigma(\sqrt{r}f_{\text{NL}}) = 7.4$	27.1
local ($r = 0.01$)	$\sigma(\sqrt{r}f_{\text{NL}}) = 15.2$	$\sigma(\sqrt{r}f_{\text{NL}}) = 0.7$	25.3
equilateral ($r = 0.01$)	$\sigma(\sqrt{r}f_{\text{NL}}) = 200.8$	$\sigma(\sqrt{r}f_{\text{NL}}) = 14.7$	13.7

Table 2-1. Forecasted constraints on local and equilateral shapes sourced by primordial and equilateral correlations of the form $\langle \gamma\zeta\zeta \rangle$ constrained through $\langle BTT \rangle$. Planck forecast is based on Blue Book values, with $f_{\text{sky}} = 0.75$. Constraints were derived using the flat-sky approximation as in Ref. [119] with $\ell_{\min} = 30$ with no cosmic variance in B . We expect similar constraints from $\langle BEE \rangle$ and $\langle BTE \rangle$. For $r = 0.01$ Planck is still noise dominated, while CMB-S4 is cosmic variance dominated.

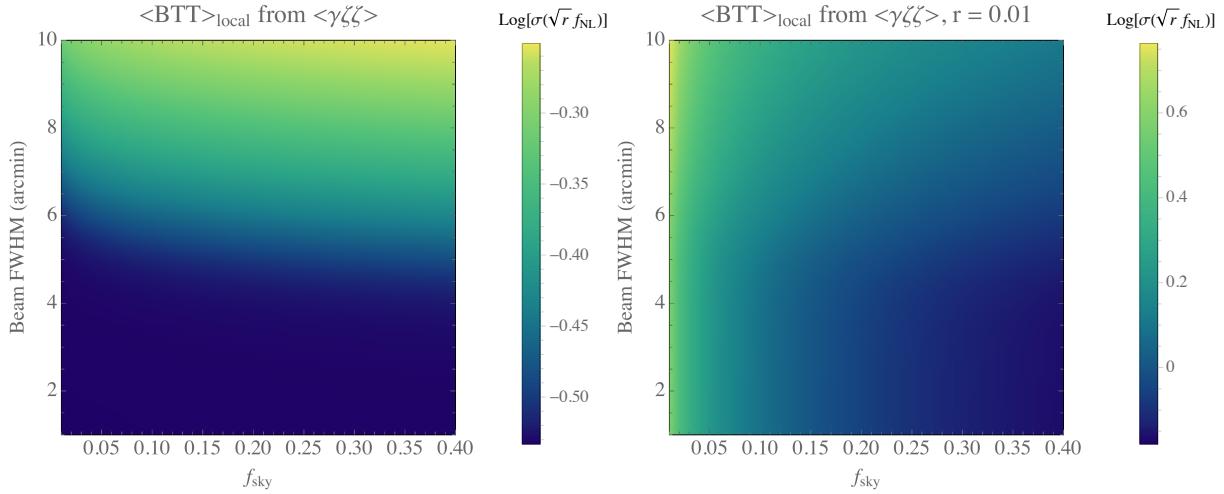


Figure 8. Left: Fixed effort density plot for local type BTT . Left, the effect of beam and f_{sky} variation in noise dominated B mode. f_{sky} has very little effect, as it is cancelled by the f_{sky} associated with the masking of the total map. Right: adding a signal changes this picture, since the cancellation no longer happens for $r > 0.001$ (i.e. when CMB-S4 becomes cosmic variance limited). A larger sky fraction benefits cosmic variance limited $\langle BTT \rangle$.

as you decrease the beam. The signal is dominated by large angle B modes correlated with small angle T modes as was pointed out in [119]. An equilateral component from the $\langle BTT \rangle$ correlation function suffers from the decaying tensor modes, which are negligible for $\ell_B > 500$. After this, no equilateral triangles can be constructed and the error saturates. The effect is that equilateral type non-Gaussianities are almost insensitive to beam size.

On the right we show what happens when B gets a primordial component. Only if $r < 0.001$ CMB-S4 noise dominates cosmic variance. We compare the potential constraining power of *Planck* to that of CMB-S4 in Tab. 2-1; as r gets larger, the relative improvement decreases.

In addition we consider and experiment with fixed effort. We show the results in Fig. 8 and 9. We find that BTT is practically insensitive to f_{sky} as long as B modes are noise dominated. For $r = 0.001$ and above a larger sky fraction helps constrain $\langle BTT \rangle$.

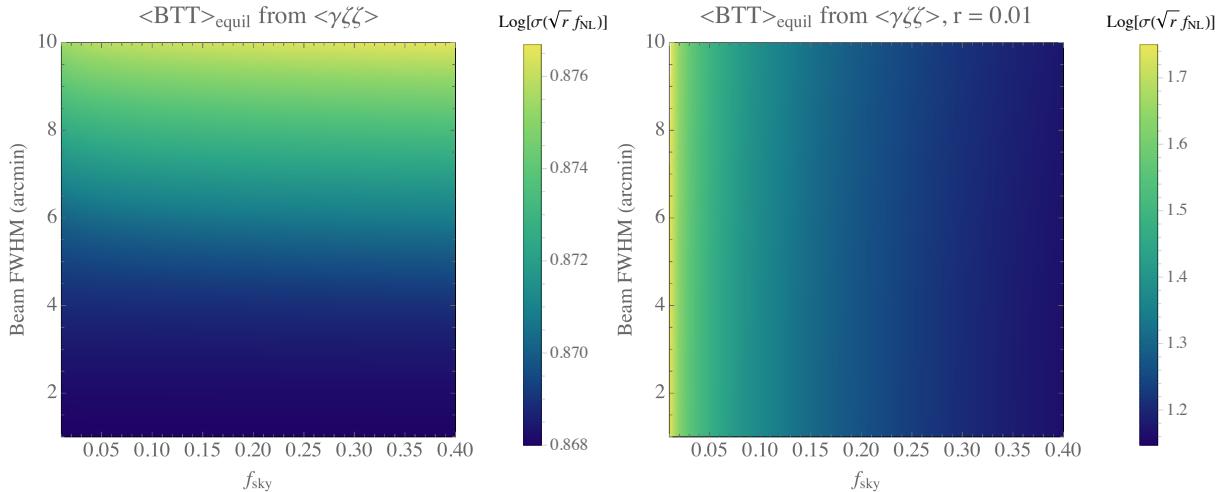


Figure 9. Left: Fixed effort density plot for equilateral type BTT. Similar to local type bispectra, a larger sky fraction benefits cosmic variance limited $\langle BTT \rangle$. Right: Fixed effort density plot for equilateral type BTT with a non-zero tensor-to-scalar ratio $r = 0.01$.

2.6.3 Constraining alternatives to inflation

Vacuum fluctuations during inflation provide a simple, elegant, and compelling mechanism to create the initial seeds required for structure formation. One of inflation’s most robust predictions is an adiabatic, nearly scale-invariant spectrum of scalar density perturbations. This prediction is in excellent agreement with observations, especially considering the need to account for a small deviation from exact scale invariance. However, it is disputable whether these observations can be considered a proof that inflation actually occurred. Clearly, a fair evaluation of the status of inflation requires the consideration of competing theories and the hope to find experimental distinctions between inflation and these alternatives.

Leading alternatives to inflation can be classified into two primary categories based on the way in which they account for the observed causality of the scalar density fluctuations. Bouncing cosmologies rely on a cold, large universe initially and a subsequent phase of slow contraction. This is then followed by a bounce which leads to an expanding and decelerating FRW cosmology. The most well studied examples are provided by Ekpyrotic/Cyclic models [3, 2] and more recently ‘matter bounce’ models [6, 7, 8]. The second class of alternatives to inflation arise from models that invoke a loitering phase of the cosmic expansion prior to the hot Big Bang – with String Gas Cosmology [120, 121, 122] providing an example.

A detailed critique of these alternatives and their relevance to the science case for a near-term CMB-based mission was presented in Appendix B of [123]. Since that publication, these alternative approaches to inflation have received considerable attention, however as science drivers for the CMB-S4 mission there are two important points to re-emphasize:

1. These alternatives invoke novel and incompletely understood physics to solve the problems associated with standard Big Bang cosmology. This implies important theoretical challenges that have to be addressed carefully before the models mature into compelling alternatives to inflation.
2. Most or all of the alternatives to inflationary cosmology predict negligible tensors on CMB scales. This strengthens the case for considering B -modes as a “smoking gun” of inflation. It should be considered an important opportunity to use CMB observations to constrain all known alternatives to inflation.

One property that is shared by many (if not all) alternatives to inflation is that they require a violation of the Null Energy Condition (NEC). Such a violation typically implies the existence of catastrophic instabilities and/or fine-tuning of initial conditions. This presents an important challenge for alternatives to inflation, but it does not imply that alternatives are impossible to realize. An example of a stable bounce violating the NEC was put forward in [124] and then used in the New Ekpyrotic scenario in [125, 126]. Although this model is consistent at the level of effective field theory, it is not clear whether it is possible to find a UV completion for it. This is a very important issue because the quantization of the New Ekpyrotic theory, prior to the introduction of a UV cutoff and a UV completion, leads to a catastrophic vacuum instability [127]. Similar challenges arise in models like String Gas Cosmology where NEC violation is required to exit the loitering phase to a radiation dominated universe [128, 129, 130, 131]. Whether such obstacles can be overcome is an area of ongoing research.

However, despite the theoretical challenges in understanding the background evolution, it has been argued that many of the observational predictions of these alternative models are independent of these issues. Most notable is that all known alternative constructions seem to predict negligible tensor modes on large scales. This was an early prediction of Ekpyrotic models, and appears true as well for the more recently studied matter bounce models when constraints on scalar non-Gaussianity are also taken into consideration [132]. It is not yet clear if an observable amplitude of non-vacuum primordial gravitational waves could be sourced during contracting phase, but see [133] for some recent work in that direction.

A detection of primordial gravitational waves with a spectrum consistent with vacuum fluctuations would rule out all currently proposed alternatives to inflation.

2.6.4 Constraints on the graviton mass

Theories of massive gravity come in many flavors (see e.g. [134, 135]), and their predictions in the scalar sector differ significantly. However, by definition, the dispersion relation for the graviton in all of them is

$$\omega^2 = p^2 + m_g^2, \quad (2.30)$$

where p is the physical momentum and m_g the possibly time-dependent graviton mass. As a consequence, gravitational waves necessarily have frequencies $\omega > m_g$. A detection of primordial B -mode polarization on angular degree scales may be considered as a detection of gravitational waves with frequencies $\omega \sim H_{\text{rec}}$ through the quadrupole they produce in the primordial plasma, where $H_{\text{rec}} \approx 3 \times 10^{-29}$ eV is the Hubble parameter at recombination. A detection then implies a model-independent bound $m_g < H_{\text{rec}}$ or

$$m_g < 3 \times 10^{-29} \text{ eV}. \quad (2.31)$$

If the graviton mass is time-dependent, this should be interpreted as a constraint on the graviton mass around the time of recombination.

Because the perturbations in the primordial plasma before and around recombination are linear, the effect of the graviton mass is straightforward to incorporate by a simple modification of the field equation for tensor metric perturbations so that the above argument can be made more quantitative. The equation of motion for the transverse traceless metric perturbation γ takes the same form as for a minimally coupled massive scalar field,

$$\ddot{\gamma}_k(\tau) + 2\frac{\dot{a}}{a}\dot{\gamma}_k(\tau) + (k^2 + m_g^2 a^2)\gamma_k(\tau) = 0. \quad (2.32)$$

Here k is the comoving momentum of the metric perturbation, and we work in the conformal coordinates so that the background cosmological metric is

$$ds^2 = a^2(\tau)(d\tau^2 - d\mathbf{x}^2). \quad (2.33)$$

The consequences of this modification are discussed in detail in [136]. The most important consequence is that superhorizon modes start to oscillate around the time τ_m when $H(\tau_m) = m_g$, and their amplitude subsequently redshifts as $a^{-3/2}$. In contrast, in the massless case all modes remain frozen until they enter the horizon. This results in a suppression of the amplitude of primordial B -modes for $m_g \gg H_{\text{rec}}$, and a detection of B -modes would rule out this possibility. For masses around H_{rec} , there is no suppression, but the angular power spectra are modified by the presence of a graviton mass, and a detection of primordial B -mode polarization would allow a measurement of the graviton mass. A detection of primordial gravitational waves with an angular B -mode power spectrum consistent with that expected in general relativity would imply $m_g < 3 \times 10^{-29}$ eV.

For comparison, the current model-independent bounds on the graviton mass arise from the indirect detection of $\sim 3 \times 10^{-5}$ Hz gravitational waves through the timing of the Hulse-Taylor binary pulsar [137], and the bound on the difference in arrival times for gravitational waves with different frequencies in the recent direct detection of astrophysical gravitational waves with LIGO [1]. The resulting bounds are $m_g \lesssim 10^{-19}$ eV and $m_g \lesssim 10^{-22}$ eV, respectively.

A detection of B -mode polarization consistent with expectations in the context of general relativity would improve current bounds on the mass of the graviton by nearly seven orders of magnitude.

We note that this improvement is calculated assuming measurements of the degree-angular-scale B modes only. Measurements of B -mode polarization on the largest angular scales would further strengthen the bound.

2.7 Improved constraints on primordial density perturbations

All current data are consistent with primordial density perturbations that are adiabatic, Gaussian, and nearly scale-invariant. With sufficient angular resolution, CMB-S4 will significantly improve current constraints on the scale dependence of the primordial power spectrum of scalar perturbations, on departures from Gaussianity, and on departures from adiabatic perturbations. In fact, it will measure anisotropies in both the temperature and E-mode polarization of the CMB to cosmic variance over the entire range of multipoles that is not contaminated by foregrounds. As a consequence, it will place the strongest constraints achievable by any ground-based CMB experiment on all observables that benefit from the number of modes measured, such as the primordial power spectrum and higher-order correlations.

2.7.1 The power spectrum of primordial density perturbations and mean curvature

The density perturbations are close to scale-invariant but not exactly so. In the context of Λ CDM, *Planck* has measured the scalar spectral index to be $n_s = 0.9677 \pm 0.0060$ and has established $n_s - 1 < 0$ at more than 5σ . CMB-S4 will improve current constraints on the spectral index by more than a factor of two to $\sigma(n_s) = 0.0028$, and this constraint will provide valuable constraints on the space of inflationary models.

As mentioned in section 2.5, a measurement of the running of the scalar spectral index with a precision of a few parts in ten thousand would allow a measurement of p in equation (2.20), or equivalently \mathcal{N}_* . This precision cannot be achieved with CMB-S4, but with a precision of one part in a thousand, it will test the idea that the lack of power on large angular scales might be explained by scale dependence of the spectral index [138].

Models of inflation that achieve super-Planckian inflaton displacements from repeated circuits of a sub-Planckian fundamental period may give rise to oscillatory features in the spectrum of primordial perturbations. The features may arise either from instanton effects or from periodic bursts of particle or string production. A search for such features is well motivated even though the amplitude is model-dependent and may be undetectably small. A detection would provide clues about the microscopic origin of the inflaton, the absence of a detection can constrain the parameter space of these models in interesting ways. CMB-S4 would tighten the constraints on the amplitude of features in the primordial power spectrum by a factor of three.

Other physical effects during inflation can lead to small features in the observed power spectrum, e.g by changing the equation of state during inflation []. Because of the stringent constraints on the minimum number of E-folds, such modifications can not last very long in order not to end inflation and associated features only affect a small range of scales. It is therefore unlikely that CMB-S4 will significantly improve constraints on these type of features unless they are on very small scales.

2.7.2 Higher order correlations of scalar modes

Any detection of departures from Gaussianity would shed light on the interactions either of the inflaton with itself or between the inflaton and other degrees of freedom. From the discussion of various scenarios to produce primordial B-modes, it is also a common theme that constraints on non-Gaussianity significantly cut into the model space of proposals to produce a B-mode signal of observable strength other than the standard inflationary signal.

The lowest order correlation function that encodes departures from Gaussianity is the three-point function

$$\langle \zeta(\vec{k}_1)\zeta(\vec{k}_2)\zeta(\vec{k}_3) \rangle = (2\pi)^3 \delta(\vec{k}_1 + \vec{k}_2 + \vec{k}_3) B_\zeta(\vec{k}_1, \vec{k}_2, \vec{k}_3), \quad (2.34)$$

where the delta function comes from translation invariance. In many scenarios invariance under rotations and translations guarantee that the bispectrum $B_\zeta(k_1, k_2, k_3)$ only depends on the magnitudes k_1 , k_2 , and k_3 . A model independent search for the bispectrum, or equivalently the angular bispectrum $b_{\ell_1 \ell_2 \ell_3}$, is not computationally tractable, and in practice searches place constraints on the amplitudes f_{NL} of certain theoretically motivated functional forms, or shapes.

Perhaps of special interest for CMB-S4 are non-Gaussian signatures that would be expected in models of large-field inflation. For example, in models in which the inflaton is an axion, there is only an approximate discrete shift symmetry. In that case instanton contributions to the potential and periodic bursts of particle or string production naturally lead to periodic features in the bispectrum. If moduli in the underlying string constructions do not evolve appreciably, instanton contributions lead to oscillations with a constant amplitude in the logarithm of k . In general, moduli evolve during inflation and cause a drift in the frequency and a scale-dependent amplitude [139]. At present, these shapes have not yet been constrained systematically. Often these contributions will lead to counterparts in the power spectrum and are expected to be detected there first [140], but this need not be the case [141]. A first attempt has been made [97] to look for resonant and local features in the bispectrum, and a more dedicated analysis is underway. Since features in the

power spectrum and the bispectrum generally contain correlated parameters [142, 143, 144, 145], statistical methods have been developed to use the power of both the power spectrum and the bispectrum to further constrain model space [146, 147]. Signatures of higher order massive spin fields [148, 149] would also lead to a bispectrum with decaying features, which will not be present in the power spectrum. So a search for these more general shapes is well motivated. Using both T and E -mode polarization, CMB-S4 will improve constraints by about of two compared to *Planck* as we will show in the next section. Bispectra containing at least one B mode, will generally be much better constrained (currently there is no bound on such correlation functions) and will benefit from CMB-S4. We will consider two examples in the next section.

More generally, one can divide the space of non-Gaussian inflationary models into those whose signals either (1) indicate fluctuations in degrees of freedom other than the inflaton, or (2) indicate non-trivial self-interactions of the effective inflaton fluctuation. Given the forecasted improvements from CMB-S4 over the *Planck*, it is unlikely that this instrument would uncover strong evidence of non-Gaussianity. However, since the *Planck* constraints have not ruled out $f_{\text{NL}} \sim \mathcal{O}(1)$, even a hint of non-Gaussianity would be extremely interesting. Here we briefly review the physics in the two cases.

A detection of the widely studied local shape would have far reaching theoretical implications. A detection of this shape would rule out all models of single clock inflation [150]. In addition, such a signal would open the door to significant cosmic variance on all scales from coupling of fluctuations within our observed volume to any super-Hubble modes [151, 152, 153]. Indeed, there would be room for a significant shift between the observed amplitude of scalar fluctuations (and so the observed r) and the mean value of fluctuations on much larger scales [154]. Any scenario that predicts local non-Gaussianity together with fluctuations on scales much larger than our observed volume predicts a probability distribution for our observed $f_{\text{NL}}^{\text{local}}$, but many well-motivated scenarios also predict a small mean value. These include the simplest modulated reheating scenario [155] and ekpyrotic cosmology [156], both of which predict mean values of $f_{\text{NL}}^{\text{local}} \sim 5$. Currently the strongest constraints on the local shape come from the *Planck* 2015 temperature and polarization analysis which finds $f_{\text{NL}}^{\text{local}} = 0.8 \pm 5.0$ [97]. A noise-free cosmic variance limited CMB experiment is expected to produce constraints on $f_{\text{NL}}^{\text{local}}$ with 1σ error bars of about 3 [157]. Therefore the improvement expected of CMB-S4 over current limits is slightly less than a factor of two. This is not sufficient to reach the interesting theoretical threshold around $|f_{\text{NL}}^{\text{local}}| \lesssim 1$ [158], but will still reduce the space of viable models or hint at a detection. CMB-S4 could, for example, provide hints for the mean level of non-Gaussianity expected from modulated reheating scenario or ekpyrotic cosmology at roughly 2σ . The simplest curvaton scenario, which predicts $f_{\text{NL}} = -5/4$ [159], will unfortunately be out of reach. Large-scale structure surveys (eg., [160]) may eventually achieve constraints $\sigma_{f_{\text{NL}}} \sim \mathcal{O}(1)$. Those observations of the inhomogeneities in the late universe would be very complementary to the results of CMB-S4.

Equilateral [161] and orthogonal shapes [162] shapes arise in a wide variety of scenarios involving interaction during inflation that respect the underlying shift symmetry (i.e. approximately scale-invariant). This may include scenarios where inflaton fluctuations have non-trivial self-interactions [163, 164, 165, 166, 167, 162] or couplings between the inflaton and other (potentially massive) degrees of freedom [168, 169, 170, 142, 171, 172, 148]. While many of these models lead to different shapes in detail, the signal-to-noise is typically dominated by equilateral configurations. One of the important features of these shapes is that single-field slow-roll inflation necessarily produces $f_{\text{NL}}^{\text{equil}} < 1$ [173] and therefore any detection of $f_{\text{NL}}^{\text{equil}} \geq 1$ would rule out this very wide class of popular models. Furthermore, a detection would imply that inflation is a strongly coupled phenomena and/or involved more than one field [109, 158, 174]. These possibilities could be distinguished, in principle, with further observations. Current constraints on the equilateral and orthogonal shapes are $f_{\text{NL}}^{\text{equil}} = -4 \pm 43$ and $f_{\text{NL}}^{\text{ortho}} = -26 \pm 21$, both (68% CL) [97]. In single-field inflation, the amplitude of the non-Gaussianity typically suggests a new energy scale, M_s , such that $f_{\text{NL}}^{\text{equil}} \propto A_s^{-1/2} (H/M_{c_s})^2$ [167, 175]. This energy scale is where self-interactions become strongly coupled and current limits translate into $M_{c_s} > \mathcal{O}(10)H$. In the presence of additional hidden sectors, the amplitude of non-Gaussianity scales with the

strength of the coupling between the inflaton and these additional fields, usually suppressed by an energy scale Λ . Current limits give $\Lambda > (10 - 10^5)H$ [176, 177] with the variation depending mostly on the dimension of the operator coupling the two sectors. For $r > 0.01$, these constraints require some of these interactions to be weaker than gravitational. The improvements from CMB-S4 would further tighten existing constraints on a wide variety of interactions of the inflaton with itself and any other fields that are excited during inflation.

2.7.3 Forecasts

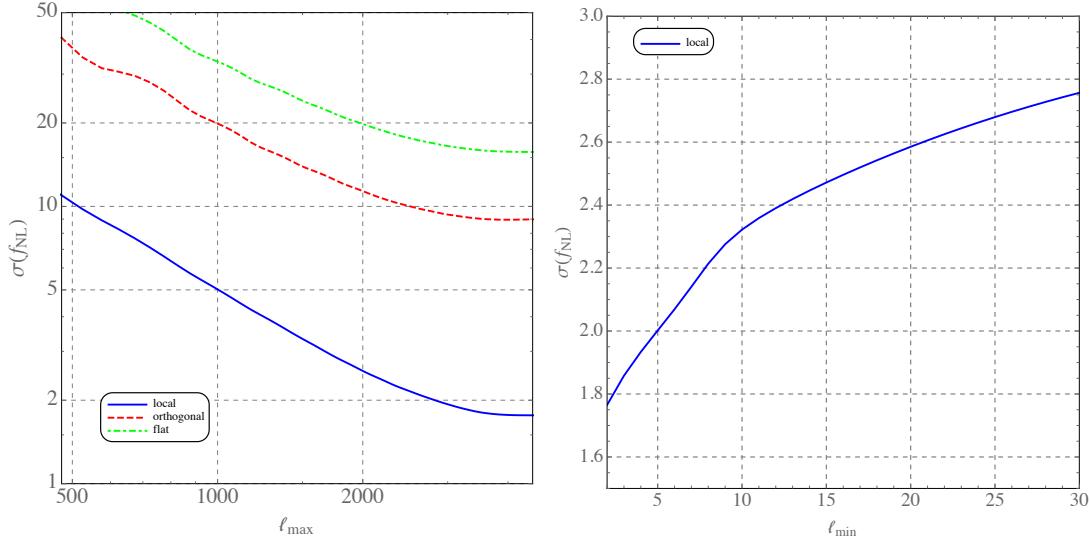


Figure 10. Left: forecasts on the constraining power of CMB-S4 combined with Planck on 3 different types of non-Gaussianities as a function of ℓ_{max} using T and E modes. Right: Sensitivity of local type non-Gaussianities as a function of the minimum multipole. The local type non-Gaussianity benefits from having the lowest multipoles, while the other standard shapes saturate to a constant at low ℓ_{min} (not shown, see Tab. 2-2). The above plot was generated using $f_{\text{sky}} = 0.4$, T -noise = $1 \mu\text{K}^-$ and E -noise = $\sqrt{2} \mu\text{K}^-$ and a beam of $1'$ and $\ell_{\text{min}} = 30$ combined with Planck Blue Book values in the range $2 \leq \ell_{\text{min}} \leq 30$.

We forecast the constraints on non-Gaussianities from CMB-S4. We consider the most well motivated non-Gaussian shapes; local, equilateral, orthogonal and enfolded with a normalization¹ as defined in [161] and [178], i.e.,

$$B_\zeta(k_1, k_2, k_3) = \frac{3}{5}(4\pi^4)B_\Phi(k_1, k_2, k_3) = \frac{3}{5}(4\pi^4)2A_s^2f_{\text{NL}}F(k_1, k_2, k_3) \quad (2.35)$$

with e.g.

$$F^{\text{local}}(k_1, k_2, k_3) = \frac{1}{k_1^{4-n_s}k_2^{4-n_s}} + 2 \text{ perm.} \quad (2.36)$$

In Fig. 10 we show the constraints as a function of the maximum multipole on the left using the following configuration $f_{\text{sky}} = 0.4$, T -noise = $1 \mu\text{K}^-$ and E -noise = $\sqrt{2} \mu\text{K}^-$ and a beam of $1'$ and $\ell_{\text{min}} = 30$. Local type non-Gaussianities benefit from large scales, and as much as 40% of the signal is lost if these modes are not available as can be seen in Fig. 10 on the right. Ideally, large scale information from Planck [97] should be included to put the best constraints on non-Gaussianities. Including Planck low ℓ (using $f_{\text{sky}} = 0.75$ [97] to determine the noise level, and $f_{\text{sky}} = 0.4$ for the maximal overlap) we can improve the forecasted bounds on

¹Note that there is a factor of $5/3$ in relating the gauge-invariant Newtonian potential Φ to the curvature ζ [115].

Type	Planck	CMB-S4	CMB-S4 + low ℓ Planck	rel. improvement
local	$\sigma(f_{\text{NL}}) = 3.1$	$\sigma(f_{\text{NL}}) = 2.8$	$\sigma(f_{\text{NL}}) = 1.8$	1.7
equilateral	$\sigma(f_{\text{NL}}) = 32.1$	$\sigma(f_{\text{NL}}) = 18.8$	$\sigma(f_{\text{NL}}) = 18.8$	1.7
orthogonal	$\sigma(f_{\text{NL}}) = 15.4$	$\sigma(f_{\text{NL}}) = 8.7$	$\sigma(f_{\text{NL}}) = 8.7$	1.8
flat	$\sigma(f_{\text{NL}}) = 26.1$	$\sigma(f_{\text{NL}}) = 15.2$	$\sigma(f_{\text{NL}}) = 15.2$	1.7

Table 2-2. Forecasted constraints on several well motivated non-Gaussian shapes using T and E modes. *Planck* forecast is based on Blue Book values, with $f_{\text{sky}} = 0.75$. The table shows we need to include low ℓ information from *Planck* for local type non-Gaussianities. *CMB-S4* is assumed to have $f_{\text{sky}} = 0.4$, T -noise = $1 \mu\text{K-}'$ and E -noise = $\sqrt{2} \mu\text{K-}'$ and a beam of $1'$ and $\ell_{\text{min}} = 30$.

local type non-Gaussianities by almost a factor of 2. Equilateral, enfolded and orthogonal non-Gaussianities are not affected by not including the lowest multipoles. We summarize the results in Tab. 2-2. Note that our forecast, using *Planck* Blue Book² values, deviate slightly from the actual bounds on non-Gaussianities obtained in Ref. [97]. The expected factor of improvement over *Planck*-only is somewhere between 1.7 and 1.8 for all shapes considered. Information saturates beyond $\ell_{\text{max}} = 4000$ for all shapes for an experiment with $1'$ beam.

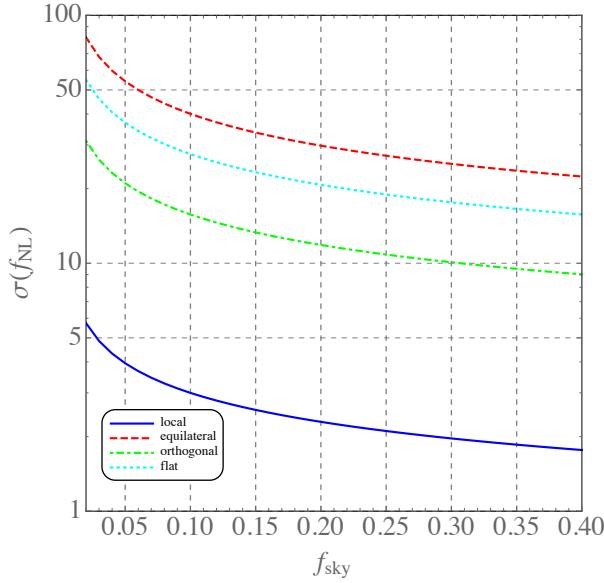


Figure 11. Fixed effort for local, equilateral, orthogonal and flat TTT bispectra with $1'$ beam.

Next, we determine the improvements as function of a fixed effort with $f_{\text{sky}} = 0.4$ corresponding to noise of $1 \mu\text{K-arcmin}$. The results are shown in Fig. 11 as a function of f_{sky} for a $1'$ beam. $\langle TTT \rangle$ generally benefits from large sky fraction which is a result of the fact that most modes are already cosmic variance limited (and hence information is saturated and does not benefit from a deeper patch).

Any one care to work this out more: Higher order statistics encode further information about particle content and interactions. The relative amplitude of certain limits of the trispectrum (the momentum space 4-point

²[http://www.rssd.esa.int/SA/PLANCK/docs/Bluebook-ESA-SCI\(2005\)1_V2.pdf](http://www.rssd.esa.int/SA/PLANCK/docs/Bluebook-ESA-SCI(2005)1_V2.pdf)

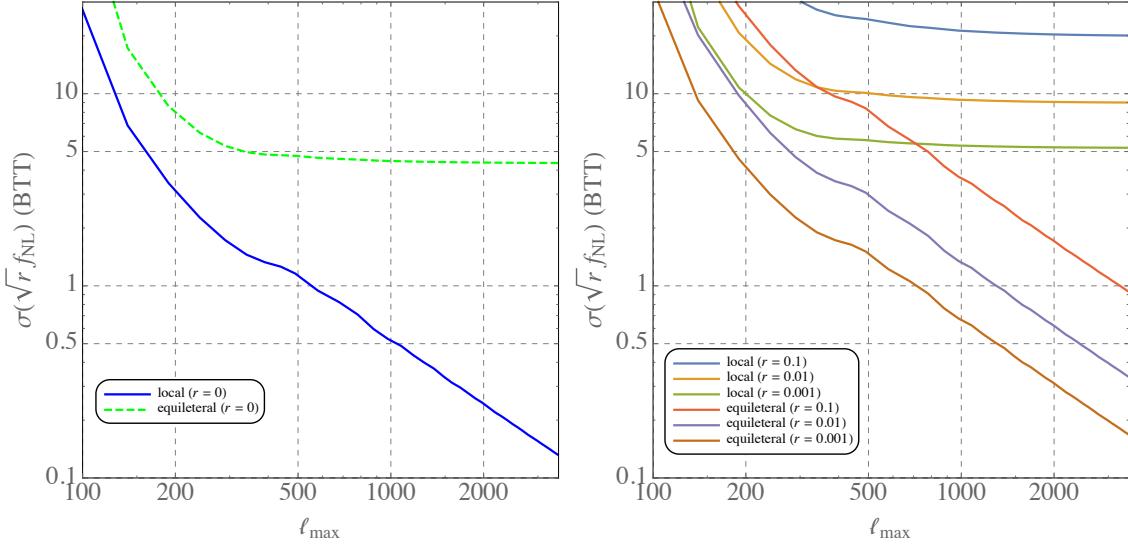


Figure 12. Left: Noise dominated B modes. Forecasts on the constraining power of CMB-S4 on two types of $\langle \gamma \zeta \zeta \rangle$ non-Gaussianities as a function of ℓ_{\max} using $\langle BTT \rangle$. Right: the effect of cosmic variance in B .

function) and of the trispectrum can also reveal whether there may be multiple sources contributing to the primordial fluctuations (and both may be different from the fluctuations of the inflaton). [more...](#)

2.8 Isocurvature

Measurements of CMB temperature/polarization power spectra indicate that the primordial initial conditions are adiabatic, that is, spatial entropy fluctuations vanish:

$$S_{i\gamma} \equiv \frac{\delta n_i}{n_i} - \frac{\delta n_\gamma}{n_\gamma} = 0. \quad (2.37)$$

The species label i can denote baryons, cold dark matter (CDM), or neutrinos. Number densities are denoted by n_i and perturbations in them by δn_i .

This can also be expressed in a gauge-invariant way for any two species i, j :

$$S_{ij} = 3(\zeta_i - \zeta_j) \quad (2.38)$$

where $\zeta_i = -\Psi - H \frac{\delta \rho_i}{\dot{\rho}_i}$.

Adiabatic perturbations are produced in models where the initial perturbations in all species are seeded by the inflaton. If fluctuations are also sourced by a second field, the initial conditions are a mixture of adiabatic and entropy (a.k.a isocurvature) perturbations, for which $S_{i\gamma} \neq 0$. These initial conditions determine the acoustic peak structure and large-scale amplitude of CMB anisotropies, as well as large-scale structure statistics [179, 180, 181, 182, 183, 184]. Observations can thus probe the number of fields during inflation. Each species can carry isocurvature perturbations in its density (e.g. Refs [185, 186, 183]).³ Indeed, the modes of the perturbation evolution equations correspond to adiabatic, CDM density isocurvature (CDI),

³Neutrinos can also carry velocity isocurvature, but this mode is not well motivated in inflationary models.

baryon density isocurvature (BDI), neutrino density isocurvature (NDI), and neutrino velocity isocurvature (NVI) initial conditions.

Data from WMAP [187], *Planck* [188], and other experiments [189, 190] indicate that perturbations are predominantly adiabatic. The limits can be stated in terms of the fractional primordial power in each isocurvature mode:

$$\beta \equiv \frac{P_{S_{i\gamma}}(k)}{P_{S_{i\gamma}}(k) + P_{\zeta\zeta}(k)}. \quad (2.39)$$

The current CMB limits (allowing one isocurvature mode at a time) are shown in Table ??, along with a forecast of CMB-S4 sensitivity.⁴ In Table. ??, “correlated” refers to totally correlated or anti-correlated (with ζ) isocurvature perturbations. All results are quoted at a fiducial $k = 0.05 \text{ Mpc}^{-1}$. Limits to BDI perturbations are not separately listed, as at linear order they are indistinguishable from CDI perturbations [191, 192, 193, 194]. CMB-S4 could improve on these limits by a factor of 2 – 5, as shown in Table. ??.

In this section we focus on two specific scenarios for isocurvature: the curvaton, compensated isocurvature perturbations (CIPs). Discussion of axion-type isocurvature is deferred to Section 5.1.2.

The curvaton scenario is an alternative to single-field inflationary models in which a sub-dominant second field σ acquires vacuum fluctuations during inflation, becomes more important later, sources ζ , and then decays [195, 196, 197, 159, 198]. Curvaton candidates include sneutrinos, string moduli, and others [199, 200, 201, 202, 203, 204, 205, 206]. Depending on whether a species i (or its quantum numbers) is produced by, before, or after curvaton decay, perturbations in i are offset from ζ , leading to isocurvature perturbations: [159, 198, 191]

$$S_{i\gamma} = \begin{cases} -3\zeta - 3(\zeta_\gamma - \zeta), & \text{if } i \text{ is produced before } \sigma \text{ decay,} \\ 3\left(\frac{1}{r_D} - 1\right)\zeta - 3(\zeta_\gamma - \zeta), & \text{if } i \text{ is produced by } \sigma \text{ decay,} \\ -3(\zeta_\gamma - \zeta), & \text{if } i \text{ is produced after } \sigma \text{ decay,} \end{cases}. \quad (2.40)$$

Here ζ_i is the density perturbation in i on surfaces of constant curvature. The parameter r_D is the fractional energy density in the curvaton when it decays.

The mixture of isocurvature modes is determined by whether or not baryon number, lepton number, and CDM are produced before, by, or after curvaton decay. Curvaton-type isocurvature is distinct from axion isocurvature, as it is correlated (or anti-correlated) ζ . If lepton number is produced by curvaton decay, the lepton chemical potential ξ_{lep} is important in setting the amplitude of NDI modes [198, 207, 208]:

$$S_{\nu\gamma} = -\frac{135}{7} \left(\frac{\xi_{\text{lep}}}{\pi}\right)^2 \zeta_\gamma. \quad (2.41)$$

There are 27 distinct curvaton decay scenarios, as baryon number, lepton number, and CDM could each be produced before, by, or after curvaton decay. Viable models are those in which one of baryon number or CDM is produced by curvaton decay, and those in which *both* baryon number and CDM are produced after curvaton decay. For curvaton-decay scenarios, we use the notation (b_x, c_y, L_z) , where $x \in (\text{before}, \text{by}, \text{after})$. Here b denotes baryon number, c denotes CDM, and L denotes lepton number. For example, $(b_{\text{before}}, c_{\text{by}}, L_{\text{by}})$ is a model in which baryon number is produced before curvaton decay, CDM by curvaton decay, and lepton number by curvaton decay.

Current isocurvature limits favor values of $r_D \simeq 1$, except for models in which baryon number is produced by curvaton decay and CDM before (or vice versa), which favor central values of $r_D \simeq 0.16$ ($r_D \simeq 0.84$).

⁴All forecasts in the isocurvature section are produced using Fisher-matrix techniques.

The current limits [209] to r_D are shown in Table 2-3, along with a forecast of CMB-S4's sensitivity to r_D via isocurvature. The dramatic improvement in the $(b_{\text{by}}, c_{\text{before}}, L_{\text{by}})$ and $(b_{\text{before}}, c_{\text{by}}, L_{\text{by}})$ scenarios because of the accompanying NDI perturbations. One unusual case is the $(b_{\text{after}}, c_{\text{after}}, L_{y_L})$ scenario. Here isocurvature just constrains the degenerate combination [209] $\chi_D \equiv \left\{ 1 + \xi_{\text{lep}}^2 / (\pi^2) (1/r_D - 1) \right\}^{-1}$, while the independent constraint to ξ_{lep}^2 is driven by the N_{eff} limit from the CMB.

Isocurvature scenario	<i>Planck</i>	CMB-S4
	$\Delta r_D/r_D^{\text{adi}}$	$\Delta r_D/r_D^{\text{adi}}$
$(b_{\text{by}}, c_{\text{before}}, L_{y_L})$	0.03	0.005
$(b_{\text{before}}, c_{\text{by}}, L_{y_L})$	0.01	0.004
$(b_{\text{by}}, c_{\text{after}}, L_{y_L})$	0.04	0.01
$(b_{\text{after}}, c_{\text{by}}, L_{y_L})$	0.008	0.002
$(b_{\text{by}}, c_{\text{by}}, L_{y_L})$	0.007	0.002
$(b_{\text{after}}, c_{\text{after}}, L_{y_L})$	$\Delta \chi_D/\chi^{\text{adi}}$	$\Delta \chi_D/\chi^{\text{adi}}$
	$\Delta \xi_{\text{lep}}^2$	$\Delta \xi_{\text{lep}}^2$
$(b_{\text{by}}, c_{\text{before}}, L_{\text{by}})$	0.02	0.002
$(b_{\text{before}}, c_{\text{by}}, L_{\text{by}})$	0.4	0.04
$(b_{\text{by}}, c_{\text{after}}, L_{\text{by}})$	0.3	0.04
$(b_{\text{after}}, c_{\text{by}}, L_{\text{by}})$	0.3	0.04
$(b_{\text{by}}, c_{\text{by}}, L_{\text{by}})$	0.3	0.04
$(b_{\text{after}}, c_{\text{after}}, L_{\text{by}})$	0.3	0.04

Table 2-3. Isocurvature constraints on r_D and ξ_{lep}^2 , both at (95% C.L.) using *Planck* TT+BAO+LowP data [209] in viable curvaton decay-scenarios, and Fisher forecasts for CMB-S4 sensitivity.

Depending on the scenario, forecasting shows that the S4 sensitivity to curvaton-sourced isocurvature should improve by a factor of 2–4 on current limits. In models with nearly-canceling CDM and baryon isocurvature perturbations, S4 limits to neutrino isocurvature drive an improvement in the sensitivity to the lepton asymmetry from $\Delta \xi_{\text{lep}}^2 \simeq 0.015$ to $\Delta \xi_{\text{lep}}^2 \simeq 0.003$. This dramatic improvement would make CMB limits comparably sensitive to BBN probes ξ_{lep}^2 (for this decay scenario).

If baryon number/CDM are produced by/before curvaton decay (or vice versa), a relative large compensated isocurvature perturbation (CIP) is produced between the baryons and CDM, that is

$$S_{bc} = \frac{\delta n_b}{n_b} - \frac{\delta n_c}{n_c} \neq 0. \quad (2.42)$$

Curvaton-generated CIPs are proportional to ζ , $S_{bc} = A\zeta$, where $A \simeq 17$ [$A \simeq -3$] in the $(b_{\text{by}}, c_{\text{before}}, L_z)$ [$(b_{\text{before}}, c_{\text{by}}, L_z)$] scenario. For CIPs, the initial relative densities of baryons and CDM vary, but with no additional overall matter or radiation density fluctuation.

CIPs are relatively unconstrained at the linear level of the CMB power-spectrum (see Ref. [210] for an exception), but would induce non-Gaussianities in the CMB [211, 212, 213, 214]. As with weak gravitational lensing [215], the CIP field $\Delta(\hat{n})$ can be reconstructed using CMB data. We find that at S4

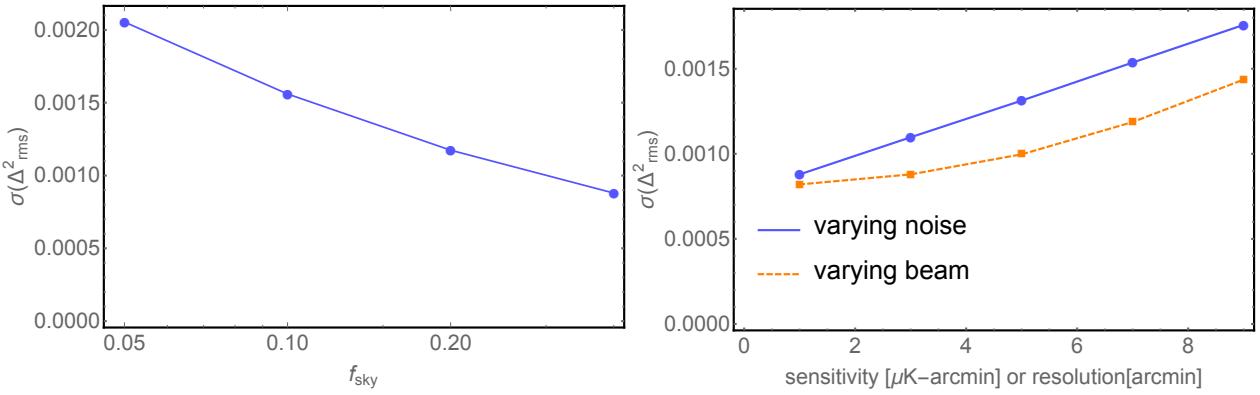


Figure 13. Constraints on compensated isocurvature perturbations as a function of observing area (left panel) and survey strategy (right panel). In changing the observed sky area, the noise is modified to maintain a fixed total observing time, while the survey parameters are varied independently in the right panel around a baseline of 1 $\mu\text{K-arcmin}$ sensitivity and 1 arcminute resolution.

sensitivity [214], the threshold for a 95% C.L. detection is $A \simeq 10$, and so a CIP test of the $(b_{\text{by}}, c_{\text{before}}, L_z)$ scenario is within reach of CMB-S4. This is a vast improvement over *Planck* sensitivity, which at 95% C.L. is $A \simeq 43$. Uncorrelated CIPs are less motivated theoretically. Updating the analysis of Ref. [214] with current parameters [188] and CMB-S4 specifications, we find that the sensitivity of CMB-S4 to a scale-invariant (SI) angular power spectrum of uncorrelated CIPs is $\Delta_{\text{cl}} = 0.003$ at the 95% C.L. detection limit. Here Δ_{cl} is the r.m.s. CIP amplitude on cluster scales. This is a vast improvement over the upper limit of $\Delta_{\text{cl}} \leq 0.077$ from WMAP [213], or the forecasted Planck [188] (including polarization) sensitivity of $\Delta_{\text{cl}} \leq 0.015$ [214].

A complementary constraint on uncorrelated CIPs can be derived from an independent search for their second-order effect on the CMB power spectrum [210]. With a CMB-S4 experiment, the power spectra would be sensitive to a CIP amplitude of $\Delta_{\text{cl}} = 0.026$ at the 95 % C.L., a factor of three better than the corresponding limit from the Planck analysis (see Ref. [210]). The CMB-S4 constraints on these models are shown in Figure 13.

2.9 Spatial Curvature

Despite the fact that inflation drives the spatial curvature to zero at the level of the background evolution, it predicts small, but non-zero curvature for a typical observer. The curvature measured in a Hubble patch receives contributions from long wavelength perturbations and is expected to be $|\Omega_k| < 10^{-4}$. Measurements of Ω_k exceeding this expectation would contain important information about the process responsible for inflation. In particular, if $|\Omega_k|$ is found to be considerably larger than this value, it would tell us that the inflaton was not slowly rolling when scales slightly larger than our observable horizon exited the horizon. Furthermore, observations of large negative Ω_k would falsify eternal inflation, while observation of positive and large Ω_k would be consistent with false vacuum eternal inflation [216, 217].

Current constraints on this parameter from the CMB alone are $\Omega_k = 0.005^{+0.016}_{-0.017}$. Including baryon acoustic oscillation (BAO) data tightens the bound to $\Omega_k = 0.000 \pm 0.005$.

Figure 14 demonstrates the Ω_k constraints achievable with CMB-S4 in isolation (solid yellow) and in combination with Planck (dashed yellow), future BAO measurements from DESI [218] (solid blue) and both (dashed blue). The constraining power of each configuration is explored as a function of CMB-S4s

beam size (left panel) and polarization noise (right panel); the noise and beam are fixed to $1 \mu\text{K}\text{-arcmin}$ and 3 arcmin in the left and right panels, respectively.

The constraints on Ω_k are relatively weakly dependent on the resolution and sensitivity of CMB-S4, inflating by 30-50% over the ranges considered. In isolation, CMB-S4 is able to place one-sigma limits of 6.7×10^{-3} on our fiducial value of zero curvature, assuming a white noise level of $1 \mu\text{K}\text{-arcmin}$ and 3 arcmin resolution. These limits are halved when including Planck data, demonstrating the importance of large-scale CMB polarization information in breaking degeneracies involving Ω_k (primarily with τ) [219]. CMB-only datasets are, however, fundamentally limited in their ability to constrain curvature by the geometrical degeneracy [220, 221]. Adding DESI measurements of the BAO standard ruler (at redshifts of 0 to 1.9) helps break this degeneracy, with a combined one-sigma limit of around 7.1×10^{-4} . Though CMB-S4 is therefore unable to measure curvature levels typical of slow-roll eternal inflation, any detection of curvature by CMB-S4 would have profound implications for the inflationary paradigm.

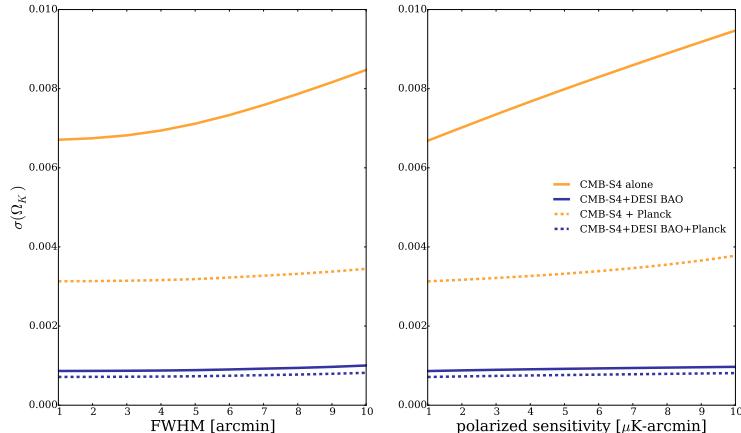


Figure 14. Results derived using the CMB4cast framework [222], assuming CMB-S4 observations with a single 150GHz channel, at $1 \mu\text{K}\text{-arcmin}$ ($1.4 \mu\text{K}\text{-arcmin}$) in temperature (polarization), with a FWHM of 3 arcmin . The instrument is assumed to observe $f_{\text{sky}} = 40\%$, covering a multipole range of $30 \leq \ell \leq 4000$. We consider no foregrounds, white noise and a single-parameter extension of the standard model: $\Lambda\text{CDM}+\Omega_k$. The constraints are derived using information from unlensed TT, TE and EE spectra along with $\kappa\kappa$ information constructed with an iterative EB quadratic estimator [223, 224]. Cases including Planck use additional TT, TE and EE information from Planck, assuming multipole coverage of $30 \leq \ell \leq 2500$ over an additional 20% of sky. A prior on τ is used: 0.06 ± 0.01 . We follow [225] for the combinations with DESI-BAO.

2.10 Microwave Background Anomalies

Several unexpected features have been observed in the temperature of the CMB sky at relatively low- ℓ or large angular scales. Some of these were first noticed in COBE data, and all have been seen in both WMAP and Planck maps. These include:

- a lack of correlation on the largest angular scales;
- alignment of the lowest multipole moments with one another and with the geometry and motion of the Solar System;

- greater power in odd-parity modes than in even-parity ones;
- a hemispherical asymmetry or dipolar modulation of the power.

Compared to the expectations of the best-fit inflationary Λ CDM model, the individual p-values of these features are in the per mille to per cent level, and therefore each anomaly has a frequentist probability at approximately the 3-sigma level or higher. Since certain pairs of anomalies are uncorrelated in Λ CDM, in combination they nominally represent a very significant detection of anomalous behaviour.

There are however two possible concerns before one can conclude that the CMB large-angle pattern is truly anomalous. First, these features were identified a posteriori and are characterized by statistics that were devised after the anomalies were first noted. Second, there is no physical understanding of how the collection of such features could arise. In order to help address or resolve these concerns, it is therefore crucial to obtain additional information about the large-scale primordial fluctuations, and to devise a successful model or other explanation.

The observed features can have two possible origins: either our cosmological model is incomplete and requires a modification (the new physics hypothesis), or we just happen to live in a realization of that model that is statistically unlikely (the fluke hypothesis). Meanwhile, cosmologists have effectively exhausted their ability to obtain further independent CMB temperature data that can test these anomalies, as observations are already cosmic-variance limited at the relevant angular scales.

It has been suggested that one may nevertheless make observational progress even in the absence of an alternative model. This can occur in two ways:

1. In the fluke hypothesis the conditional probability distributions of Λ CDM for correlation functions of CMB polarization (and other observables) with CMB temperature and with one another are altered by the observed temperature anomalies (Dvorkin et al 2008, Copi et al 2013, Yoho et al. 2013).
2. In the new physics hypothesis, a given phenomenological model that explains the anomalies will have observational consequences for other observable quantities [Yoho et al. 2015]. For example, the absence of large-angle correlations in T may reflect a lack of long-distance correlation in a fundamental physical quantity like the potential; similarly, a hemispherical asymmetry in TT power could cause a similar asymmetry in EE.

A variety of ideas have been proposed to explain the anomalies, ranging from Solar system dust artifacts to anisotropic models of inflation (for a summary, see Copi et al, 2016). Unfortunately, none of those ideas leads to a convincing explanation, as it is simply difficult to find models that explain the alignments of the largest primordial structures in the universe while at the same time lowering the amplitude of large-angle temperature correlations (e.g. Gordon et al, 2005).

Additional information from polarization would be of great help. To address the first three anomalies we list above would probably require a space mission, due to the need to access very low ℓ . But CMB-S4 can shed light on the fourth one.

CMB-S4, alone or in combination with other data, can begin to explore both the cosmological and the fluke explanation for the hemispherical power asymmetry. For example, Λ CDM instructs us how to remove the part of the E-mode signal that is correlated to temperature; the remainder should be Gaussian random and statistically isotropic. If it contains a hemispherical anomaly (especially one aligned with the temperature asymmetry), that would be evidence against the fluke hypothesis [Copi, Knox, ODwyer and Starkman, contribution to March 9/10 S4 meeting, in preparation].

2.11 Summary

CMB-S4 is an ideal tool to test the inflationary paradigm and competing theories of the early universe. On the one hand, its exquisite sensitivity will allow a detection of degree scale B-modes in the CMB or achieve upper limits on the amount of B-mode polarization that improve current constraints on the tensor-to-scalar ratio by almost two orders of magnitude. In particular, it is sensitive enough to detect the level of B-mode polarization predicted in a wide range of well-motivated inflationary models. In doing so, it would provide invaluable information about physics at energy scales far outside the reach of any terrestrial particle physics experiment. In the absence of a detection it would exclude large classes of inflationary models. On the other hand, with sufficient angular resolution, it will measure anisotropies in both the temperature and E-mode polarization of the CMB to cosmic variance over the entire range of multipoles that is not contaminated by unresolved foregrounds, and it will extend our window to the early universe by almost one e -fold beyond the reach of current experiments. As a consequence, it will provide the best constraints achievable by any ground-based CMB experiment on any observable that benefits from the number of modes measured, such as the primordial power spectrum, and hence the spectral index, its running, features in the power spectrum, as well as higher order correlations.

3

Neutrinos and Light Species from the Cosmic Microwave Background

(send feedback on this chapter to s4_neutrinos@cosmo.uchicago.edu)

3.1 Introduction

Direct interactions between neutrinos and observable matter effectively ceased about one second after the end of inflation. Nevertheless, the total energy density carried by neutrinos was comparable to other components through recent cosmological times. As a result, the gravitational effect of the neutrinos is detectable both at the time of recombination and in the growth of structure at later times [226], leaving imprints in the temperature and polarization spectrum as well as in CMB lensing.

CMB-S4 can improve our understanding of neutrino physics in regimes of interest for both cosmology and particle physics. Arguably the most important parameters of interest will be the sum of the neutrino masses ($\sum m_\nu$) and the effective number of neutrino species (N_{eff}). These two parameters have natural targets that are within reach of a CMB-S4 experiment:

- $\sum m_\nu \gtrsim 58$ meV is the lower bound guaranteed by observations of solar and atmospheric neutrino oscillations. A CMB experiment with $\sigma(\sum m_\nu) < 20$ meV would be guaranteed a detection of at least 3σ . At this level, one can detect the overall scale of the neutrino masses even for a normal ordered hierarchy.
- $\Delta N_{\text{eff}} \geq 0.027, 0.047,$ and 0.054 are predicted for models containing additional light particles of spin 0, 1/2 and 1 that were in thermal equilibrium with the particles of the Standard Model. A CMB experiment reaching $\sigma(N_{\text{eff}}) \lesssim 0.01$ would be sensitive to all models in this very broad class of extensions of the Standard Model, which includes a wide range of models predicting axions and axion-like particles. At lower sensitivity ($\sigma(N_{\text{eff}}) \lesssim 0.03$), we can test the full range of models containing light fermions and vectors, including gravitinos and dark photons that freeze-out before the QCD phase transition.

Current CMB data already provides a robust detection of the cosmic neutrino background at $\sim 10\sigma$. A CMB-S4 experiment will provide an order of magnitude improvement in sensitivity that opens a new window back to the time of neutrino decoupling and beyond.

Section 2 will review the motivation for studying neutrino masses with cosmological probes, and specifically with the CMB. We will explain why cosmology is sensitive to $\sum m_\nu$ via different probes and how it is complementary to the laboratory-based experimental neutrino effort. Section 3 will review the physics of N_{eff} and its role as probe of the CMB and as sensitive tool for beyond the Standard Model physics. We will emphasize the unique impact N_{eff} has on the CMB that makes it distinguishable from other extensions of Λ CDM. In Section 4, we will discuss the implications for a variety of well motivated models, including sterile neutrinos and axions. In Section 5, we will discuss the relation between CMB and BBN based constraints.

3.2 Neutrino Mass

3.2.1 Theory Review

Cosmic background neutrinos are nearly as abundant in the universe as CMB photons. In the standard cosmological model, neutrinos ceased to scatter with other particles at temperatures $\sim 1 \text{ MeV}$. The relic neutrinos were relativistic at decoupling, but as the universe expanded and cooled the neutrino momenta redshifted as $p_\nu \propto 1/a$ and eventually the energy of most relic neutrinos came to be dominated by their rest mass, rather than their momentum. The energy density in nonrelativistic neutrinos therefore contributes to the matter budget of the universe today. The neutrinos, however, were relativistic for much of the history of the Universe so their gravitational clustering is qualitatively different from that of cold dark matter (CDM) particles. This difference can be used to distinguish the neutrino and cold dark matter contributions to the matter density [227, 228, 229]. In this section, we review how neutrino mass affects the evolution of the neutrino energy density and the gravitational clustering of matter in the universe.

As discussed more detail in the N_{eff} section, cosmic background neutrinos have been detected indirectly through their contribution to the energy density in radiation in the early universe. The current CMB constraints from N_{eff} are in excellent agreement with the Standard Model expectation of three species of neutrinos and anti-neutrinos each described by a relativistic thermal Fermi-Dirac distribution [230]. The distribution function for each species of neutrinos and anti-neutrinos is given by

$$f_\nu(p) = \frac{1}{e^{ap/(k_B T_{\nu 0})} + 1}, \quad (3.1)$$

where $T_{\nu 0} \approx 1.95K$ or $k_B T_{\nu 0} \approx 1.68 \times 10^{-4} \text{ eV}$ is the temperature today. Note that the spectral shape of the neutrino phase space distribution is preserved with the expansion of the universe so relic neutrinos have retained the relativistic Fermi-Dirac momentum distribution inherited from decoupling even as the individual neutrinos became non-relativistic.

The neutrino energy density is given by

$$\rho_\nu = \sum_i \int \frac{d^3 \mathbf{p}}{(2\pi\hbar)^3} \frac{\sqrt{p^2 + m_{\nu i}^2}}{e^{ap/(k_B T_{\nu 0})} + 1} \quad (3.2)$$

where $m_{\nu i}$ are the three neutrino mass eigenstates. For $T_\nu/a \gg m_{\nu i}$ the neutrino energies are dominated by their momenta and the total energy density behaves like radiation

$$\begin{aligned} \rho_\nu \Big|_{\text{early}} &\approx \frac{7\pi^2}{40} \frac{(k_B T_{\nu 0})^4}{\hbar^3 c^3} \frac{1}{a^4} \\ &\propto a^{-4} \end{aligned} \quad (3.3)$$

While for $T_{\nu 0}/a \ll m_{\nu i}$ the energy density behaves like matter

$$\begin{aligned} \rho_\nu \Big|_{\text{late}} &\approx \sum_i m_{\nu i} \bar{n}_\nu \\ &\propto a^{-3} \end{aligned} \quad (3.4)$$

where \bar{n}_ν is the number of neutrinos and antineutrinos in each mass eigenstate

$$\bar{n}_\nu = \int \frac{d^3 \mathbf{p}}{(2\pi\hbar)^3} \frac{2}{e^{ap/(k_B T_{\nu 0})} + 1} \approx \frac{113}{a^3} \text{ cm}^{-3}. \quad (3.5)$$

For a neutrino of mass $m_{\nu i}$ the transition between these two regimes ($T_\nu(a) \sim m_{\nu i}$) occurs at redshift $z_{\text{nr}} \sim 300(m_{\nu i}/0.05\text{eV})$. Using Eq. (3.4) the fractional energy density in neutrinos today can be written as

$$\Omega_\nu h^2 \approx \frac{\sum_i m_{\nu i}}{93 \text{ eV}}. \quad (3.6)$$

The individual masses of the neutrino states are unknown but neutrino oscillation data specifies the square of two mass splittings $\Delta m_{12}^2 = 7.54 \times 10^{-5} \text{ eV}$, $|\Delta m_{13}^2| \approx 2.4 \times 10^{-3} \text{ eV}$ [231]. These mass splittings, in combination with the neutrino number density, give a lower limit on the contribution of neutrinos to the cosmic energy budget

$$\Omega_\nu h^2 \gtrsim 0.0006. \quad (3.7)$$

At $z \ll z_{\text{nr}}$ the matter density of the universe, which enters into the Hubble equation, is the sum of the CDM, baryon, and massive neutrino energy densities $\Omega_m = \Omega_c + \Omega_b + \Omega_\nu$. Whereas, at $z \gg z_{\text{nr}}$ the matter density is solely made up of the baryon and CDM parts while neutrinos contribute to the radiation density.

Neutrinos do not participate in gravitational collapse until late times when they have become nonrelativistic. Prior to this transition, the neutrinos *free-stream* out of gravitational wells, leaving the CDM and baryons behind [232, 233, 234, 227]. Primordial fluctuations in the neutrino density are therefore damped away on scales smaller than the horizon at z_{nr} . In comoving units, this scale corresponds to a wave number

$$k_{\text{nr}} \equiv a_{\text{nr}} H(a_{\text{nr}})/c \approx 0.003 \left(\frac{\Omega_m}{0.3} \frac{m_\nu}{0.05 \text{ eV}} \right)^{1/2} h/\text{Mpc}. \quad (3.8)$$

Once the neutrinos are non-relativistic, their finite velocity dispersion still prevents them from clustering on scales smaller than the typical distance a neutrino travels in a Hubble time, $v_\nu/H(a)$ where $v_\nu \approx 3.15 T_{\nu 0}/(am_\nu)$ the mean neutrino velocity. In analogy with the Jeans criterion for gravitational collapse, the neutrino free-streaming scale is defined by [232, 228]

$$k_{\text{fs}}(a) \equiv \sqrt{\frac{3}{2} \frac{aH(a)}{v_\nu(a)}} \approx 0.04 a^2 \sqrt{\Omega_m a^{-3} + \Omega_\Lambda} \left(\frac{m_\nu}{0.05 \text{ eV}} \right) h/\text{Mpc} \quad (3.9)$$

in comoving coordinates.

On scales larger than k_{nr} (adiabatic) perturbations in the density of neutrinos, baryons, and CDM are coherent and can be described by a single perturbation to the total matter density $\delta_m = \delta\rho_m/\rho_m$. On smaller scales where the neutrino perturbations have decayed, only the perturbations to the CDM and baryons remain so that $\delta_m = \delta_{cb} (\Omega_c + \Omega_b)/\Omega_m$. The remaining CDM and baryon perturbations also grow more slowly because the neutrino energy density contributes to the expansion rate, but not to the source potentials. These two effects cause a suppression in the amplitude and the growth rate of matter perturbations wavenumbers $k > k_{\text{fs}}$ relative to a universe with massless neutrinos (and also relative to density perturbations with $k < k_{\text{nr}}$). The net change in the amplitude of perturbations with $k > k_{\text{nr}}$ primarily depends on the fractional energy density in massive neutrinos (keeping $\Omega_c + \Omega_b$ fixed) but retains a small sensitivity to the individual neutrino masses through a dependence on a_{nr} .

An estimate of the effect of massive neutrinos on the growth of structure can be made by studying the evolution of matter perturbations in the two regimes $k \ll k_{\text{fs}}$ and $k \gg k_{\text{fs}}$. In the synchronous gauge, linear perturbations to the matter density with wavenumbers $k \ll k_{\text{fs}}$ evolve as

$$\ddot{\delta}_m + 2H(a)\dot{\delta}_m - \frac{3}{2}\Omega_m H_0^2 a^{-3} \delta_m = 0 \quad \text{for } k \ll k_{\text{nr}} \quad (3.10)$$

which has solutions $\delta_m \propto a, a^{-\frac{3}{2}}$ during the matter dominated era.

On scales where the neutrino perturbations have decayed, perturbations to matter density are just in the CDM and baryon components

$$\delta_m(k \gg k_{\text{fs}}) \approx (\delta\rho_c + \delta\rho_b)/\rho_m = (1 - f_\nu)\delta_{cb} \quad (3.11)$$

where $f_\nu = \Omega_\nu/\Omega_m$ and $\delta_{cb} = (\delta\rho_c + \delta\rho_b)/(\rho_c + \rho_b)$, but the neutrino energy density still contributes to the Hubble friction. In this limit, linear perturbations to the CDM and baryon density evolve as

$$\ddot{\delta}_m + 2H(a)\dot{\delta}_m - \frac{3}{2}\Omega_{cb}H_0^2a^{-3}\delta_m = 0 \quad \text{for } k \gg k_{\text{fs}} \quad (3.12)$$

where $\Omega_{cb} = \Omega_c + \Omega_b$ and $\Omega_{cb} < \Omega_m$ for a cosmology with massive neutrinos. Equation (3.12) has the approximate solutions during the matter dominated era of $\delta_{cb} \propto a^{1-\frac{3}{5}f_\nu}, a^{-\frac{3}{2}+\frac{3}{5}f_\nu}$ for $f_\nu \ll 1$.

The matter dominated solutions give a simple estimate of the net effects of massive neutrinos on the amplitude of matter perturbations. For fixed Ω_ch^2 , the evolution of perturbations in a cosmology with $f_\nu \neq 0$ is the same as a cosmology with $f_\nu = 0$. After a_{nr} , the perturbations with $k \gg k_{\text{fs}}$ grow more slowly (according to Eq. (3.12), the growing mode solution grows as $\propto a^{1-\frac{3}{5}f_\nu}$) than those with $k \ll k_{\text{fs}}$ (according to Eq. (3.10), $\propto a$). At scale-factor a during the matter dominated era, the total difference in growth or perturbations with $k \gg k_{\text{fs}}$ is roughly

$$\frac{\delta_{cb}(k \gg k_{\text{fs}}, a|f_\nu)}{\delta_{cb}(k \gg k_{\text{fs}}, a|f_\nu = 0)} \sim \left(\frac{a}{a_{\text{nr}}}\right)^{-\frac{3}{5}f_\nu}. \quad (3.13)$$

The resulting difference in the amplitude of the matter power spectra is then

$$\frac{P_{mm}(k \gg k_{\text{fs}}, a|f_\nu)}{P_{mm}(k \gg k_{\text{fs}}, a|f_\nu = 0)} \sim (1 - 2f_\nu) \frac{P_{cc}(k \gg k_{\text{fs}}, a|f_\nu)}{P_{cc}(k \gg k_{\text{fs}}, a|f_\nu = 0)} \sim \left(1 - 2f_\nu - \frac{6}{5}f_\nu \ln(a/a_{\text{nr}})\right). \quad (3.14)$$

On the other hand, the evolution of the large scale modes is identical,

$$\frac{P_{mm}(k \ll k_{\text{fs}}, a|f_\nu)}{P_{mm}(k \ll k_{\text{fs}}, a|f_\nu = 0)} = 1. \quad (3.15)$$

The above expression overestimates the effect of neutrino mass by assuming the transition from relativistic to non-relativistic is instantaneous. It also ignores the effects of the cosmological constant at late times. Using the true evolution of δ_{cb} through a_{nr} and allowing for the cosmological constant gives

$$\frac{P_{mm}(k \gg k_{\text{fs}}|f_\nu)}{P_{mm}(k \gg k_{\text{fs}}|f_\nu = 0)} \approx 1 - 6f_\nu \quad (3.16)$$

at $a = 1$. Note that this expression assumes fixed Ω_ch^2, Ω_bh^2 so that matter-radiation equality is not changed by neutrino mass and that $\Omega_\Lambda = 1 - \Omega_m$ is fixed by adjusting h so that the onset of cosmological constant domination is also unchanged. Alternatively, assuming fixed Ω_m and decreasing Ω_{cb} to account for Ω_ν makes matter-radiation equality, which occurs while the neutrinos are relativistic, slightly later so that the suppression is increased to

$$\frac{P_{mm}(k \gg k_{\text{fs}}|f_\nu)}{P_{mm}(k \gg k_{\text{fs}}|f_\nu = 0)} \approx (1 - 8f_\nu). \quad (3.17)$$

3.2.2 Observational Signatures and Target Sensitivity

As explained in the previous section, the signature of massive neutrinos manifests through the energy density Ω_ν , which is related to the mass through

$$\Omega_\nu h^2 \simeq \frac{\sum m_\nu}{93 \text{ eV}} \gtrsim 0.0006. \quad (3.18)$$

The lower limit on $\Omega_\nu h^2$ is a reflection of the lower limit on the sum of the masses, $\sum m_\nu \gtrsim 58 \text{ meV}$, that is determined from neutrino oscillation experiments[Citation needed]. This sets a clear observational target for future observations.

Any probe of P_{mm} late times is, in principle, sensitive to the sum of the neutrino masses. The question we will be most interested in is whether a given probe is sensitive to the lower limit, $\sum m_\nu = 58 \text{ meV}$ (or $\Omega_\nu h^2 = 0.0006$) under realistic circumstances. In this subsection, we will discuss the two methods through which CMB Stage IV can directly constrain the neutrino mass, CMB lensing and SZ cluster abundances. We will also compare these observables to other cosmological probes of the neutrino mass from upcoming large scale structure surveys such as DESI and LSST.

3.2.3 CMB Lensing

Likely the cleanest probe of the neutrino mass in the CMB is through gravitational lensing [235], which directly measures the matter distribution along the line of sight. To be concrete, in the Limber approximation, the lensing power spectrum is given by

$$C_\ell^{\phi\phi} = \frac{8\pi^2}{\ell^2} \int_0^{\chi_*} \chi d\chi P_\Psi(\ell/\chi; \eta_0 - \chi) \frac{(\chi_* - \chi)^2}{\chi_* \chi} \quad (3.19)$$

$$P_\Psi(k; \eta) = \frac{9\Omega_m^2(\eta) H^4(\eta)}{8\pi^2} \frac{P_{mm}(k; \eta)}{k} \quad (3.20)$$

where χ (χ_*) is the co-moving distance (to the last scattering surface) and η (η_0) is conformal time (today). More details regarding CMB lensing, including current and future measurements, will be discussed in Chapter **.

For the purposes of the neutrino mass measurement, the advantage of lensing over other probes is that it is largely free of astrophysical uncertainties. As we see from the lensing power spectrum, we are directly sensitive to the matter power spectrum (rather than a biased tracer) and the relevant scales are in the linear regime where modeling should be reliable.

The primary challenges for the lensing measurement are degeneracies with other cosmological parameters. The two primary degeneracies in Λ CDM are

- Optical depth, τ : The suppression of small scale power at low redshift requires a reliable measurement of the amplitude of the power spectrum at high redshift. In principle, this is measured by the primary CMB anisotropies, but the overall normalization is degenerate with τ for $\ell \gtrsim 20$. A precise measurement of τ is therefore crucial to calibrate the suppression at low redshifts. Such a measurement will likely come from $\ell \lesssim 20$ polarization data from CMB-S4 and/or other CMB experiments. It should be emphasized that Stage IV sensitivity is not needed for the measurement of τ and such a measurement could be performed by Planck or a Stage III experiment.
- $\Omega_m h^2$: The amount of lensing is controlled by the total amount of matter. Therefore, we can compensate for a suppression from neutrinos by increasing the matter power spectrum. This degeneracy will be broken by DESI BAO measurements of the expansion history.

In addition to degeneracies in Λ CDM there can be degeneracies with possible extensions. Most notably:

- N_{eff} : The density of neutrinos after they become non-relativistic is given by $\rho_\nu \simeq m_\nu n_\nu$ where n_ν is the number density. Therefore, we only measure the mass if we know the number density to sufficient

accuracy. Fortunately, as we will discuss in the next section, measurements of the neutrino energy density from the primary CMB will be sufficiently accurate as to make this degeneracy insignificant under plausible assumptions.

In principle, measurement of the free streaming scale directly in the matter power spectrum would separate the neutrino mass from most other physical quantities. Unfortunately, given current limits on the neutrino mass, the change to the shape of the lensing potential power spectrum is not expected to drive future constraints.

Status of current observations – Planck has provided a strong constraint of $\sum m_\nu < 0.194$ eV when combining both temperature and polarization data with the CMB lensing power spectrum and external data. A weaker constraint of $\sum m_\nu < 0.492$ eV can be derived using only the temperature and polarization data. This constraint arises through the effect of massive neutrinos on the primordial TT and EE power spectra. For sufficiently large masses, the neutrinos do not behave as radiation around the time of recombination which impacts the damping tail and locations of the acoustic peaks. Improvements in the limits on the sum of the neutrino masses will be driven primarily by lensing given that current limits imply that the neutrinos are effectively massless from the point of view of the primary CMB anisotropies. External data (BAO) will continue to be important in breaking the degeneracy with Ω_m .

3.2.4 Other Cosmological Probes

3.2.4.1 Galaxy Cluster Abundance

Galaxy clusters form from rare high peaks in the matter density field. A galaxy cluster of mass M forms from a region of size $R \sim (M/(4/3\pi\bar{\rho}_m))^{1/3}$, which is smaller than the neutrino free streaming scale for even the most massive galaxy clusters so long as $m_{\nu i} \lesssim 0.1$ eV. The neutrino free-streaming therefore slows the growth of structure on cluster scales, suppressing the abundance of galaxy clusters.

The number density of clusters with mass M can be expressed by (e.g. [236, 237]),

$$\frac{dn}{dM}(M, z) = \frac{\rho}{M} \frac{d \ln \sigma^{-1}}{dM} f(\sigma, z) \quad (3.21)$$

where $\sigma = \sigma(M, z)$ is the variance of linear perturbations in CDM and baryons on mass scale M given by

$$\sigma^2(M, z) = \int \frac{dk}{k} \frac{4\pi}{(2\pi)^3} P_{cc}(k, z) |W(kR)|^2 \quad (3.22)$$

where $R = (3M/(4\pi\rho_{cb}))^{1/3}$, $P_{cc}(k)$ is the power spectrum of CDM and baryons, and $W(kR) = 3(\sin(kR)/(kR)^3 - \cos(kR)/(kR)^2)$ is a top-hat window function [238, 239]. The cluster abundance is extremely sensitive to $\sigma(M, z)$, and therefore M_ν via the suppression in P_{cc} discussed in §3.2.1.

Current constraints on neutrino mass from cluster abundance, in combination with the primary CMB and BAO, are $M_\nu \lesssim 0.2$ – 0.3 eV at 95% confidence [240, 241, 242, 243]. To date, the constraints have been driven by the difference between (or consistency of) the matter power spectrum amplitude measured at late times, from clusters, and at early times, from the CMB. An additional signal is present internally to the cluster data, namely the time-dependent influence of massive neutrinos on cluster growth through Eqs. 3.21 and 3.22, which can potentially provide tighter constraints [244].

Making these measurements of cluster abundance and growth require cluster surveys with well understood selection functions extending to high redshift ($z \sim 2$). Galaxy clusters can be identified from CMB data via the thermal Sunyaev-Zel'dovich (tSZ) effect, the frequency shift of CMB photons that have scattered off of electrons in the hot, intra-cluster gas. CMB S4 is projected to detect a nearly mass-limited sample of $\mathcal{O}(100,000)$ galaxy clusters.

Given such a survey, the primary systematic limitation on cluster measurements of the neutrino mass, through either of the approaches above, is cluster mass estimation. This can be usefully broken into two problems: that of absolute mass calibration, i.e. our ability to measure cluster masses without bias on average, and relative mass calibration. The former is required for accurate inference of the power spectrum amplitude, while the latter significantly boosts cosmological constraints by providing more precise measurements of the shape and evolution of the mass function. Precise relative mass information can be provided by X-ray observations of the intracluster medium using existing (*Chandra*, *XMM-Newton*) and future (eROSITA, ATHENA) facilities; for the most massive SZ-selected clusters, found in existing CMB surveys, this work is already well advanced [243, 245].

For absolute mass calibration, the most robust technique currently is galaxy-cluster weak lensing, which (with sufficient attention to detail) can provide an unbiased results [246, 247]. At redshifts $z \lesssim 1$, residual systematic uncertainties in the lensing mass calibration are at the $\sim 7\%$ level currently [248], and reducing these systematics further is the focus of significant effort in preparation for LSST and other Stage 4 data sets, with 1–2% mass calibration at low redshifts seen as an achievable goal [249]. Lensing of the CMB by clusters has also been measured [250, 251], and provides an additional route to absolute mass calibration. CMB-cluster lensing is particularly well suited to calibrating clusters at high redshifts ($z \gtrsim 1$) where ground-based galaxy-cluster lensing becomes less efficient. CMB data with sufficient resolution and depth (especially in polarization) can potentially provide a percent-level mass calibration at these redshifts [252], comparable to galaxy-cluster lensing at lower redshifts.

As with dark energy studies, there are strong synergies between the SZ cluster catalog and CMB-cluster lensing information that CMB S4 can provide and external cluster surveys. The complementarity is particularly strong with LSST, which will provide highly mass-complete clusters catalogs out to redshifts ~ 1.2 and competitive lensing measurements for clusters at $z \lesssim 1$. In contrast, CMB S4 will cleanly select the most massive clusters at all redshifts of interest for cluster cosmology through the SZ effect, and CMB lensing data can provide the key mass calibration for clusters at the highest redshifts detected by either survey. The latter especially provides the longest possible lever arm for detecting the time-dependent impact of neutrino mass on the growth of clusters. Because their systematic uncertainties are not identical, the combination of galaxy-cluster and CMB-cluster lensing information can straightforwardly produce tighter constraints on the power spectrum amplitude than either alone. The prospects for Stage 4 cluster data sets to definitively detect the neutrino mass are thus strong [241, 244].

3.2.4.2 CMB Measurements in Context with Other Datasets

Current and future large-scale structure surveys, such as BOSS, DES, DESI, LSST, Euclid, and WFIRST, provide maps of the distribution of mass and galaxies in the late universe¹. Large-scale structure datasets are primarily sensitive to the neutrino mass scale via two means: (i) the suppression of the matter power spectrum, which can be inferred from weak gravitational lensing [253], fluctuations in the number of galaxies [254, 255], or fluctuations in the opacity of intervening gas [256], for example, and (ii) the change in the growth rate of matter perturbations which is inferred from redshift-space distortions (RSD)[257]. The first effect, the

¹While these surveys measure structure by a variety of means (the distribution of galaxies or quasars, weak gravitational lensing, and the Lyman- α forest, for example) we refer to all of them as galaxy surveys.

Datasets	σ_{M_ν} (eV)
Planck + DES lensing and galaxy clustering	0.041
Planck + DESI Lyman- α Forest + BAO	0.098
Planck + DESI Galaxy Power spectrum + BAO	0.024
Planck + LSST Lensing and Galaxy Clustering	0.02

Table 3-1. Forecasted constraints on neutrino mass from future galaxy surveys in combination with Planck CMB from [218]

suppression in the matter power spectrum, is the same effect tested by CMB lensing. The primary qualitative difference between information from galaxy surveys and the CMB, is that galaxy surveys measure structure at multiple epochs in cosmic history whereas the CMB provides a map of the integrated mass distribution out to the surface of last scattering. The LSS information content in galaxy surveys is therefore greater than the CMB, but interpreting the data can be considerably more complex because the structure measured in galaxy surveys is typically more nonlinear and the relationship between the galaxy and mass distribution is less well-understood (In fact, massive neutrinos can make this even more complicated [258]). Importantly, many of the observational and astrophysical systematics in the CMB and galaxy surveys are different, so the two approaches to measuring the neutrino mass scale are complementary. A summary of the forecasted constraints on M_ν from galaxy surveys is given in Table 3-1 [218].

There are a number synergistic opportunities between CMB S4 and external cosmological datasets. For instance, the large-scale structure measured by galaxy surveys gravitationally lenses the CMB so there is a physical correlation between the CMB lensing convergence and e.g. the galaxy distribution or weak lensing shear maps inferred from galaxy surveys. Constraints on neutrino mass can therefore be tightened by cross-correlating maps of structure from galaxy surveys with the lensing information from the CMB (e.g. [259, 260]). Or, the CMB can be cross-correlated with galaxy survey data to constrain neutrino mass via the mean pair-wise momentum of galaxy clusters [261]. Additionally, CMB S-4 data indirectly aid measurements of neutrino mass from galaxy surveys because the CMB data can be used to calibrate systematics in weak lensing shear data [262].

3.2.5 Forecasts

As we described in Section 3.2.3, the most concrete constraints on $\sum m_\nu$ from CMB Stage IV are derived from the CMB lensing power spectrum, as seen in Equation (3.19). Our sensitivity to $\sum m_\nu$ is limited by the error in the reconstruction of ϕ from the T or E/B maps. The noise for the lensing reconstruction is discussed in detail in Section 6.1 and the noise curves for both the TT - and EB -estimators are shown in Figure 38.

Forecasts for CMB Stage IV with and without DESI BAO are shown in Figure 15, following the methodology outlined in Section 7.10. As we explained in Section 3.2.3, the main signature neutrino mass in CMB lensing is degenerate with both τ and $\Omega_m h^2$. The sensitivity is therefore strongly dependent on the constraints on these parameters both internally and with external data. We can see the effect of the measurement of $\Omega_m h^2$ by comparing the results with and without DESI BAO. In particular, we note that the constraints improve by a factor of three by including DESI.

We have assumed that $\ell \geq 30$ and therefore we do not constraint τ directly. We instead assume an external prior of $\tau = 0.06 \pm 0.01$ which is consistent with the current Planck measurement and sensitivity. Note that

this is somewhat larger than the Planck Blue Book values which leads to weaker constraints. In particular, we see that $\sigma(\sum m_\nu) > 26$ meV for the full range of beam size and sensitivity. These results improve significantly with a stronger prior on τ .

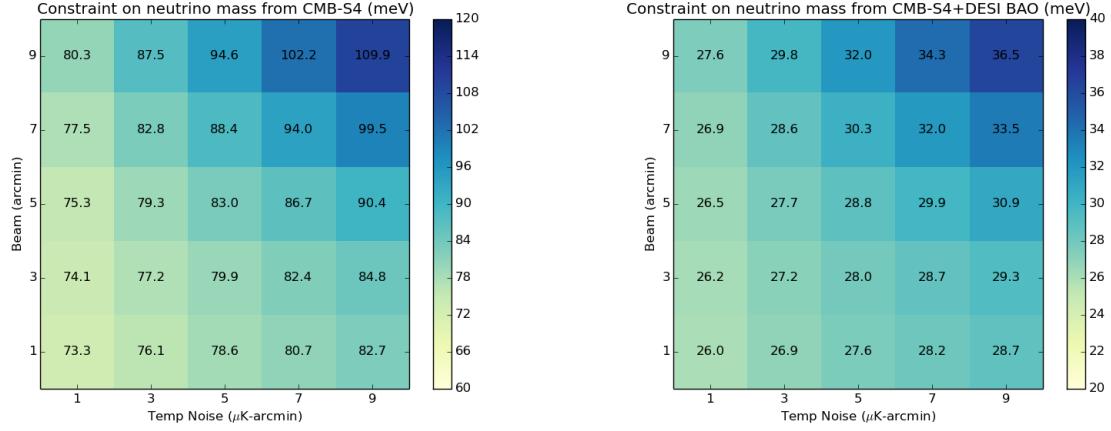


Figure 15. Forecasts for $\sigma(\sum m_\nu)$ assuming $\Lambda\text{CDM} + \sum m_\nu$. Both figures vary beam size in arcmin and effective detector noise in $\mu\text{K-arcmin}$. These limits assume an external prior on $\tau = 0.06 \pm 0.01$. Left: CMB Stage IV alone. Right: CMB Stage IV with DESI BAO.

3.2.6 Relation to Lab Experiments

Measurements of $\sum m_\nu$ by CMB-S4 will be interesting within the context of the broader neutrino experimental program. Neutrino flavor oscillations are described by a model where the neutrino flavor eigenstates are a mixture of massive neutrino eigenstates. The mixing is parameterized by the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix,

$$\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix} = \begin{pmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} \end{pmatrix} \cdot \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix},$$

where ν_i , $i = 1, 2, 3$ are the neutrino mass eigenstates. U_{PMNS} depends upon six real parameters: three mixing angles, θ_{12} , θ_{23} , θ_{13} that correspond to the three Euler rotations in a 3-dimensional space, and three phases, δ , α_1 , α_2 . A suitable parametrization is

$$U_{\text{PMNS}} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{13}s_{23}e^{i\delta} & c_{12}c_{23} - s_{12}s_{13}s_{23}e^{i\delta} & c_{13}s_{23} \\ s_{12}s_{23} - c_{12}s_{13}c_{23}e^{i\delta} & -c_{12}s_{23} - s_{12}s_{13}c_{23}e^{i\delta} & c_{13}c_{23} \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{i\alpha_1/2} & 0 \\ 0 & 0 & e^{i(\alpha_2/2)} \end{pmatrix}$$

where $c_{ij} \equiv \cos \theta_{ij}$ and $s_{ij} \equiv \sin \theta_{ij}$. The phases δ ($\equiv \delta_{CP}$) and α_1 , α_2 are Dirac-type and Majorana-type CP violating phases, respectively.

Experiments have measured the three mixing angles of U_{PMNS} and the two mass splittings, Δm_{21}^2 (the “solar” mass splitting) and Δm_{32}^2 (the “atmospheric” mass splitting), but fundamental aspects of neutrino mass and mixing remain. These include:

- measuring the absolute mass scale,
- determining the mass ordering,
- Lepton number violation (i.e., whether neutrinos are Majorana particles) and
- CP violation (measuring δ_{CP}).

Exploring this physics is the focus of current and upcoming neutrino experiments. CMB-S4 will measure $\sum m_\nu$ with sufficient sensitivity to be relevant to these open issues. Fig. 19 shows the relationship between CMB-S4 measurements and other measures of neutrino mass and ordering and Fig. 17 shows the connection between cosmological measurements of neutrino mass and neutrinoless double beta decay experiments.

3.2.6.1 Determining the neutrino mass scale

As discussed above, CMB-S4 will make a cosmological measurement of $\sum_i m_{\nu i}$ and thus the neutrino mass scale. This approach complements terrestrial measurements of the neutrino mass using radioactive decay. These kinematic measurements of neutrino mass focus on one of two processes, beta-decay or electron-capture, where the decay spectra near the decay endpoint is particularly sensitive to the mass of the neutrino (see Fig. 16).

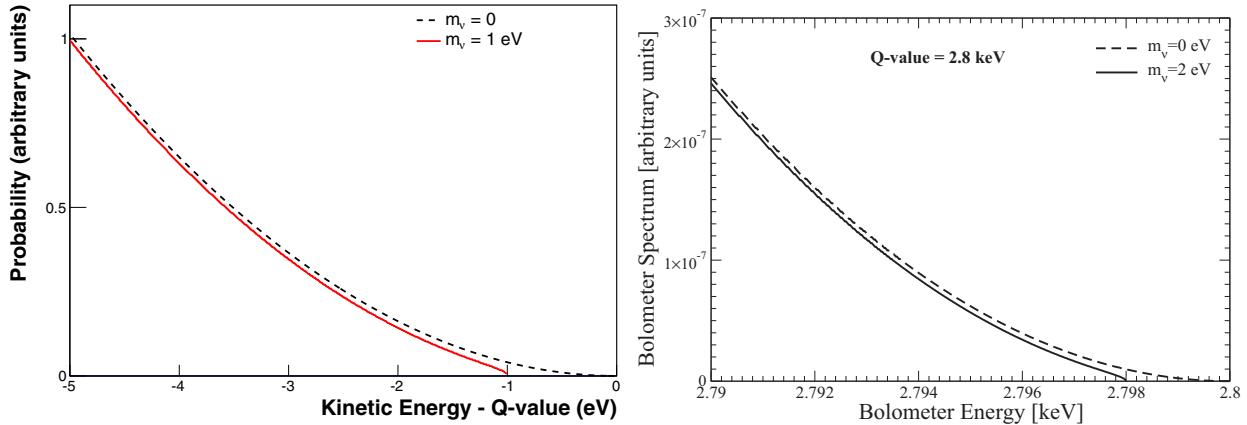


Figure 16. Illustration of the signal for kinematic measurements of the electron neutrino mass.

Current kinematic measurements from Mainz [263] and Troitsk [264] limit the neutrino mass to < 2.0 eV. The KATRIN experiment [265] will begin taking data in 2016 and is expected to improve this limit by a factor of ten.

Within the standard neutrino mass and cosmological paradigm, the kinematic and cosmological measurements of the neutrino mass are connected through the PMNS matrix. Thus, the combination of cosmological and terrestrial neutrino mass measurements tests our cosmological neutrino model. A discrepancy could point to new physics (e.g. modified thermal history through neutrino decay).

Improving kinematic measurements beyond KATRIN's 0.2 eV limit will require new technology since KATRIN will be limited by the final state spectrum of the source itself, specifically rotational-vibrational states of molecular Tritium. One of the new approaches is a calorimetric measurement of the electron-capture spectrum of ^{163}Ho . The calorimetric measurement of the ^{163}Ho endpoint is insensitive to the details of

the source configuration and may provide an avenue for eventually surpassing the KATRIN sensitivity. Interestingly, upcoming experiments such as ECHO [266], HOLMES [267], and NuMECS [268] utilize multiplexed superconducting detectors, the same technology baselined for the CMB-S4 experiment. Another promising direction for direct neutrino mass measurement is the frequency-based technique employed by the Project-8 experiment [269]. Project-8 aims to measure the beta-decay spectrum of Tritium by measuring the frequency of cyclotron radiation emitted by the decay electrons when trapped in a magnetic field. An exciting aspect to this frequency-based technique is the potential to trap atomic Tritium which is not subject to the rotational-vibrational excitations of molecular Tritium. A spectroscopic measurement using atomic Tritium could eventually achieve sensitivities of < 0.04 eV, a level comparable to cosmological measurements.

3.2.6.2 Lepton number violation: Majorana vs. Dirac neutrinos

One of the more interesting connections between cosmological measurements of neutrino mass and terrestrial experiments is the complementarity between cosmological neutrino mass measurements and the search for neutrinoless double beta decay (NLDBD). NLDBD is a hypothetical decay mode of certain nuclei where two neutrons convert to two protons and two electrons with no emission of neutrinos. The observation of NLDBD would be transformational demonstrating that neutrinos are Majorana particles and revealing a new lepton-number-violating mechanism for mass generation. This new physics could potentially explain both the smallness of neutrino masses and matter-antimatter asymmetry in the universe.

Initial results from the current generation of NLDBD searches limit the NLDBD half life, $T_{1/2}^{0\nu}$, to be larger than $\sim 2 \times 10^{25}$ years [270, 271, 272]. The full sensitivity of these experiments is expected to be in the range of $10^{25} - 10^{26}$ years [273]. Planning and technology development is already underway for next generation “ton-scale” NLDBD searches which would achieve sensitivities of $10^{27} - 10^{28}$ years [273].

We can illustrate the connection between NLDBD searches to cosmological determinations of neutrino mass by examining the simplest case where NLDBD is mediated by exchange of light Majorana neutrinos. Within the context of this mechanism, we can define an “effective neutrino mass,” $m_{\beta\beta}$, given by

$$m_{\beta\beta}^2 = \left(\sum_i U_{ei}^2 m_{\nu i} \right)^2 \quad (3.23)$$

where $m_{\nu i}$ are the light neutrino masses and U_{ei} is the usual PMNS mixing matrix including two unknown Majorana phases. The NLDBD half-life is then given by

$$(T_{1/2}^{0\nu})^{-1} = G^{0\nu} \cdot (M^{0\nu})^2 \cdot m_{\beta\beta}^2, \quad (3.24)$$

where $G^{0\nu}$ is a phase space integral and $M^{0\nu}$ is the nuclear matrix element. In this simple scenario, the signal from NLDBD experiments can be directly related to other measures of neutrino mass. Figure 17 illustrates this relationship between the effective neutrino mass and the lightest neutrino mass including limits and sensitivities of current and next generation NLDBD searches.

The complementarity between cosmological neutrino mass measurement and NLDBD can be understood by considering scenarios where NLDBD experiments either observe or fail to observe NLDBD. In the absence of a signal in next generation NLDBD searches, a cosmological measurement constraining $\sum_i m_{\nu i} > 100$ meV (corresponding to either the inverted hierarchy or a minimum neutrino mass of 50 meV) would strongly point to neutrinos being Dirac particles (see Fig. 17). On the other hand, if NLDBD is observed, equation 3.23 shows that cosmological measurements of $\sum_i m_{\nu i}$ are sensitive to the Majorana phases. For example, Fig. 18 shows that in the inverted mass hierarchy cosmological measurements together with NLDBD measurements

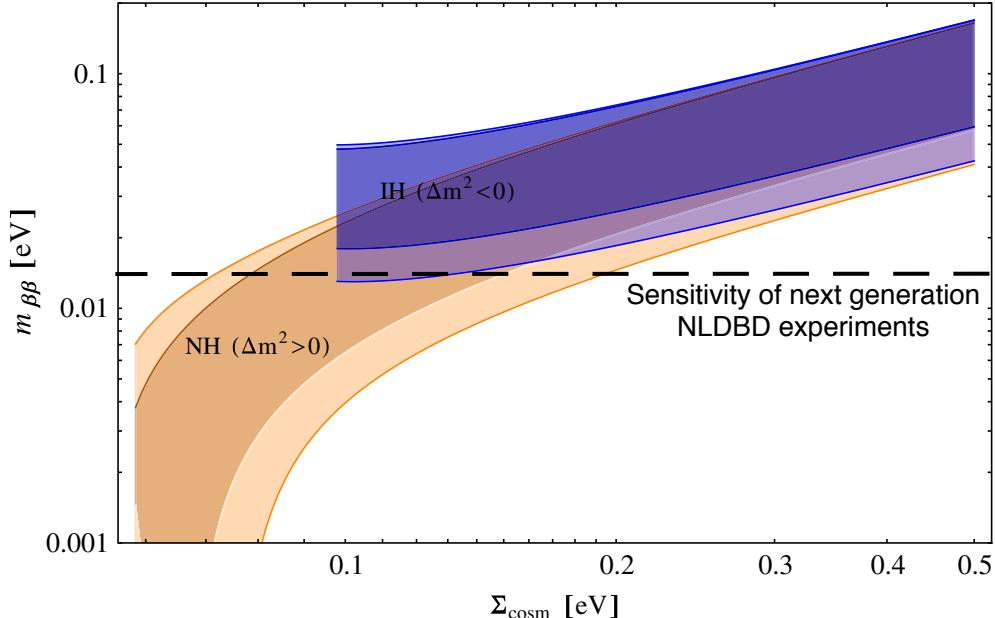


Figure 17. Plot of effective neutrino mass versus $\sum_i m_{\nu i}$ in the scenario where NLDBD is mediated by light neutrino exchange. The blue band corresponds to inverted ordering and the orange band corresponds to normal ordering. Next generation “ton-scale” NLDBD searches will have sensitivities down to $m_{\beta\beta} > 15$ meV (dashed line). Figure from [?]

can constrain one of the Majorana phases. Perhaps even more interesting would be the situation where cosmological and NLDBD measurements violate equation 3.23 indicating new physics beyond the simple model of light Majorana neutrino mediated decay.

3.2.6.3 Neutrino mass ordering and CP violation

In the case of normal ordering with non-degenerate neutrino mass, the CMB-S4 measurement of $\sum_i m_{\nu i}$ will provide a 2–4 σ determination of the neutrino mass ordering. Fully characterizing neutrino mass ordering and CP violation is one of the goals of the terrestrial neutrino physics program [274]. The upcoming reactor neutrino experiment JUNO [275] is scheduled to start data taking around ~ 2020 and will have ~ 2 –3 σ sensitivity to neutrino ordering after six-years of operation. Future experiments measuring atmospheric neutrino oscillations (e.g Hyper-K [276], DUNE [277], KM3NeT/ORCA [278]) can also resolve the neutrino mass ordering. For example, KM3NeT/ORCA forecasts a 3 σ measurement of the mass ordering by around 2023. Accelerator neutrino experiments are the only known method for exploring neutrino CP violation and in some cases, are also sensitive to neutrino mass ordering. In a ~ 5 -year timescale, the currently operating NO ν A experiment [279] may determine the neutrino ordering at the 2–3 σ level, provided that δ_{CP} falls into a favorable range. Hyper-K will measure δ_{CP} , though it requires external input regarding the neutrino ordering (e.g. from Hyper-K atmospheric neutrinos or from cosmology). The next-generation US-based long-baseline neutrino-oscillation experiment, DUNE, is planned to start operation around 2024, and will measure both the neutrino mass ordering (at the 2–4 σ level) and δ_{CP} . External input on neutrino ordering from other sources such as CMB-S4 would provide a strong consistency check of DUNE results and test the three-neutrino paradigm.

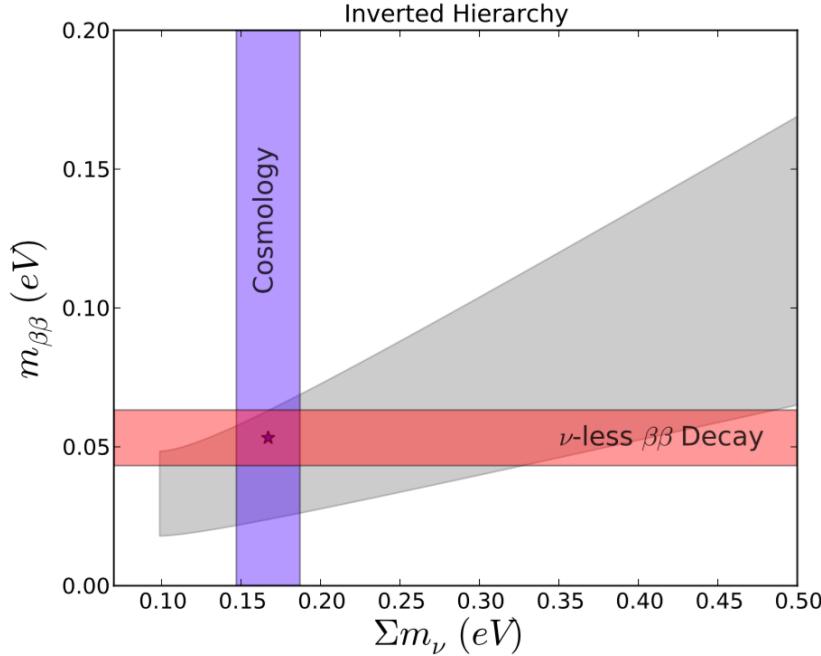


Figure 18. Relationship between effective neutrino mass as measured by NLDBD experiments versus $\sum_i m_{\nu i}$ as measured by cosmology for the inverted hierarchy. The gray band corresponds to a region allowed by existing measurements where the width of the band is determined by the unknown Majorana phase.

In the scenario where the neutrino mass spectrum is normally ordered and non-degenerate, CMB-S4 would be a strong complement to terrestrial experiments by providing a measurement of neutrino ordering that is independent of oscillation parameters and δ_{CP} . Under all circumstances, the combination of CMB-S4 with terrestrial determinations of neutrino ordering will provide a definitive measurement of the neutrino mass spectrum.

3.3 Relativistic Species at Recombination

The angular power spectrum of the cosmic microwave background (CMB) at small angular scales is quite sensitive to the radiation content of the early universe, usually parametrized by a quantity N_{eff} which will be defined more precisely below. In the Standard Models of cosmology and particle physics, N_{eff} is a measure of the energy density of the cosmic neutrino background. More generally, however, N_{eff} receives contributions from all forms of radiation apart from photons which are present in the early universe. Due to its sensitivity to N_{eff} , the CMB can be used as a tool to probe physics of the Standard Model and beyond which are difficult to measure through other means. Here we will give an overview of the theoretical issues related to N_{eff} and outline the expectations from the Standard Model.

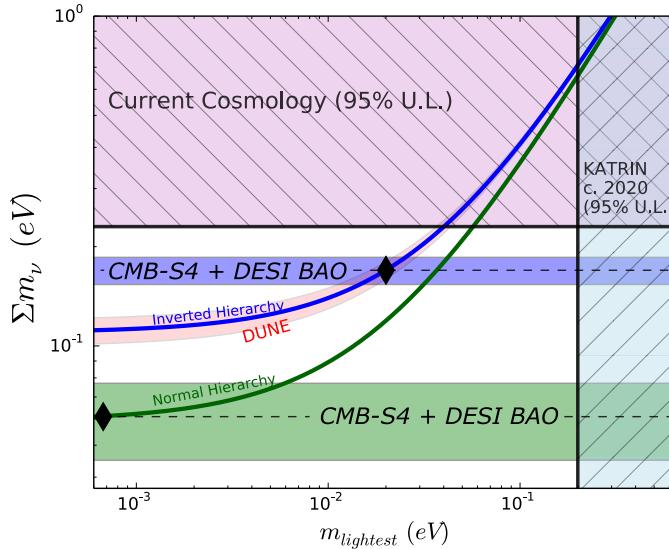


Figure 19. Shown are the current constraints and forecast sensitivity of cosmology to the neutrino mass in relation to the neutrino mass hierarchy. In the case of an “inverted ordering,” with an example case marked as a diamond in the upper curve, the CMB-S4 (with DESI BAO prior) cosmological constraints would have a very high-significance detection, with 1σ error shown as a blue band. In the case of a normal neutrino mass ordering with an example case marked as diamond on the lower curve, CMB-S4 would detect the lowest $\sum m_\nu$ at $\gtrsim 3\sigma$. Also shown is the sensitivity from the long baseline neutrino experiment (DUNE) as the pink shaded band, which should be sensitive to the neutrino hierarchy. Figure adapted from the Snowmass CF5 Neutrino planning document.

3.3.1 Thermal History of the Early Universe

In this section, we will give a sketch of the thermal history of the standard hot big bang universe when the temperature of the plasma was falling from about 10^{11} K to about 10^8 K following Section 3.1 of [280]. For other reviews see [281, 231]. During this era, there are two events of particular interest: neutrinos decoupled from the rest of the plasma, and a short time later electrons and positrons annihilated, heating the photons relative to the neutrinos. Our task is to follow how these events impact the evolution of the energy densities of the photons and neutrinos.

For massless particles described by the Fermi-Dirac or Bose-Einstein distributions, the energy density is given by

$$\rho(T) = \begin{cases} g \frac{\pi^2 k_B^4}{30 \hbar^3 c^3} T^4 & \text{Boson} \\ \frac{7}{8} g \frac{\pi^2 k_B^4}{30 \hbar^3 c^3} T^4 & \text{Fermion} \end{cases} \quad (3.25)$$

where g counts the number of distinct spin states. The entropy density for massless particles is given by

$$s(T) = \frac{4\rho(T)}{3T}. \quad (3.26)$$

It is convenient to define a quantity g_* which counts the spin states for all particles and antiparticles, with an additional factor $\frac{7}{8}$ for fermions. With this definition, the total energy density and entropy density of the

universe during radiation domination are given by

$$\begin{aligned}\rho(T) &= g_* \frac{\pi^2 k_B^4}{30 \hbar^3 c^3} T^4, \\ s(T) &= \frac{4}{3} g_* \frac{\pi^2 k_B^4}{30 \hbar^3 c^3} T^3.\end{aligned}\quad (3.27)$$

In an expanding universe, the first law of thermodynamics implies that for particles in equilibrium, the comoving entropy density is conserved

$$a^3 s(T) = \text{const.} \quad (3.28)$$

One straightforward consequence of this conservation is that for radiation in free expansion, the temperature evolves as the inverse of the scale factor

$$T \propto \frac{1}{a}. \quad (3.29)$$

Let us now apply this to the physics of the early universe.

At a temperature of 10^{11} K ($k_B T \sim 10$ MeV), the universe was filled with photons, electrons and positrons, and neutrinos and antineutrinos of three species, all in thermal equilibrium with negligible chemical potential, along with a much smaller density of baryons and dark matter both of which are unimportant for the present discussion. As the temperature of the plasma dropped below about 10^{10} K (about 1 second after the end of inflation), the rate of collisions between neutrinos and electrons and positrons could no longer keep up with the expansion rate of the universe, and neutrinos began to fall out of equilibrium and begin a free expansion. This is just above the temperature for which $m_e c^2 \sim k_B T$, and so for slightly lower temperatures electrons and positrons rapidly disappeared from equilibrium. We will simplify the discussion by assuming that neutrinos decoupled instantaneously before electron-positron annihilation and comment below how a more detailed calculation modifies the results. Non-zero neutrino masses can safely be neglected here as long as $m_\nu c^2 \lesssim 1$ keV which is guaranteed by current observational bounds.

From this point on, we will distinguish the temperature of neutrinos T_ν from that of the photons T_γ . Before neutrino decoupling, frequent interactions kept neutrinos and photons in equilibrium, ensuring they had a common falling temperature. After the universe became transparent to neutrinos, the neutrinos kept their relativistic Fermi-Dirac distribution with a temperature which fell as the inverse of the scale factor. The photons, on the other hand, were heated by the annihilation of the electrons and positrons. Comoving entropy conservation allows us to compute the relative temperatures at later times.

After neutrino decoupling, but before electron positron annihilation, the thermal plasma contained two spin states of photons, plus two spin states each of electrons and positrons, which means that during this period,

$$g_*^{\text{before}} = 2 + \frac{7}{8}(2+2) = \frac{11}{2}. \quad (3.30)$$

After electron positron annihilation, only the two spin states of photons remained, and so

$$g_*^{\text{after}} = 2. \quad (3.31)$$

Since $T_\nu \propto a^{-1}$ during this period, we can express the condition of comoving entropy conservation as follows

$$\frac{g_*^{\text{before}} T_{\gamma, \text{before}}^3}{T_{\nu, \text{before}}^3} = \frac{g_*^{\text{after}} T_{\gamma, \text{after}}^3}{T_{\nu, \text{after}}^3}. \quad (3.32)$$

Using the fact that $T_{\gamma, \text{before}} = T_{\nu, \text{before}}$, we find as a result

$$\frac{T_{\gamma, \text{after}}}{T_{\nu, \text{after}}} = \left(\frac{11}{4} \right)^{1/3}. \quad (3.33)$$

We find that in the instantaneous neutrino decoupling limit, the annihilation of electrons and positrons raised the temperature of photons relative to that of neutrinos by a factor of $(11/4)^{1/3} \simeq 1.401$.

After electron positron annihilation, assuming three species of light neutrinos and antineutrinos, each with one spin state, the radiation density of the universe is

$$\rho_r = \frac{\pi^2 k_B^4}{30 \hbar^3 c^3} \left[2T_\gamma^4 + 3 \frac{7}{8} T_\nu^4 \right] = \frac{\pi^2 k_B^4}{15 \hbar^3 c^3} \left[1 + 3 \frac{7}{8} \left(\frac{4}{11} \right)^{4/3} \right] T_\gamma^4. \quad (3.34)$$

It is conventional to define a quantity N_{eff} which gives the radiation energy density in terms of the effective number of neutrino species as

$$\rho_r = \frac{\pi^2 k_B^4}{15 \hbar^3 c^3} \left[1 + \frac{7}{8} \left(\frac{4}{11} \right)^{4/3} N_{\text{eff}} \right] T_\gamma^4. \quad (3.35)$$

In the instantaneous neutrino decoupling approximation described above, we found $N_{\text{eff}} = 3$. In the real universe, however, decoupling of neutrinos is not instantaneous, and the residual coupling of neutrinos at the time of electron positron annihilation increases N_{eff} by a small amount in the Standard Model.

The current best estimate of N_{eff} in the Standard Model is $N_{\text{eff}} = 3.046$ [282]. Unlike photon decoupling at temperature $T \sim 0.2$ eV, active neutrino decoupling at $T \sim 10$ MeV – 0.1 MeV takes place over many tens of Hubble times, with the result that we expect distortions in the relic neutrino energy spectra relative to thermal-shaped, Fermi-Dirac black bodies. Standard Model physics Boltzmann neutrino transport calculations show that these distortions could change N_{eff} from 3 to something close to 3.05. This result is largely due to (1) the incomplete decoupling of neutrinos during electron-positron annihilation and (2) QED plasma effects. While both effects have been calculated independently quite accurately, there is some theoretical uncertainty in this quantity due to the various numerical approximations that are made in the calculations when both effects are included simultaneously (see e.g. [283] for discussion).

3.3.2 Natural Target

As will be discussed in more detail below, the CMB angular power spectrum is sensitive to the value of N_{eff} in our universe. Measurement of the value of N_{eff} provides a huge amount of insight into the early universe, and is in fact an observational probe of the conditions at very early times, well before recombination. Even within the Standard Model, N_{eff} provides an observational handle on the thermal history back to about one second after the end of inflation. The true power of measuring N_{eff} , however, comes from the realization that it is sensitive not just to the neutrinos of the Standard Model, but it in fact receives contributions from all forms of radiation apart from photons present in the early universe and is thus a probe of new physics.

Collider experiments are known to provide a measurement of the number of neutrino species (or more precisely the number of species of fermions coupling to the Z boson with mass below $m_Z/2$) and find very close agreement with three families of light active neutrinos [284]. Cosmological measurements of N_{eff} provide complimentary constraints and are sensitive to the total energy density of radiation whether constituted of active neutrinos or other light species.

If the measured value of N_{eff} exceeds the Standard Model prediction, it would be an indication that there is additional radiation content in the early universe or that the thermal history is somehow modified. Additional radiation which contributes to N_{eff} is often referred to as dark radiation. There is a large number of possible sources for dark radiation, including sterile neutrinos [285, 286, 287], gravitational waves [288, 289, 107],

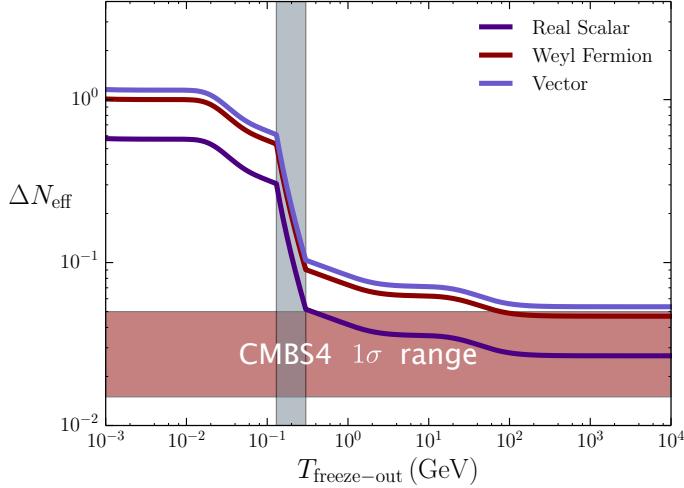


Figure 20. Contribution to N_{eff} from a massless field that was in thermal equilibrium with the Standard Model at temperatures $T > T_{\text{freeze}}$. For $T_{\text{freeze}} \gg m_{\text{top}}$, these curves saturate with $\Delta N_{\text{eff}} > 0.027$. The region in red shows the range of forecasts for $\sigma(N_{\text{eff}})$ for plausible CMB-S4 configurations.

dark photons [290, 291, 292], and many more [293, 294]. It is also possible that the measured value of N_{eff} could be found below the Standard Model prediction. This can happen if for example photons are heated relative to neutrinos after decoupling [295, 296].

One of the features that make N_{eff} a compelling theoretical target is the degree to which broad classes of models fall into two basic levels of ΔN_{eff} . As illustrated in Figure 20, any species that was in thermal equilibrium with the Standard Model degrees of freedom produces a characteristic correction to N_{eff} that depends only on its spin and its freeze-out temperature. For freeze-out after the QCD phase transition, one finds $\Delta N_{\text{eff}} \gtrsim 0.3$. Freeze-out before the QCD phase transition instead produces $\Delta N_{\text{eff}} > 0.027$. This first category has been tested by Planck. The second category, which is sensitive freeze-out temperatures as early as reheating, falls into the level of sensitivity attainable by CMB Stage IV.

The contributions to N_{eff} from hot thermal relics are relatively easy to understand from the discussion of neutrino decoupling. After freeze-out, the temperature of a relativistic species redshift like a^{-1} and therefore is only diluted relative to the Standard Model when energy is injected. The annihilation of heavy Standard Model particles into photons conserves the comoving entropy and therefore, the diluted temperature of a relic before neutrino decoupling is given by

$$\left(\frac{T_{\text{relic}}}{T_\nu}\right)^3 = \frac{g_*^{\nu-\text{freeze-out}}}{g_*^{\text{relic freeze-out}}} = \frac{43/4}{g_*^{\text{relic freeze-out}}}, \quad (3.36)$$

where g_* is defined as above to be the number of independent spin states including an additional factor of $\frac{7}{8}$ for fermions. The order of magnitude difference in ΔN_{eff} before and after the QCD phase transition comes from order of magnitude drop in g_* below the QCD scale. At temperatures well above to top mass, the Standard Model gives $g_* = 106.75$.

Even a measurement of N_{eff} which agrees with the Standard Model prediction to high precision would be very interesting due to the constraints it would place on physics beyond the Standard Model. Some specific implications for sterile neutrinos, axions, and other popular models will be discussed below. Broadly speaking, constraining ΔN_{eff} at the 10^{-2} level would constrain or rule out a wide variety of models that are consistent with current cosmological, astrophysical, and lab-based constraints. Furthermore, because of the sharp change in ΔN_{eff} at the QCD phase transition, the improvement from current constraints to projections for CMB Stage IV can be quite dramatic.

For the minimal scenario of a single real scalar, reaching $\sigma(N_{\text{eff}}) \sim 1 \times 10^{-2}$ would push the constraint on freeze-out temperatures from electroweak scale to the reheat temperature. This broad reach to extremely high energies and very early times demonstrates the discovery potential for a precision measurement of N_{eff} with the CMB. Furthermore, the CMB power spectrum has the ability to distinguish among certain types of dark radiation based on the behavior of its density perturbations [297, 298]. This point will be discussed further below.

A measurement with a slightly larger error on N_{eff} would still be extraordinarily valuable for higher spin fields, multiple light scalars, and modifications to the thermal history up to the electroweak scale. In particular, massless fermions and vectors have two helicity states which imply contributions of $\Delta N_{\text{eff}} \geq 0.047$ and $\Delta N_{\text{eff}} \geq 0.054$ respectively. As a result, a less sensitive instrument is still capable of probing physics back to reheating since it could detect or rule out the existence of light thermal relics with non-zero spin. In addition, there is good reason to think dark sectors could contain multiple light fields which could appear at any level of ΔN_{eff} above the minimum contribution from a single scalar field. Still, the most dramatic jumps in discovery potential occur at the critical values $\Delta N_{\text{eff}} = 0.027, 0.047$, and 0.054 .

3.3.3 Observational Signatures

Cosmic neutrinos and other light relics play two important roles in the CMB that are measured by N_{eff} . They contribute to the total energy in radiation which controls the expansion history and, indirectly, the damping tail of the power spectrum. The fluctuations of neutrinos and any other free streaming radiation also produces a constant shift in the phase of the acoustic peaks. These two effects drive both current and future constraints on N_{eff} .

The effect of neutrinos on the damping tail drives the constraint on N_{eff} in the CMB in $\Lambda\text{CDM} + N_{\text{eff}}$. The largest effect is from the mean free path of photons, which introduces a suppression $e^{-(k/k_d)^2}$ of short wavelength modes, with [299]

$$k_d^{-2} = \int \frac{da}{a^3 \sigma_T n_e H} \frac{R^2 + \frac{16}{15}(1+R)}{6(1+R)^2}, \quad (3.37)$$

where R is the ratio of the energy in baryons to photons, n_e is the density of free electrons, and σ_T is the Thompson cross-section. The damping scale is sensitive to the energy density in all radiation through $H \propto \sqrt{\rho_{\text{radiation}}}$ during radiation domination (which is applicable at high ℓ), and is therefore sensitive to N_{eff} or any form of dark radiation. From this discussion, we can also see the origin of the degeneracy between N_{eff} and n_e which may be altered by the primordial helium fraction, Y_p .

In reality, the effect on the damping tail is subdominant to the change to the scale of matter radiation equality and the location of the first acoustic peak [300]. As a result, the effect of neutrinos on the damping tail is more accurately represented by holding the first acoustic peak fixed. This changes the sign of the effect on the damping tail, but the intuition for the origin of the effect (and degeneracy) remains applicable.

sensitivity / beamsize	1'	2'	3'
1 $\mu\text{K}\text{-arcmin}$	0.030	0.032	0.35
2 $\mu\text{K}\text{-arcmin}$	0.035	0.038	0.041
3 $\mu\text{K}\text{-arcmin}$	0.040	0.042	0.045

Table 3-2. Forecasts for $\sigma(N_{\text{eff}})$ with varying beamsize in arcmin ('') and detector sensitivity. These forecasts assume $f_{\text{sky}} = 0.4$.

In addition to the effect on the Hubble expansion, perturbations in neutrinos affect the photon-baryon fluid through their gravitational influence. The contributions from neutrinos are well described by a correction to the amplitude and the phase of the acoustic peaks in both temperature and polarization [301]. The phase shift is a particularly compelling signature as it is not degenerate with other cosmological parameters [301, 298]. This effect is the result of the free-streaming nature of neutrinos that allows propagation speeds of effectively the speed of light (while the neutrinos are relativistic). Any gravitationally coupled free-streaming light relics will also contribute to the amplitude and phase shift of the acoustic peaks.

E-mode polarization will play an increasingly important role for several reasons. First of all, the acoustic peaks are sharper in polarization which makes measurements of the peak locations more precise, and therefore aid the measurement of the phase shift. The second reason is that polarization breaks a number of degeneracies that would also affect the damping tail [298].

Status of current observations – Planck has provided a strong constraint on $N_{\text{eff}} = 3.15 \pm 0.23$ when combining both temperature and polarization data. The addition of polarization data has both improved the constraint on N_{eff} and reduced the impact of the degeneracy with Y_p . Recently, the phase shift from neutrinos has also been established directly in the Planck temperature data [302]. This provides the most direct evidence for presence of free-steaming radiation in the early universe, consistent with the cosmic neutrino background.

3.3.4 Forecasts

As we described in Section 3.3.3, the most concrete signatures of N_{eff} from CMB Stage IV are derived from the damping tail at high- ℓ in both T and E . In addition, there is the effect of the shift in the phase of the acoustic peaks that is present at $\ell > 500$. Within ΛCDM these parameters are robust to degeneracies, having already accounted for the degeneracy of the location of the first acoustic peak with H_0 . The damping tail is generate with other extensions, like Y_p , as we will explore in Section 3.5.4.

Forecasts for $\sigma(N_{\text{eff}})$ under various experimental configurations are given in Table 3-2. Since the damping tail extends to very high- ℓ , constraints on N_{eff} are sensitive to ℓ_{max} and the beamsize of the experiment. We assume $\ell_{\text{max}} = 5000$ except for C^{TT} where we assume $\ell_{\text{max}} = 3000$ due to foregrounds. One can see the impact of the highest- ℓ modes by looking at the effect of the beamsize.

The other important parameter for N_{eff} is the sky fraction of the survey, f_{sky} . Since the signature of neutrinos depends on the detailed measurement of the shape of the power spectra, constraints depend on the number of observed modes. At fixed noise levels this implies that $\sigma(N_{\text{eff}}) \propto f_{\text{sky}}^{-1/2}$. Of course, with a fixed number of detectors the sensitivity, S , is proportional to $S \propto f_{\text{sky}}^{1/2}$. Since $\sigma(N_{\text{eff}})$ grows more slowly than S , we see that we mostly gain information by increasing f_{sky} . This trend is shown in Figure 21.

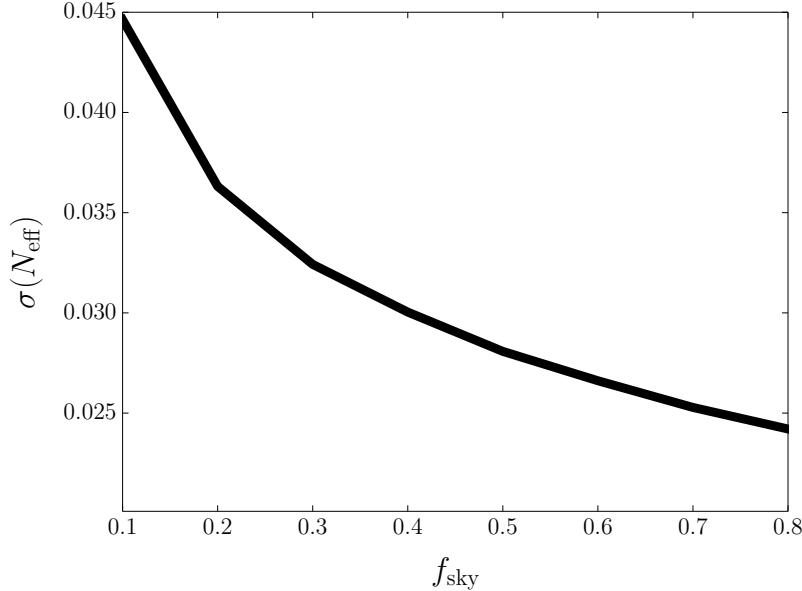


Figure 21. Forecasts for $\sigma(N_{\text{eff}})$ as a function of sky fraction. The sensitivity have been normalized to 1 $\mu\text{K}\text{-arcmin}$ for $f_{\text{sky}} = 0.4$ and is scaled according to $S \propto f_{\text{sky}}^{1/2}$ for different sky fractions. For a fixed number of detectors, we see that $\sigma(N_{\text{eff}})$ is minimized by increasing sky fraction.

We see from these results that for sufficiently large sky fraction, the thresholds for the light fermions and vectors are accessible at 2σ for plausible experimental configurations. These forecasts reach that minimum threshold for a light scalar of $\Delta N_{\text{eff}} = 0.027$ at 1σ for plausible configurations. Reaching at least 2σ sensitivity may be possible with more extreme design configurations (i.e. lower noise at large sky fraction).

These results are tantalizingly close to some exciting theoretical targets. This raises the question of whether we can reach these targets at higher significant by other means. First, constraints on N_{eff} improve significantly if $\ell_{\text{max}}^{TT} > 3000$. These modes are currently ignored due to foregrounds, but foreground removable could makes some of the modes accessible. Second, one could also imagine the sensitivity could be improved in combination with large scale structure surveys such as DESI and LSST, both by breaking degeneracies and providing a future measurement of the primordial power spectra.

3.4 Neutrino constraints on BSM physics

We primarily detect neutrinos in cosmology through their gravitational influence. Light relics can contribute to the same observables, $\sum m_\nu$ and N_{eff} , and therefore these measurements constrain additional light degrees of freedom. The precise relation between these observables and the underlying models depends on spin. This section is organized by the spin of the relevant light particle from axions (spin 0) to gravitons (spin 2).

3.4.1 Axions

Light particles of spin-0 (scalar fields) are highly constrained by naturalness. In the absence of a symmetry, one would expect them to be heavy and thus a poor candidate for dark radiation. On the other hand, (pseudo) Goldstone bosons are naturally light and they appear generically from spontaneous breaking of some high energy global symmetry. A ubiquitous example in Beyond the Standard model physics are axions and/or axion-like particles (ALPs). Axions have been introduced to solve the strong-CP problem [303], the hierarchy problem [304] and the naturalness of inflation [20]. Furthermore, they appear generically in string theory, in large numbers, leading to the qualitative phenomena described as the string axiverse [305]. They may even be tied to the origin of the breaking of the flavor, baryon/lepton number symmetries of the standard model.

At low energy, the mass of the ALP is protected by an approximate shift symmetry of the general form $a \rightarrow a + c$ where a is the axion and c is a constant (for non-abelian Goldstone bosons, this transformation will include higher order terms in a). We will define an ALP to be any such particle where all of the couplings of the axion to the Standard model respect such a symmetry. This symmetry may be softly broken with an explicit mass term, although this is highly restricted in the case of the QCD axion.

Two very common features of these models are that the axions are typically light (in many cases, $m \ll 1$ eV) and their interactions are suppressed by powers of f_a , the dimensional scale that controls the symmetry breaking. These two features make ALPs a particularly compelling target for N_{eff} . Because of the small masses, they will often behave as relativistic species in the CMB. Furthermore, because their production rate will scale as T^{2n+1}/f_a^{2n} for some $n \geq 1$, they are likely to be thermalized at high temperatures. Given that $\Delta N_{\text{eff}} \geq 0.027$ under such circumstances, a CMB experiment with sensitivity at this level will be sensitive to a very wide range of ALP models.

A CMB experiment sensitive to $\Delta N_{\text{eff}} \leq 0.027$ would have a dramatic impact on axion physics. We provide a detailed discussion of the constraints on individual couplings in Section 5.1.3, but here we will illustrate the impact qualitatively. At this sensitivity, any freeze-out at $T < T_{\text{reheat}}$ would be detectable in the CMB since they would all produce ΔN_{eff} . Since freeze-out is controlled by f_a we can translate the measurements of N_{eff} into limits on f_a . Since it is possible that $T_{\text{reheat}} > 10^{10}$ GeV, freeze-out is generally sensitive to f_a as high as 10^{12-14} GeV. This easily surpasses the sensitivity from stellar cooling ($f_a \lesssim 10^{10}$ GeV), which to date is one of the most broadly sensitive probes of axions. As long as $T_{\text{reheat}} > 10^4$ GeV, the CMB will be a more sensitive than stellar cooling to these models.

Axions can also produce $\Delta N_{\text{eff}} > 0.027$ in a wide variety of well-motivated circumstances. A case of particular interest is when axions couple directly to Standard model fermions. The Standard model has an $U(3)^5$ flavor symmetry and it is for axions couplings to arise if the axion is a goldstone boson that results from spontaneous breaking of this symmetry (or a sub-group). Given the non-abelian nature of the flavor symmetry, these scenarios can often lead to many axions (also known as familons). Under such circumstances, the contribution to freeze-out is given by

$$\Delta N_{\text{eff}} = N_a \times 0.027 \quad (3.38)$$

where N_a is the number of axions / number of broken generators of the symmetry group. It is easy to find scenarios where $N_a \sim \mathcal{O}(10)$ which is at the current level of sensitivity.

3.4.2 Sterile Neutrinos

Mechanisms of introducing neutrino mass generally include sterile neutrinos, with both Majorana and Dirac terms potentially contributing (e.g., Ref. [306]):

$$\mathcal{L}_D = -m_D (\bar{\nu}_L \nu_R + \bar{\nu}_R \nu_L) \quad (3.39)$$

$$\mathcal{L}_M = -\frac{1}{2}m_T (\bar{\nu}_L \nu_L^c + \bar{\nu}_L^c \nu_L) - \frac{1}{2}m_S (\bar{\nu}_R \nu_R^c + \bar{\nu}_R^c \nu_R) = -\frac{1}{2}m_T (\bar{\nu}_a \nu_a) - \frac{1}{2}m_S (\bar{\nu}_s \nu_s), \quad (3.40)$$

where $\nu_a \equiv \nu_L + (\nu_L)^c$ and $\nu_s \equiv \nu_R + (\nu_R)^c$ are active and sterile Majorana two component spinors, respectively. The mass m_T can be generated by a Higgs triplet, i.e., $m_T = y_T \langle \phi_T^0 \rangle$, or from a higher-dimensional operator involving two Higgs doublets with coefficients C/M . For dimension 5 operators, this becomes the Type-I seesaw mechanism, where both Majorana and Dirac terms are present and $m_S \gg m_D$.

A number of recent neutrino oscillation experiments have reported anomalies that are possible indications of four or more neutrino mass eigenstates. The first set of anomalies arose in short baseline oscillation experiments. First, the Liquid Scintillator Neutirno Detector (LSND) experiment observed electron antineutrinos in a pure muon antineutrino beam [307]. The MiniBooNE Experiment also observed an excess of electron neutrinos and antineutrinos in their muon neutrino beam [308]. Two-neutrino oscillation interpretations of these results indicate mass splittings of $\Delta m^2 \approx 1 \text{ eV}^2$ and mixing angles of $\sin^2 2\theta \approx 3 \times 10^{-3}$ [308]. Another anomaly arose from re-evaluations of reactor antineutrino fluxes that indicate an increased flux of antineutrinos as well as a lower neutron lifetime and commensurately increased the antineutrino events from nuclear reactors by 6%. This caused previous agreement of reactor antineutrino experiments to have a $\approx 6\%$ deficit [309, 310]. Another indication consistent with sterile neutrinos was observed in radio-chemical gallium experiments for solar neutrinos. In their calibrations, a 5-20% deficit of the measured count rate was found when intense sources of electron neutrinos from electron capture nuclei were placed in proximity to the detectors. Such a deficit could be produced by a $m_S > 1 \text{ eV}$ sterile neutrino with appreciable mixing with electron neutrinos [311, 312]. Some simultaneous fits to the short baseline anomalies and reactor neutrino deficits, commensurate with short baseline constraints, appear to prefer at least two extra sterile neutrino states [313, 314], but see Ref. [315]. Because such neutrinos have relatively large mixing angles, they would be thermalized in the early universe with a standard thermal history, and affect primordial nucleosynthesis [316] and CMB measurements of N_{eff} .

To accomodate $m_S = \mathcal{O}(\text{eV})$ with some mixing between active and sterile states in the neutrino mass generation mechanism discussed above requires mixing between active and sterile states with the same chirality, which does not occur for pure Majorana or Dirac mass cases or for the conventional seesaw mechanism. One proposed mechanism is the minimal mini-seesaw ($m_T = 0$ and $m_D \ll m_S \sim \mathcal{O}(\text{eV})$, e.g. Ref.[317, 318]). In such models, the sterile neutrinos can have the appropriate masses and mixings to accommodate the short baseline anomalies. For standard thermal histories, these sterile neutrinos are typically fully thermalized [316]. However, it is possible they are partially thermalized in two extra neutrino models [319].

Interestingly, there are combinations of CMB plus LSS datasets that are in tension, particularly with a smaller amplitude of fluctuations at small scale than that inferred in zero neutrino mass models. This would be alleviated with the presence of massive neutrinos, extra neutrinos, or both. In particular, cluster abundance analyses [320, 242] and weak lensing analyses [321] indicate a lower amplitude of fluctuations than zero neutrino mass [322]. Baryon Acoustic Oscillation measures of expansion history are affected by the presence of massive neutrinos, and nonzero neutrino mass may be indicated [257], though 2015 Planck results show a lack of such alleviation in cases with massive or extra neutrinos [230].

There is a potential emergence of both laboratory and cosmological indications of massive and, potentially, extra neutrinos. However, the combined requirements of the specific masses to produce the short baseline results, along with mixing angles that require thermalized sterile neutrino states, are inconsistent at this point with cosmological tension data sets [323, 324]. The tension data sets are not highly significant at this point ($\lesssim 3\sigma$), and there are a significant set of proposals for short baseline oscillation experiment follow up [325]. Future high-sensitivity probes of neutrino mass and number such as CMB-S4 will be able to definitively test for the presence of extra neutrino number and mass consistent with sterile neutrinos.

3.4.3 Light Fermions and Vectors

A general approach to interpreting N_{eff} constraints in terms of massless fields of arbitrary spin was undertaken in [326]. One identifies the symmetry that is required to explain the small mass of the particle, and then writes the most general interactions with the Standard Model consistent with the symmetry. The couplings will take the form

$$\mathcal{L} \supset \sum_{\Delta_h, \Delta_s} \frac{1}{\Lambda^{\Delta_h + \Delta_s - 4}} \mathcal{O}_{h, \Delta_h} \cdot \mathcal{O}_{s, \Delta_s} \quad (3.41)$$

where $\mathcal{O}_{h, \Delta_h}$ is an operator of dimension Δ_h constructed only from hidden sector fields (and similarly for $\mathcal{O}_{s, \Delta_s}$ and Standard Model fields). The total operator must be a scalar under both the Lorentz transformations and the symmetry that protects the mass of the hidden sector field(s). The bounds on axions discussed in Section 3.4.1 are one such example, where the axion is protected by a shift symmetry. For the purpose of this discussion, we have classified all scalars of this type as axions.

For a single Weyl fermion, χ , the leading couplings to the Standard Model are through the anapole moment and four-fermion interactions

$$\mathcal{L} \supset \frac{\chi^\dagger \bar{\sigma}^\mu \chi}{\Lambda_\chi^2} \left(d_a \partial^\nu F_{\mu\nu} + d_f \bar{\psi} \gamma_\mu \psi \right), \quad (3.42)$$

where we have chosen one of several four-fermion interactions for illustration and d_f, d_a are order one numbers. Current experimental constraint from LEP and the LHC limit $\Lambda_\chi \gtrsim 1$ TeV. Similar bounds are set by Planck by excluding the contributions to ΔN_{eff} from freeze-out after the QCD phase transition. If we are sensitive to the minimal contribution from a Weyl fermion of $\Delta N_{\text{eff}} = 0.047$, then for a reheat temperature of $T_{\text{reheat}} \sim 10^{10}$ GeV, we would be sensitive to $\Lambda_\psi \lesssim 10^{12}$ GeV. We see that for an order of magnitude improvement in sensitivity to N_{eff} , we get as much as a nine order of magnitude improvement in sensitivity to Λ_ψ .

For a hidden Dirac fermion with a $U(1)$ global symmetry, X , the leading² coupling is through an effective dipole interaction. A similar interaction also permits a hidden $U(1)$ gauge boson, A'_μ , to couple to Standard Model fermions.

$$\mathcal{L} \supset \frac{1}{\Lambda_X} \bar{X} \sigma_{\mu\nu} X F^{\mu\nu} + \frac{1}{\Lambda_{A'}^2} H \bar{\psi} \sigma_{\mu\nu} \psi F'_{\mu\nu}, \quad (3.43)$$

where $F'_{\mu\nu}$ is the field strength of A'_μ . The We have included the Higgs field, H , in the coupling to Standard Model fermions as it is required by gauge invariance above the scale of EWSB. Stellar cooling provides a strong constraint of $\Lambda_X \gtrsim 10^9$ GeV and $\Lambda_{A'} \gtrsim 10^5$ GeV.

Freeze-out above the scale of EWSB for a Dirac fermion produces $\Delta N_{\text{eff}} = 0.094$ and $\Delta N_{\text{eff}} = 0.054$ for a hidden photon. For a reheating temperature of 10^{10} GeV, we are sensitive to $\Lambda_X \lesssim 10^{13}$ GeV and $\Lambda_{A'} \lesssim 10^{11}$

²The dipole operator preserves the same $U(1)$ symmetry that allows a Dirac mass. It is unclear if one can UV-complete this model with a small Dirac mass.

GeV respectively. The Dirac fermion will be accessible with CMB Stage IV and will improve on the stellar cooling constraint for a reheat temperature $T_{\text{reheat}} > 10^4$ GeV.

3.4.4 Gravitinos

One of the most popular extensions of the Standard Model is supersymmetry, which is motivated both by naturalness and gauge coupling unification. Although the most generic possibilities are under significant tension for the LHC, there are still a variety of possibilities consistent with low-scale supersymmetry.

One of the universal predictions of supersymmetry is the existence of a spin-3/2 partner to the graviton, the gravitino. The gravitino mass is determined by the absolute scale of supersymmetry breaking,

$$m_{\frac{3}{2}} = \frac{|F|}{\sqrt{3}M_{\text{pl}}} \quad (3.44)$$

where $|F|$ is the order parameter for the scale of SUSY breaking (the vacuum expectation value of the auxiliary field F). This result does not depend on details of the mechanism of SUSY breaking unlike the super-partners of the rest of the Standard Model particles.

The typical coupling strength of the gravitino is the same as the graviton, $8\pi G = M_{\text{pl}}^{-2}$. However, the strength of the coupling to the helicity-1/2 component of the gravitino is enhanced by M_{pl}^2/F . This is simply the statement that the goldstino of SUSY breaking is coupled with strength F^{-1} (but is ‘eaten’ by the gravitino). Due to the enhanced coupling, the gravitino can be in thermal equilibrium with the Standard Model at plausible temperatures in the early universe. The gravitino therefore behaves just like a Weyl fermion in figure 20.

For $m_{3/2} \lesssim 10$ eV, hot relic gravitinos free stream on the scale of the CMB and therefore lead to observable signatures. Current data from Planck already requires that $m_{3/2} < 4.7$ eV from a combination of the primary CMB and CMB lensing [327]. To probe lower masses, note that for $m_{3/2} < \mathcal{O}(1)$ eV a gravitino will behave as free-streaming radiation from the point of view of the CMB. One finds that for these low values of the mass, gravitinos contribute a shift to N_{eff} ,

$$(\Delta N_{\text{eff}})_{3/2} \gtrsim 0.057 . \quad (3.45)$$

At $\sigma(N_{\text{eff}}) \sim 0.03$, CMB Stage IV can rule out all low-scale SUSY breaking models allowed by current cosmology.

3.4.5 Gravitational Waves

Since gravitational waves are massless and free-streaming, any gravitational waves which were present in the early universe naturally contribute to the total radiation energy density, and can therefore be constrained with N_{eff} [328, 288, 289, 107, 108].

Let us briefly review how to compute the energy density of a stochastic background of gravitational waves, following the treatment of [329, 330, 331, 332]. We will take the metric of spacetime to be given by a background component \bar{g}_{μ} described by the flat Friedmann-Robertson-Walker metric and a perturbation $\delta g_{\mu\nu}$. We will take the characteristic frequency of $\delta g_{\mu\nu}$ to be much higher than that of $\bar{g}_{\mu\nu}$. In particular, we will focus on gravitational waves whose wavelengths are much shorter than scales of cosmological interest.

The Ricci tensor can be expanded in powers of δg as

$$R_{\mu\nu} = \bar{R}_{\mu\nu} + R_{\mu\nu}^{(1)} + R_{\mu\nu}^{(2)} + \dots \quad (3.46)$$

We are interested in determining how spacetime is curved by the presence of small scale gravitational waves, or in other words, how $\bar{R}_{\mu\nu}$ is affected by terms containing $\delta g_{\mu\nu}$. Since $R_{\mu\nu}^{(1)}$ is linear in $\delta g_{\mu\nu}$, it contains only high frequency components, while on the other hand, $R_{\mu\nu}^{(2)}$ has both low and high frequency parts.

The high frequency part of Einstein's equations governs how gravitational waves propagate in a curved background, and is not necessary here. For the low frequency part, we can take an average over several cycles of the high frequency modes, which then gives

$$\bar{R}_{\mu\nu} = -\left\langle R_{\mu\nu}^{(2)} \right\rangle + 8\pi G \left\langle T_{\mu\nu} - \frac{1}{2}g_{\mu\nu}T \right\rangle. \quad (3.47)$$

We can then read off the effective energy-momentum tensor of small scale fluctuations

$$T_{\mu\nu}^{\text{GW}} = -\frac{1}{8\pi G} \left\langle R_{\mu\nu}^{(2)} - \frac{1}{2}\bar{g}_{\mu\nu}R^{(2)} \right\rangle + \mathcal{O}(\delta g^3). \quad (3.48)$$

If we take our perturbation to be of the form $\delta g_{ij} = a^2 h_{ij}$ with $h_{ij,j} = 0$ and $h_{ii} = 0$, we can compute the energy density of gravitational waves explicitly in the transverse traceless gauge

$$\rho_{\text{GW}} = T_{00}^{\text{GW}} = \frac{1}{32\pi G} \delta^{ik} \delta^{j\ell} \left\langle \dot{h}_{ij} \dot{h}_{k\ell} \right\rangle + \mathcal{O}(\delta g^3). \quad (3.49)$$

We will define the gravitational wave power spectrum as

$$\left\langle h_{ij}(\eta, \mathbf{x}) h^{ij}(\eta, \mathbf{x}) \right\rangle \equiv \int d\log k \mathcal{P}_t(k) [\mathcal{T}(\eta, k)]^2, \quad (3.50)$$

where $\mathcal{P}_t(k)$ is the primordial power spectrum of gravitational waves and $\mathcal{T}(\eta, k)$ is the tensor transfer function. The energy density of gravitational waves is then given by

$$\rho_{\text{GW}} = \frac{1}{32\pi G a^2} \int d\log k \mathcal{P}_t(k) [\mathcal{T}'(\eta, k)]^2, \quad (3.51)$$

where the prime denotes a derivative with respect to conformal time η .

Direct searches for the stochastic gravitational wave background are often quoted in terms of the normalized energy density per logarithmic scale

$$\Omega_{\text{GW}}(k) \equiv \frac{8\pi G}{3H_0^2} \frac{d\rho_{\text{GW}}}{d\log k} = \frac{\mathcal{P}_t(k)}{12H_0^2 a_0^2} [\mathcal{T}'(\eta_0, k)]^2. \quad (3.52)$$

On the other hand, constraints on N_{eff} provide an integral constraint on the spectrum of gravitational waves since waves of all frequencies contribute to the total energy density.

If for example, we take the primordial gravitational wave power spectrum to be a simple power law of the form $\mathcal{P}_t(k) = A_t \left(\frac{k}{k_*}\right)^{n_t}$ (as in Eq. (2.10)), we can use a constraint on N_{eff} to place bounds on the tensor spectral tilt n_t . The contribution to N_{eff} can then be approximated for $n_t > 0$ as [107]

$$\Delta N_{\text{eff}} \simeq \left(3.046 + \frac{8}{7} \left(\frac{11}{4} \right)^{4/3} \right) \frac{A_t}{24n_t} \left(\frac{k_{\text{UV}}}{k_*} \right)^{n_t}, \quad (3.53)$$

where k_{UV} represents the ultraviolet cutoff of the primordial tensor power spectrum. While this constraint does not probe the regime of great interest for inflationary models, it does provide a useful constraint on early universe alternatives to inflation which predict positive tensor tilt.

3.5 Complementarity of CMB and BBN

Primordial light element abundances have historically been an interesting observational test of hot big bang cosmology. The process by which light elements form in the early universe known as big bang nucleosynthesis (BBN) was worked out theoretically in the early days of the development of the hot big bang model of cosmology [333]. It is a process which depends on all four fundamental forces, that unfolded during the first three minutes of our current phase of expansion, and which has long provided a useful constraint on physics beyond the Standard Model. The current observational limits on primordial abundances overall show good agreement with the predictions of standard BBN. Primordial light element abundances are sensitive to the radiation content of the universe as measured through N_{eff} , which also affects the angular power spectrum of the CMB. Combining these two probes provides useful insight into the physics of the early universe which neither could achieve alone.

3.5.1 Standard Big Bang Nucleosynthesis

In this section, we will briefly review the physics of big bang nucleosynthesis in the Standard Model. For more extensive reviews see for example [280, 231, 334].

At temperatures above $k_B T \sim 1$ MeV, weak interactions kept neutrons and protons in thermal equilibrium, fixing their number densities to have the ratio $n_n/n_p = e^{-Q/k_B T}$, where $Q = 1.293$ MeV is the mass difference between neutrons and protons. At lower temperatures, interactions which convert protons to neutrons could not keep up with the expansion rate, leaving free neutron beta decay as the only channel by which protons and neutrons interconverted. The initial ratio of their number densities at the freeze-out temperature $k_B T_{\text{fr}} \simeq 0.8$ MeV was therefore

$$n_n/p_n = e^{-Q/k_B T_{\text{fr}}} \simeq 1/5. \quad (3.54)$$

After freeze-out, neutrons decayed until becoming bound into nuclei. This process proceeded primarily through two-body processes, starting with the formation of deuterium. The very small number density of baryons compared to that of photons (parametrized through $\eta \equiv n_b/n_\gamma$) delayed the start of these nuclear reactions until well after the temperature dropped below the binding energy of deuterium due to photo-dissociation of deuterium. The condition of the onset of deuterium formation is set by requiring that the number of photons per baryon with energy above the binding energy of deuterium drops below unity

$$\eta^{-1} e^{-|B_D|/k_B T_D} \simeq 1. \quad (3.55)$$

With the deuterium binding energy given by $|B_D| = 2.23$ MeV, and $\eta \sim 6 \times 10^{-10}$, we find that deuterium began to form when the temperature dropped below about $k_B T_D \simeq 0.1$ MeV. By this time, due to free neutron decay, the neutron to proton ratio had dropped to about $n_n/n_p \simeq 1/7$. Once deuterium was able to form, nearly all of the neutrons quickly became bound into the most energetically favorable light nucleus, which is ${}^4\text{He}$. We can estimate the mass fraction of primordial ${}^4\text{He}$, $Y_p \equiv \frac{\rho({}^4\text{He})}{\rho_b}$ to be

$$Y_p = \frac{2(n_n/n_p)}{1 + n_n/n_p} \simeq 0.25. \quad (3.56)$$

In addition to ${}^4\text{He}$, BBN produces a small amount of D, ${}^3\text{He}$, ${}^6\text{Li}$, and ${}^7\text{Li}$ (and also ${}^7\text{Be}$ which subsequently decays by electron capture to ${}^7\text{Li}$). While Y_p is primarily sensitive to the neutron lifetime, the primordial

abundances of the other light elements depend in complicated ways various nuclear rates and generally require numerical computation (see for example [335, 336, 337]).

Standard BBN is a one parameter model, depending only on the baryon to photon ratio η . The theory predicts several abundances which can be used to fix η and check the consistency of the theory, or alternatively, to constrain new physics. Current observations agree quite well with the predictions of standard BBN, with the exception of ^7Li . It is unclear whether this disagreement points to a problem with the astrophysical determination of the primordial abundance or a problem with the standard theory. The cosmological lithium problem remains unsolved [338]. From here on, however, we will ignore the lithium problem and focus on how measurements of the other abundances (primarily D and ^4He) can be used to constrain the physics of the early universe.

3.5.2 Beyond the Standard Model

Moving beyond standard BBN, measurements of primordial abundances have the ability to constrain many deviations from the standard thermal history and the Standard Model of particle physics. Because BBN is sensitive to all fundamental forces, changes to any force can in principle impact light element abundances. Of primary interest for our purpose is that BBN is sensitive to the expansion rate between about one second and a few minutes after the end of inflation. The expansion rate is in turn determined by the radiation content of the universe during this period, and thus BBN is sensitive to N_{eff} .

More specifically, the expansion rate determines the freeze-out temperature setting the initial ratio of neutrons to protons and the amount of time free neutrons have to decay. Additional radiation compared to the Standard Model gives a higher expansion rate, which leads to a higher freeze-out temperature and less time for free neutron decay, leading to a larger primordial ^4He abundance. The freeze-out temperature also depends weakly on the distribution function of electron neutrinos, though this is subdominant to the dependence N_{eff} for small non-thermal distortions [339].

Historically, Y_p had provided the best constraint on N_{eff} . Recent advancements in the determination of primordial deuterium abundance have made constraints on N_{eff} from deuterium competitive with those from Y_p [340]. The precision with which primordial abundances constrain N_{eff} is now comparable to that of constraints the CMB power spectrum, and there is no evidence for deviation from the Standard Model [230].

3.5.3 Complementarity with the CMB

The CMB can be used to quite precisely constrain η by measurement of the baryon fraction of the critical density, which is related to η by

$$\Omega_b h^2 \simeq \frac{\eta \times 10^{10}}{274}. \quad (3.57)$$

Using the value of η determined from CMB measurements as an input for BBN makes standard BBN a theory without free parameters which agrees very well with all observations (apart from the aforementioned disagreement with the observed lithium abundance). The CMB and BBN are sensitive to the baryon density measured at different times. While BBN is sensitive to the baryon to photon ratio up to a few minutes after the end of inflation, the CMB is sensitive to the baryon density at much later times, closer to recombination about 380,000 years later. Combining constraints from BBN and CMB on the baryon fraction therefore allows constraints on models where the photon or baryon density changes between these times.

$\sigma(N_{\text{eff}})/\sigma(Y_p)$	1'	2'	3'
1 $\mu\text{K-arcmin}$	0.082 / 0.0049	0.085 / 0.0050	0.090 / 0.0053
2 $\mu\text{K-arcmin}$	0.093 / 0.0054	0.096 / 0.0056	0.10 / 0.0058
3 $\mu\text{K-arcmin}$	0.10 / 0.0058	0.10 / 0.0059	0.11 / 0.0062

Table 3-3. Forecasts for $\sigma(N_{\text{eff}})$ and $\sigma(Y_p)$ with varying beam size in arcmin ('') and detector sensitivity. These forecasts allow both Y_p and N_{eff} to vary (along with the parameters of ΛCDM). We assume $f_{\text{sky}} = 0.4$.

The precision with which the CMB can constrain N_{eff} will soon come to surpass the constraints from BBN, but the value of the latter will not be totally eclipsed. BBN and the CMB probe the physics at different times, and so combining constraints can give insight into models where N_{eff} changes in time. If it were measured for example that $N_{\text{eff}}^{\text{BBN}} < N_{\text{eff}}^{\text{CMB}}$, this could be explained by the late decay of some unstable particles [341, 342, 343]. Alternatively, if observations revealed that $N_{\text{eff}}^{\text{BBN}} > N_{\text{eff}}^{\text{CMB}}$, this might signal late photon heating [293, 344].

The power spectrum of the CMB is also directly sensitive to Y_p . Since helium recombines earlier than hydrogen, the density of helium present at the time of recombination affects the free electron density, and thereby affects the damping tail of the CMB (though in a way which can be distinguished from the effects of N_{eff}) [301, 300, 302, 298]. The degeneracy between Y_p and N_{eff} is more strongly broken with precise CMB polarization data.

CMB Stage IV will provide constraints on N_{eff} which are about an order of magnitude better than the current best constraints, and will also improve on the measurement of Y_p by about a factor of two compared to the current best astrophysical measurements. Combined with measurements of other primordial abundances, this will help to provide a very thorough check of our understanding of the early universe and provide the opportunity to discover physics beyond the Standard Model.

3.5.4 Forecasts

Constraints on the combined Y_p - N_{eff} parameter space is a useful probe of both the physics of BBN and recombination and possible evolution in between. It is therefore useful to consider joint constraints on these two parameters available from the CMB alone.

The leading effect of both Y_p and N_{eff} is to change the damping tail and they are therefore degenerate. However, this degeneracy is broken both by the phase shift induced by neutrinos and by the E -mode spectra. The phase shift is sensitive to delensing of the T and E -modes and proper forecasting must account for realistic delensing, and the non-trivial covariance induced by the imperfect delensing. Forecasts including both effects are given in Table 3-3.

3.6 Detection Scenarios for Labs and Cosmology

The measurement of $\sum m_\nu$ and N_{eff} from cosmology are complimentary to a wide variety of experiments. In principle one or more experiment may observe deviations from the standard scenario and the interplay between the different observations will help narrow the space of possibilities.

3.6.1 Neutrino Physics

As discussed in Section 3.2.6, the measurements of the absolute scale of the neutrinos masses from the lab and from cosmology are complimentary in that they are sensitive to different parameters. In principle, there are a variety of possible scenarios where detections are made both in cosmology and in the lab. However, given current constraints, most scenarios that involve mostly conventional neutrino physics will result in upper limits from the lab based measurements and a detection of $\sum m_\nu$ and/or $\Delta N_{\text{eff}} \equiv N_{\text{eff}} - 3.046$. A plausible list of possibilities is shown in Table 3-4:

- Conventional neutrino mass scenarios imply majorana masses with a normal or inverted hierarchy. The normal hierarchy with $\sum m_\nu \simeq 58$ meV is perhaps the most conventional as it reflects the same hierarchical / non-degenerate masses that appear in the charged fermions of the Standard Model. This scenario is only detectable in the near term via cosmology due to the small size of the neutrino masses. Somewhat more exotic is the case of a Dirac mass, as it predicts the existence of new light states.
- The more exotic possibility is that there could be sterile neutrinos that are consistent with a variety of anomalies, as discussed in Section 3.4.2. In this case, we would observe a correlated signature in both a excess in $\sum m_\nu$ and ΔN_{eff} . The sterile neutrino parameters that are most consistent with the anomalies in short-baseline experiments are already in tension with cosmology but would be detected at high significance if true.
- Given the current cosmological constraints on $\sum m_\nu$, detections of m_β and $m_{\beta\beta}$ in near term experiments would require a significant change to the thermal history. In particular, a detection of a Majoron mass at the 0.25 eV level would predict a $\sum m_\nu$ that is already excluded by cosmology. Making the current (or future) limit consistent then requires a mechanism that satisfies both the present bound on $\sum m_\nu$ and the current constraints on N_{eff} .
- There are a variety of scenarios that produce $\Delta N_{\text{eff}} \gtrless 0$ without changing neutrino physics. In this case, the neutrino physics may follow a conventional pattern like the normal hierarchy. In principle, one would distinguish scenarios where there is no change to the number density of neutrinos (dark radiation) from scenarios where the neutrinos are diluted or enhanced by a change to the thermal history (late decay) as the interpretation of $\sum m_\nu$ depends on the neutrino number density. However, given that current measurements allow for < 10 percent change to the neutrino number density, we would need to detect $\sum m_\nu$ at 10σ to be sensitive to such a change. Nevertheless, dark radiation and changes to the thermal history can make correlated predictions for other experiments as we will discuss in the next subsection.

3.6.2 Dark Sectors and Particle Physics

Deviations from $N_{\text{eff}} = 3.046$ can arise from a wide variety of changes to the particle content and thermal history of the universe. In most cases, the physics responsible fundamentally requires a coupling of new particles to the Standard Model in regimes where they often can, in principle, be detected by other means. Cosmology is a very broad tool for searching for beyond the Standard Model physics, but it is also very complimentary to more targeted searches. A list of plausible detection scenarios is shown in Table 3-5:

- Evidence for new massless particles from either experiments or astrophysical observations have immediate implications for cosmology. Any non-cosmological probe of light particles necessarily requires a

Scenario	$m_{\beta\beta}$	m_β	$\sum m_\nu$	ΔN_{eff}	Conclusion
Normal hierarchy	$< 2\sigma$	$< 2\sigma$	60 meV	0	Normal neutrino physics; no evidence for BSM
Dirac Neutrinos	$< 2\sigma$	$< 2\sigma$	350 meV	0	Neutrino is a dirac particle
Sterile Neutrino	$< 2\sigma$	$< 2\sigma$	350 meV	> 0	Detection of sterile neutrino consistent with short-baseline
Diluted Neutrinos	0.25 eV	0.25 eV	< 150 meV	< 0	Modified thermal history (e.g. late decay)
Exotic Neutrinos	0.25 eV	0.25 eV	< 150 meV	0	e.g. Modified thermal history; (e.g. neutrino decay to new particle)
Excluded	0.25 eV	0.25 eV	500 meV	0	Already excluded by cosmology
Dark Radiation	$< 2\sigma$	$< 2\sigma$	60 meV	> 0	Evidence for new light particles; normal hierarchy for neutrinos
Late Decay	$< 2\sigma$	$< 2\sigma$	60 meV	< 0	Energy-injection into photons at temperature $T \lesssim 1$ MeV

Table 3-4. Relation between neutrino experiments and cosmology. We include the measurement of the Majorona mass via NLDBD ($m_{\beta\beta}$) or a kinematic endpoint (m_β) compared to the cosmological measurement of the sum of the masses $\sum m_\nu$ and the CMB measurement of N_{eff} . Here $< 2\sigma$ indicates an upper limit from future observations. For Section 3.2.6, one can use $\sigma(m_{\beta\beta}) \approx 0.075$ eV and $\sigma(m_\beta) \approx 0.1$ eV for observations on the timescale of CMB-Stage IV. For ΔN_{eff} the use of $\gtrless 0$ indicates a significant deviation from the Standard Model value.

coupling of the field to the Standard Model. A measurement suggesting the strength of the coupling for this particle implies an upper-limit on the freeze-out temperature. One can then look for the contribution to ΔN_{eff} associated with the particle. If no such contribution is detected, we could place an upper limit on the reheating temperature (or a lower limit for a detection). Axions present the simplest such examples, as there are a number of experiments that could directly detect axion dark matter (such as ADMX and CASPER). A detection in either experiment would predict $\Delta N_{\text{eff}} \geq 0.027$ unless reheating was at sufficiently low temperatures. The inferred bound in either experiment can be read off of Figures 35 or 36.

- A more complicated example is if a deviation for typical stellar cooling is observed, as has been suggested for white dwarfs. In this case, there are a number of models that could produce the necessary additional cooling, but would predict $\Delta N_{\text{eff}} \geq 0.027$. The precise value of ΔN_{eff} depends on the spin of the particle and nature of the coupling (which cannot be unambiguously inferred from cooling). Evidence for additional dark radiation would provide strong support that there is anomalous cooling and would imply a spin for the new light particle through the observed ΔN_{eff} using Figure 20.
- Evidence for supersymmetry at the LHC would suggest the existence of a gravitino. The collider phenomenology of most SUSY models is driven by the precise spectrum of masses of the super partners of the Standard Model fields. A priori, one may not be able to determine the mass or coupling of the gravitino in a collider. However, if the gravitino mass is low (\lesssim MeV) then a cosmological abundance of gravitinos is expected.

Scenario	ΔN_{eff}	Experimental Input	Conclusion
Axions	≥ 0.027	Direct detection of axions	lower-limit on the reheat temperature
Low-reheating	0	Direct detection of axions	upper-limit on the reheat temperature
Stellar Cooling	≥ 0.027	Anomalous Stellar cooling (e.g. white dwarfs)	Evidence for new light particle ; Spin determined by ΔN_{eff} .
Gravitinos	≥ 0.057	LHC evidence for SUSY	upper limit on scale of SUSY breaking
Late Decays	< 0	Spectral Distortions observed (μ or y)	evidence for new massive particle; energy injection
Evolving N_{eff}	> 0	Primordial abundances (BBN) consistent with $\Delta N_{\text{eff}} = 0$	Radiation density changed between BBN and Recombination

Table 3-5. Relation between particle physics experiments and cosmology.

- There are a variety of possibilities where we may observe $\Delta N_{\text{eff}} \gtrless 0$ which results from the decay of a massive particle after neutrino decoupling. A change to N_{eff} after BBN would imply that N_{eff} as measured in the CMB could differ significantly from the value inferred from primordial abundances [341]. From the CMB, we can measure Y_p and N_{eff} simultaneously which implies such a signature can even be internal to the CMB. Similarly decays to photons after BBN can also produce μ - or y -distortions to the CMB spectrum which could be correlated with $\Delta N_{\text{eff}} < 0$.

3.6.3 Dark Sectors and Neutrino Mass

CMB Stage IV will provide compelling sensitivity in the $N_{\text{eff}} - \sum m_\nu$ plane. From cosmology alone a measurement of $\sum m_\nu \gtrsim 60$ meV is consistent with conventional neutrino physics and therefore does not point to more exotic beyond the standard model physics without N_{eff} . However, an upper limit or detection of $\sum m_\nu < 60$ meV would provide evidence of unconventional cosmology on its own and combined with N_{eff} may give future insight into possible modifications to the cosmological history and/or neutrino physics necessary to accommodate such an observations. An illustration of possible scenarios is shown in Figure 22 and can be compared directly to the forecasted region in Figure 23:

- Conventional neutrino physics predicts $\sum m_\nu \gtrsim 60$ meV. A measurement of $\Delta N_{\text{eff}} < 0.02$ (i.e. consistent with zero) is the expectation from ΛCDM . As described in the previous subsection, $\Delta N_{\text{eff}} > 0$ suggests dark radiation or additional energy in neutrinos and $\Delta N_{\text{eff}} < 0$ would imply energy injection into photons after neutrino decoupling. The amount of energy injection that is consistent with the current constraint from Planck, $N_{\text{eff}} = 3.04 \pm 0.18$, is insufficient to significantly alter $\sum m_\nu$, as it would correspond to at most a five percent change in $\Omega_\nu h^2$.
- If $\sum m_\nu < 60$ meV, the most natural explanation is that the number density of neutrinos was lowered due to a change to the standard cosmological evolution. However, in order to significantly lower

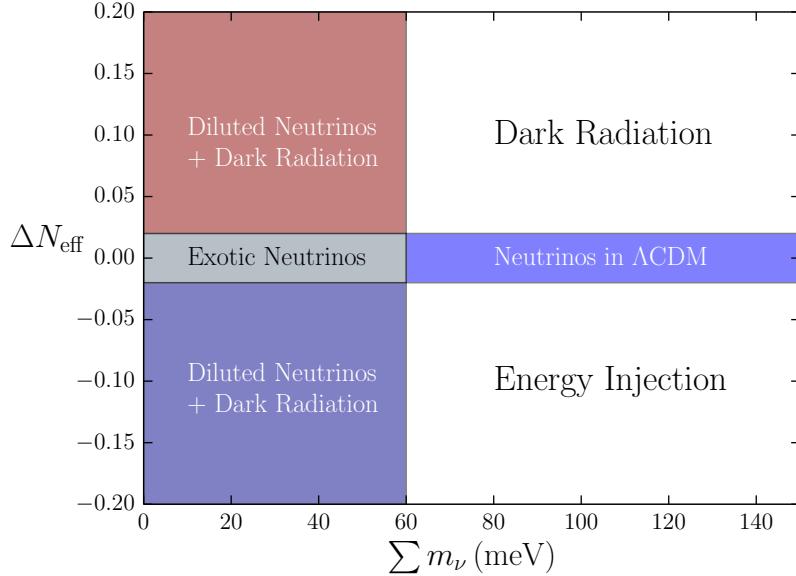


Figure 22. Physical mechanisms behind different regions in the space ΔN_{eff} - $\sum m_\nu$.

the number density to explain such an observation, the neutrinos number would have to be changed dramatically which, on its own, would be in contradiction with current N_{eff} constraints. Therefore, to satisfy existing constraints on N_{eff} from Planck, some other form of radiation (i.e. other than Standard neutrinos) would also be required to make $|\Delta N_{\text{eff}}| < 0.2$. Seeing a deviation of the form $\Delta N_{\text{eff}} \gtrless 0$ would give additional evidence for a modification of the thermal history.

- If $\sum m_\nu < 60$ meV and $\Delta N_{\text{eff}} = 0$, then it suggests that either the mass for the neutrinos is generated after the CMB (late mass) or that the heavy neutrinos decayed to a lighter specifics in some novel way. This situation would be unusual in that the limits on $\sum m_\nu$ would suggest deviations from the Standard thermal history without any other hints. Presumably this scenario would be scrutinized heavily to check that the amplitude of the power spectrum is normalized correctly. Finally, one might also allow for a delicate cancelation between the dilution of the neutrinos and the additional dark energy to be consistent with $\Delta N_{\text{eff}} = 0$.

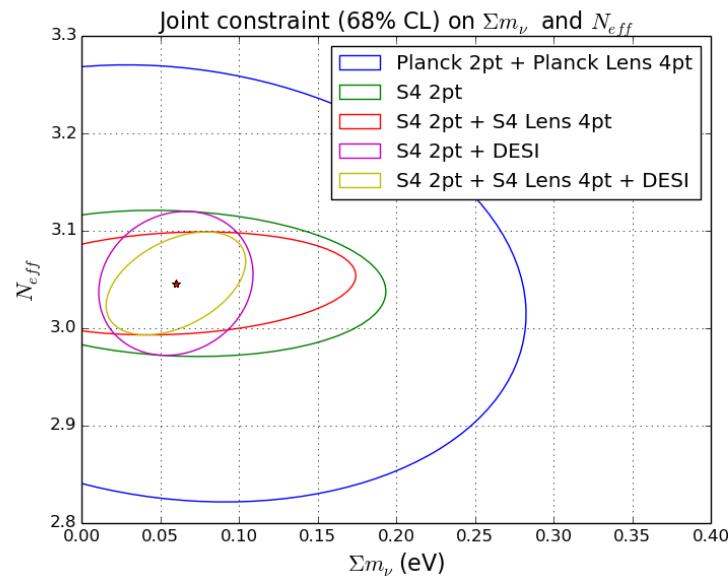


Figure 23. Forecasts 2d parameter space $\sigma(N_{eff})$ and $\sigma(\sum m_\nu)$. These constraints assume $f_{sky} = 0.4$ and 1 μK -arcmin noise. A prior of $\tau = 0.06 \pm 0.01$ was also assumed.

4

Dark Energy and Dark Matter

4.1 Dark Energy and Modified Gravity

(send feedback on this chapter to s4-de@cosmo.uchicago.edu)

The enigma of cosmic acceleration is among the most challenging problems in physics. Our most basic understanding about gravity – that objects fall towards one and other under mutual gravitational attraction – simply does not apply on the largest distance scales. Instead, gravity is apparently repulsive at large distances and late times; the scale of spacetime itself is currently not only expanding but accelerating. The implication is either that our understanding of gravity is incomplete, or some other causative agent – dark energy – with exotic gravitational properties fills the universe. In both cases, new physics is required beyond the four fundamental forces described by the Standard Model and general relativity.

The working hypothesis is that the cosmic acceleration is due to an exquisitely small cosmological constant, that Einstein’s general relativity is valid from millimeter to beyond gigaparsec scales, and that dark matter consists of a single species of a cold, collisionless particle. Yet none of these offer insight or reflect the unity of physics demonstrated elsewhere as in the Standard Model of particle physics.

In particular, the cosmological constant suffers from a naturalness problem whose resolution may lie in a dynamical dark energy, quintessence. Theories of quintessence posit a new scalar field and predict a variety of testable phenomena. They can also unveil new links to dark matter, neutrino physics, and cosmic parity violation. In their most general form, they represent a scalar-tensor theory of gravity which can be described by the effective field theory (EFT) of cosmic acceleration in the linear regime. CMB-S4 can provide the hard evidence needed to pare down these possibilities and potentially discover clues to the enigma of cosmic acceleration that will enable the development of compelling theoretical alternatives to the cosmological constant.

In summary, the current observational evidence suggests a new frontier for physics at low energies and weak coupling, implied by the cosmological scales that characterize cosmic acceleration. CMB lensing, thermal SZ cluster counts and mass measurements, and the kinematic SZ effect all measure the influence of cosmic acceleration on the growth of structure. As such, CMB-S4 would be capable of helping to answer basic questions about dark energy and gravity in a manner complementary to ongoing precision measurements of the expansion history.

4.1.1 Models and parameters

In this section, we briefly review the models and frameworks that have been proposed over the past years to test dark energy and modified gravity. These fall into three families: “trigger”, equations of motion, and theory parametrizations. The first ones are aimed at testing and falsifying the standard model of Λ CDM, a cosmological constant with cold dark matter, and are agnostic as to its alternatives. Given precise measurements from primary CMB anisotropy of the high redshift universe, all low redshift observables related

to the expansion history and growth of structure are potential triggers. Trigger parameters thus have the benefit that their relationship to the raw observables can be made as direct as desired. The drawback is that deviant values for the trigger may not have any physical motivation. Instead they help pare down the possibilities for the more model dependent and theory oriented tests.

In the next section we discuss the cluster abundance, CMB lensing and pairwise kinematic SZ effects as the building blocks of triggers when combined with other measurements such as BAO and SNIa. CMB-S4 will also enhance the precision and robustness of these other tests by measurements of the primary E -mode polarization. For example the cold dark matter $\Omega_c h^2$ and effective relativistic degrees of freedom N_{eff} enter into the calibration of the BAO scale as well as inferences on H_0 .

In addition to triggers based on the expansion history, CMB-S4 provides triggers based on the growth of structure. The Λ CDM model predicts that the growth of structure will slow in a precisely known manner as the expansion starts to accelerate. For example the rms amplitude of linear matter fluctuations at the $8h^{-1}\text{Mpc}$ scale $\sigma_8(z)$ is a trigger parameter that can be closely associated with the cluster abundance. The linear growth rate index γ is another that is closely related to peculiar velocities and the kSZ observables.

The second way of parametrizing deviations from Λ CDM is by modifying the equations of motion for dark energy in a manner consistent with conservation laws. These have the benefit of attempting to tie distance and growth tests together in a physical, yet still phenomenological manner. The next step up in complexity from a cosmological constant is a model where the dark energy is dynamical but spatially smooth relative to the dark matter. In these models, the expansion history can deviate from that of Λ CDM due to evolution in the dark energy equation of state $w(z)$, yet still predict the growth of structure. A common parameterization of this phenomenology is

$$w(z) = w_0 + w_a \frac{z}{1+z}. \quad (4.1)$$

The figure of merit defined by the DETF is the inverse area of the 95% CL region in the $w_0 - w_a$ plane.

There are generalizations of this type of parameterization that separate the expansion history from the growth of structure. A complete parametrization for observables for scalar-tensor theories in the linear regime would include in addition the gravitational slip or effective anisotropic stress (the ratio of the space curvature potential and Newtonian potential), the effective Newton constant, and c_T the speed of tensor perturbations. [intentionally vague on which notation to use: forecaster will decide the favorite one and supply details. Possibly add graviton mass as it is mentioned in Inflation chapter.]

The last way of studying deviations from Λ CDM consists in directly testing theories beyond it. Given the lack of a compelling specific theory to test, we can still make progress by parameterizing all possible Lagrangians for fluctuations that are consistent with the given symmetry. This approach maintains a strong connection with the underlying theory at the price of complicating the relation to the raw data.

More specifically, a systematic implementation of this approach is the effective field theory (EFT) of cosmic acceleration [345, 346], inspired by the EFT of inflation described in the Inflation Chapter [124, 167, 347, 348, 349, 350]. The EFT of cosmic acceleration describes the cosmological phenomenology of all universally coupled single scalar field dark energy and modified gravity models. Specifically, the EFT action is constructed in a unitary gauge to preserve isotropy and homogeneity of the cosmological background and

reads:

$$\begin{aligned} \mathcal{S}_{\text{EFT}} = \int d^4x \sqrt{-g} & \left\{ \frac{m_0^2}{2} [1 + \Omega(\tau)] R + \Lambda(\tau) - c(\tau) a^2 \delta g^{00} + \frac{M_2^4(\tau)}{2} (a^2 \delta g^{00})^2 \right. \\ & - \frac{\bar{M}_1^3(\tau)}{2} a^2 \delta g^{00} \delta K_\mu^\mu - \frac{\bar{M}_2^2(\tau)}{2} (\delta K_\mu^\mu)^2 - \frac{\bar{M}_3^2(\tau)}{2} \delta K_\nu^\mu \delta K_\mu^\nu \\ & + m_2^2(\tau) (g^{\mu\nu} + n^\mu n^\nu) \partial_\mu (a^2 g^{00}) \partial_\nu (a^2 g^{00}) + \frac{\hat{M}^2(\tau)}{2} a^2 \delta g^{00} \delta \mathcal{R} + \dots \Big\} \\ & + S_m[g_{\mu\nu}, \chi_m] \end{aligned} \quad (4.2)$$

where R is the four-dimensional Ricci scalar, δg^{00} , δK_ν^μ , δK_μ^μ and $\delta \mathcal{R}$ are, respectively, the perturbations of the upper time-time component of the metric, the extrinsic curvature and its trace and the three dimensional spatial Ricci scalar of constant-time hypersurfaces. Finally, S_m denotes the action for all the matter fields conventionally considered in cosmology.

In the action (4.2), the extra scalar degree of freedom is hidden inside metric perturbations. To study the dynamics of linear perturbations, however, it is convenient to make it explicit by means of the Stückelberg technique i.e. performing an infinitesimal coordinate transformation such that $\tau \rightarrow \tau + \pi$, where the field π describes the extra propagating degree of freedom. This approach allows us to maintain a direct link to the underlying theory so that we can keep under control its theoretical viability while exploring the cosmological implications of any of the models included in this language [351].

Since the choice of the unitary gauge breaks time diffeomorphism invariance, each operator allowed by the residual symmetry, in action (4.2), can be multiplied by a time-dependent coefficient that we shall call EFT function. To fully specify the phenomenology of linear perturbations only a restricted set of EFT functions are needed. These can be either parametrized to explore agnostically the space of dark energy and modified gravity models [352, 353, 354, 355] or can be fixed to reproduce exactly the phenomenology of some model of interest such as $f(R)$ gravity, quintessence and, more generally, the Horndeski class of theories and beyond [356, 357, 358].

4.1.2 CMB Dark Energy Observables

4.1.2.1 Cluster abundance and mass

Clusters of galaxies are sensitive to the content, geometry, and growth of structure in the universe. In the report of the DETF, they were highlighted as having the highest sensitivity to dark energy parameters but simultaneously the largest astrophysical systematic due to uncertainties in the mass scaling of the cluster observables. In the “Stage IV” era of dark energy probes, CMB-S4 will play a critical role in overcoming this challenge.

Clusters of galaxies are the most massive ($\sim 10^{14}\text{-}10^{15} M_\odot$) objects in the universe to have undergone gravitational collapse, having formed from regions $\sim 10\text{-}40$ Mpc. This property makes clusters representative of the overall content of the universe, and also makes them important tracers of the evolution of large-scale structure, sampling the most extreme peaks in the large-scale matter distribution. These properties have enabled clusters to make important contributions to cosmology: the discovery of dark matter in the Coma cluster [359], providing early evidence for $\Omega_m < 1$ [360, 361, 362], and constraining the physical nature of dark matter [363]. More recently, measurements of clusters have been used to constrain the properties of dark energy and modifications to gravity [364, 365, 366, 367, 368, 241]. In the future era of “Stage IV”

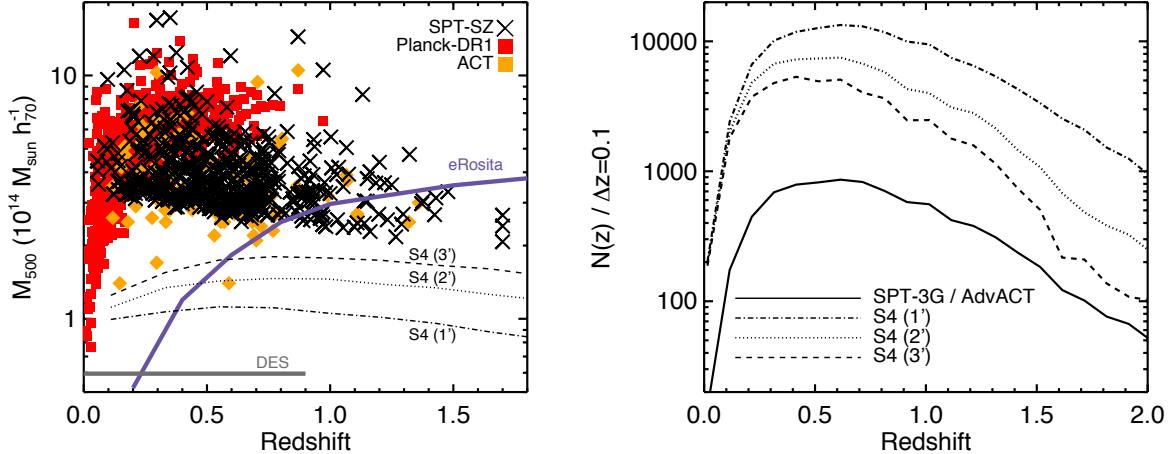


Figure 24. (Left) The 50% mass-completeness limits for three possible CMB-S4 instrumental configurations with either 1, 2, or 3 arc minute angular resolution. For comparison, this can be compared with existing SZ-selected cluster catalogs from Planck [379], SPT-SZ [378], and ACT [240], and future thresholds expected for the optical Dark Energy Survey and the X-ray eRosita survey [380]. (Right) The projected cluster counts for the three possible CMB-S4 configurations described above. For comparison, the projected cluster counts from the SPT-3G [381] and AdvACT surveys.

dark energy facilities such as DESI and LSST, measurements of the abundance of galaxy clusters can make complementary and competitive constraints on the dark energy equation of state and deviations from General Relativity [369].

CMB measurements find clusters through the inverse Compton scattering of CMB photons off of intra-cluster gas, otherwise known as the Sunyaev-Zel'dovich (SZ) effect [370]. SZ cluster surveys have two important advantages: the SZ surface brightness is redshift independent, and the integrated SZ signal is expected to be a relatively low-scatter cluster observable [371, 372, 373]. These properties enable SZ surveys to provide relatively clean, mass-limited cluster catalogs out to high-redshift ($z > 1$). Since the first SZ-discovered clusters were reported in 2009 [374], SZ surveys have produced catalogs of over 1000 SZ-selected clusters extending out to $z \sim 1.7$ [375, 376, 240, 377, 378, 379].

Figure 25 shows an estimate of the mass sensitivity, expressed as the one-sigma mass uncertainty as a function of redshift, for two possible CMB-S4 configurations and compares them to planned CMB experiments. In Figure 24, we show projections for the mass-threshold and total cluster counts for three possible CMB-S4 configurations. The 50% mass-completeness threshold for the CMB-S4 cluster survey would be relatively flat with redshift out to $z \sim 2$. The mass-threshold increases from $\sim 1-2 \times 10^{14} M_\odot$, going from a CMB-S4 angular resolution of 1 to 3 arc-minutes, which is > 2 times lower than current SZ surveys even in the worse case scenario for CMB-S4. In addition, the lower mass threshold and larger sky area of CMB-S4, would translate to a nearly ~ 100 fold increase in the number of SZ-identified clusters. At a 99% purity threshold, for a configuration with a 1, 2, and 3 arc-minute angular resolution, CMB-S4 would identify $\sim 40,000$, $70,000$, and $140,000$ clusters, respectively.

CMB-S4 will also enable a new means of calibrating cluster masses through CMB lensing. Accurate masses are crucial for catalogs to provide constraints on dark energy and modified gravity. With sufficient angular resolution, CMB-S4 opens tremendous possibilities for measuring cluster, more generally halo masses. Figure

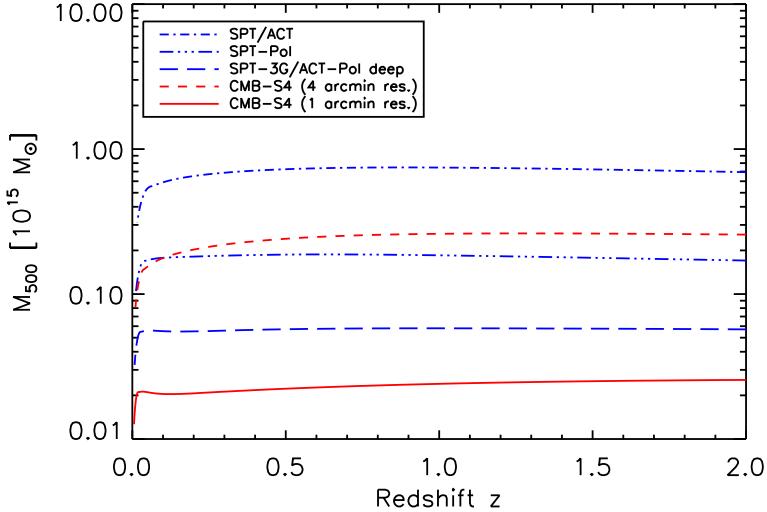


Figure 25. Cluster mass sensitivity of CMB lensing.

25 shows an estimate of the mass sensitivity, expressed as the one-sigma mass uncertainty as a function of redshift, for two possible CMB-S4 configurations and compares them to planned CMB experiments.

The estimation is made assuming foreground subtraction to reach the quoted CMB map noise level at the given angular resolution. The method (Melin & Bartlett 2015) employs an optimal filter matched to the NFW profile and applied to reconstructions of the lensing potential with a quadratic estimator (Hu & Okamoto 2002). The method has already been successfully applied to the Planck cluster cosmology sample of more than 400 objects (Planck Collaboration XXIV 2015). This figure shows the sensitivity obtained with just CMB temperature lensing reconstruction. Including polarization will significantly improve it. We see that the mass sensitivity remains flat with redshift, a remarkable property that enables mass estimation out to redshifts unreachable with galaxy shear measurements. This is a powerful and unique capability of CMB lensing. The figure also demonstrates the important gains attained with high angular resolution. At one arcmin resolution, S4 achieves a mass sensitivity of $2 \times 10^{13} M_\odot$ with temperature alone. This unprecedented sensitivity not only ensures robust mass estimation for cluster cosmology over the large redshift interval where S4 will detect clusters through the SZ effect, but also paves the way to numerous cluster and large-scale structure studies. The rapidly increasing science reach enabled by high angular resolution, approaching one arcmin, is an important consideration in CMB-S4 objectives.

4.1.2.2 Lensing

As described in the Lensing Chapter, the CMB lensing deflection map measures the projected mass density all the way back to the decoupling epoch at $z \sim 1100$, with the majority of the contributions coming from $z > 1$. CMB lensing is also dominated by structure on large scales in the linear regime. Thus, CMB lensing provides a clean probe of a particular integral over the linear growth of structure, e.g. $\sigma_8(z)$, weighted by distances. The lensing power spectrum shape is predicted from the background cosmology; shape deviations indicate scale-dependent effects on the growth, including those caused by modified gravity or the gravitational

effects of dark energy. CMB lensing complements other dark energy probes by providing a handle on effects at high redshift, e.g. in so-called early dark energy scenarios.

Cross-correlating the CMB lensing with other tracers of structure further permits extraction of information about the growth rate of structure in the universe that is localized in redshift. To the extent that other tracers have well-understood redshift distributions, cross-correlating a set of them to the CMB constitutes a tomographic study probing the evolution of the dark energy and its impact on the growth rate. The Lensing Chapter catalogs two broad categories of other tracers: galaxy density fields (and by extension the CIB) and galaxy shear maps. Combining lensing maps with maps of large scale flows from the kSZ will provide further constraints on the dark energy.

Another key contribution from CMB lensing to the study of dark energy will be its complementarity to other Stage IV experiments, including DESI and LSST, as well as EUCLID and WFIRST. (CMB Lensing chapter already mentions forecasting the improvement on calibrating multiplicative bias for LSST – would it make sense to move that here?) Not only will CMB lensing provide a new redshift kernel for tomography studies, it will also validate or improve the calibration for LSST, increasing its DE FOM by a factor of xxx.

4.1.2.3 Kinematic SZ

CMB-S4 will map with unprecedented precision the momentum field of the large scale structure via measurements of the kinematic Sunyaev Zel'dovich (kSZ) effect. Multi-frequency data can be used to remove other foregrounds and isolate the kSZ signal. CMB-S4 measurements with sufficient angular resolution can be used to reconstruct the diffuse kSZ anisotropy signal enabling sub-percent precision measurements of the amplitude of the matter density fluctuations σ_8 (see for example [382]) while measurements of the patchy kSZ can place strong constraints on the time and duration of reionization.

The combination of CMB-S4 with data from galaxy surveys will be able to measure the kSZ effect associated with galaxy clusters, which is proportional to their peculiar momentum. The large scale structure momentum field is an important cosmological observable that can place strong constraints on the cosmological parameters [383, 384, 385, 261] complementary to density fluctuations field measurements. The mean pairwise velocity of galaxy clusters is sensitive to both the growth of structure and the expansion history of the universe and it is an excellent probe for gravity on large scales. Being a differential measurement it is also particularly stable against residual foregrounds that might survive the frequency cleaning process. In [385, 261] it has been shown that a S4 survey with high resolution can constrain the redshift dependent growth of structure at $\lesssim 5\%$ precision in generic models allowing also for a redshift dependent equation of state of the dark energy. These measurements will be able to distinguish dark energy from modified gravity and will provide complementary constraints to redshift space distortions and weak lensing measurements, probing larger physical scales.

kSZ pairwise measurements can also constrain the sum of neutrino masses $M_\nu = \sum m_\nu$ with a 1σ uncertainty of 0.030eV for a 1 arcmin CMB-S4 overlapping 10000 deg^2 with a galaxy survey able to identify $M > 10^{13} M_\odot$ clusters. With a 5 arcmin resolution separating the CMB background from the kSZ signal would be more difficult, providing $\sigma_{M_\nu} = 0.076 \text{ eV}$ (de Bernardis et al., in preparation). These forecasts include only priors on the 6 standard cosmological parameters from Planck temperatures data and show the potential of the kSZ pairwise signal to provide constraints on the neutrino mass.

4.1.3 Forecasts

MR: the following section is highly work in progress. Any feedback is very welcome!

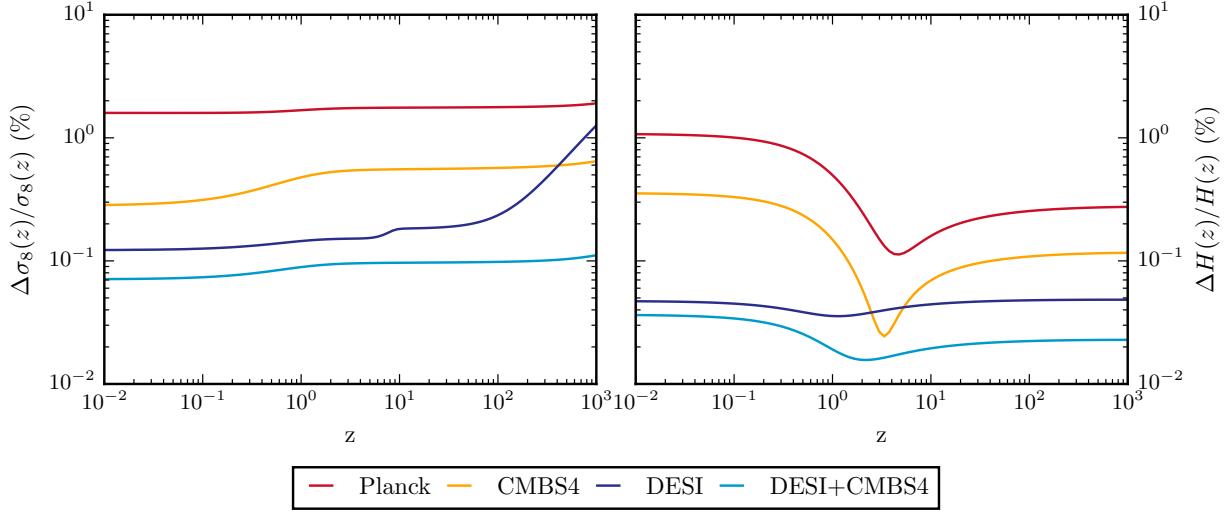


Figure 26. Relative 68% C.L. error on σ_8 and H as a function of redshift. Different colors correspond to different experiments, as shown in legend.

MR: shall we discuss here forecast specifications? Just in case, discussion of forecast specifications below.

To forecast CMB-S4 performances on Dark Energy and Modified Gravity models we shall use the following specifications. CMB-S4 is assumed to measure CMB fluctuations in temperature and polarization over 0.4 of the sky with a $1\,\mu\text{K}\,\text{arcmin}$ sensitivity in temperature and $1.4\,\mu\text{K}\,\text{arcmin}$ sensitivity in polarization, with a beam with 1 arcmin FWHM. This is added to *Planck* measurements of CMB fluctuations on the remaining part of the sky with specifications from [386]. To reproduce the noise levels of real *Planck* measurements at large angular scales in polarization the E and B mode polarization sky fraction is reduced to 0.01. Along with CMB probes we shall use DESI to exploit the complementary sensitivity of LSS measurements and investigate the synergies with CMB-S4 in constraining DE/MG models. We shall assume pessimistic specifications for the DESI survey as in [218]. When both CMB-S4 and DESI are considered we include in the forecast all the cross correlations between these two probes.

When a result is presented it is always marginalized over all the other parameters of the model. In particular, when considering DESI, we marginalize over a constant scale independent bias, different in all the survey redshift bins.

4.1.3.1 Trigger

We consider $\sigma_8(z)$, $H(z)$ and the growth index γ_L as trigger parameters. Figure 26 shows the relative error on $\sigma_8(z)$ and $H(z)$, assuming an underlying ΛCDM model. The central panel of Figure 27 shows the marginal forecast constraint on γ_L .

As we can see from the left panel of Figure 26 CMB-S4 will improve on *Planck* determination of $\sigma_8(z)$ substantially, pushing the sensitivity to this parameter to sub-percent accuracy, especially at late times. This level of accuracy is comparable with DESI measurements that are, in turn, just a factor two tighter. At early times CMB-S4 sensitivity to σ_8 is slightly lower but significantly better than DESI, as soon as redshift increases. Noticeably when CMB-S4 and DESI are joined σ_8 gets constrained to $\sim 0.1\%$ at all times. The

Datasets	$\sigma_{S_8(z=0)}$	Planck-KiDS tension	Planck-DES tension
<i>Planck</i>	0.025	-	-
KiDS-450	0.038	2.3σ	-
DES	0.06	-	0.6σ
CMB-S4	0.003	2.7σ	0.6σ
<i>Planck</i> and DESI	0.0009	4.1σ	1.5σ
CMB-S4 and DESI	0.0003	33σ	12σ

Table 4-1. Forecasted constraints on $S_8 \equiv \sigma_8 \sqrt{\Omega_m/0.3}$ and statistical significance of the discrepancy between *Planck* and the DES and KiDS surveys.

gain in the joint constraint is higher than the gain in sensitivity in going from CMB-S4 to DESI, thanks to the cross correlation between the two surveys. A similar picture emerges from the right panel of Figure 26 with the noticeable difference that CMB measurements have a peak in sensitivity around $z \sim 3$ that make CMB-S4 stronger than DESI. The joint DESI CMB-S4 constraints reflect this. The addition of CMB-S4 measurements improves the constraint on the expansion history significantly at redshifts higher than three.

We considered $\sigma_8(z)$, $H(z)$ as trigger parameters because these levels of sensitivity will be the key to resolve tensions between different experiments. CMB measurements and LSS surveys display a marginal disagreement on the determination of the growth of cosmic structures but these tensions are still in a low statistical significance phase. In particular *Planck* data are in tension with measurements of the Canada-France-Hawaii Telescope Lensing Survey (CFHTLenS) [387] and the Kilo Degree Survey (KiDS) [388] when considering the parameter $S_8 \equiv \sigma_8 \sqrt{\Omega_m/0.3}$. On the other hand the disagreement with the Dark Energy Survey (DES) is only marginal [389]. Two experiments with similar, high, sensitivity can either confirm or falsify these tensions to high statistical significance making CMB-S4 and DESI instrumental to each other.

In Table 4-1 we investigate the expected statistical significance of these two tensions, when assuming the *Planck* mean S_8 value and the KiDS and DES ones. As we can see if we replace the *Planck* error with the forecasted CMB-S4 one the statistical significance of these tensions is limited by the sensitivity of the weak lensing surveys. When considering *Planck* and DESI sensitivities the statistical significance improves becoming almost decisive but still being limited by *Planck*. Only when considering both CMB-S4 and DESI we will achieve definitive sensitivity and this will allow us to establish whether these discrepancies are due to new physical phenomena or just statistical fluctuations.

The power of CMB-S4 in constraining the growth of structures and its synergy with LSS surveys clearly shows when considering the Growth index γ_L . As we can see from both from the central panel of Figure 27 and Table 4-2 CMB-S4 will give stronger constraints with respect to *Planck* due to the additional leverage of CMB lensing. These constraints will be comparable with DESI ones and displaying a slightly different degeneracy with the amplitude of scalar perturbations. Leveraging on the precision of both CMB and LSS measurements, the joint constraints with CMB-S4 and DESI are significantly stronger than the single probes considered alone.

4.1.3.2 Equation of motion parametrization

MR: w0wa here?

Datasets	$\sigma(\gamma_L)$	$\sigma(\Omega_0)$	$\sigma(\gamma_0^{(2)})$	$\sigma(\gamma_0^{(3)})$	\tilde{M}_0	α_0^B	α_0^T
<i>Planck</i>	0.02	0.03	0.4	0.01	0.03	0.02	0.02
CMB-S4	0.007	0.02	0.1	0.01	0.02	0.02	0.008
DESI	0.007	0.2	0.4	0.1	0.02	0.07	0.03
CMB-S4 + DESI	0.002	0.01	0.05	0.002	0.006	0.02	0.001

Table 4-2. Forecasted 68% C.L. constraints on different models: the trigger parameter γ_L ; constant EFT couplings Ω_0 , $\gamma_0^{(2)}$ and $\gamma_0^{(3)}$; constant Horndeski couplings \tilde{M}_0 , α_B and α_T .

Datasets	Ω_{early}	Ω_{late}	\tilde{M}_{early}	\tilde{M}_{late}	α_{early}^B	α_{late}^B	α_{early}^T	α_{late}^T
<i>Planck</i>	0.08	0.05	0.2	0.1	0.05	0.2	0.03	0.04
CMB-S4	0.04	0.04	0.05	0.02	0.05	0.1	0.02	0.01
DESI	0.3	0.2	1.0	0.04	0.4	0.4	0.08	0.03
CMB-S4 + DESI	0.03	0.01	0.04	0.007	0.04	0.08	0.01	0.001

Table 4-3. Forecasted 68% C.L. constraints on early and late time values of different EFT couplings.

We parametrize deviations of the speed of GWs from the speed of light, with the parameter c_{GW}^2 . If the effect of primordial GWs on the B-mode polarization of the CMB is detected then the same observations will be capable of constraining their speed at the time of recombination. In the left panel of Figure 27 we show the marginalized joint forecast constraint on the tensor to scalar ratio and the speed of GWs. As we can see these two parameters are only weakly degenerate.

MR: uploading tomorrow the same forecast with different r fiducial

4.1.3.3 Theory parametrization

MR: beware of inconsistent notation in the figures $f_0 \rightarrow f_{\text{early}}$ and $f_1 \rightarrow f_{\text{late}}$. Will be fixed tomorrow.

We consider two parametrization basis for the functions describing the EFT of cosmic acceleration and, for the sake of simplicity, we focus on the Horndeski class of models. The first parametrization is obtained by making the couplings in action (4.2) dimensionless, as in [390]. The second one consists in re-parametrizing the couplings explicitly targeting the phenomenological features of Horndeski [391]. In both cases we consider two functional forms: we assume all the couplings constant in time; we allow all the EFT couplings to have different early and late time values, with a smooth transition in between, inspired by [392]. Specifically this second parametrization is given by $f(a) = 1/2(f_{\text{early}} + f_{\text{late}}) + (f_{\text{late}} - f_{\text{early}})\text{ArcTan}[(a - a_T)/\Delta a]/\pi$ where f_{early} and f_{late} are respectively the early and late time values of the considered EFT function, a_T is the transition scale factor assumed to correspond to $z = 10$ and $\Delta a = 0.01$ is the transition sharpness.

In Figure 29 and Table 4-2 we show the forecast constraints on constant EFT couplings. As we can see the sensitivity of CMB probes are unmatched when measuring the conformal coupling to gravity Ω_0 . CMB-S4 measurements, in addition, are found to be the most constraining measurements on the other two EFT higher order operators, $\gamma_0^{(2)}$ and $\gamma_0^{(3)}$. Confirming the picture presented in the previous sections the synergy between CMB-S4 measurements and DESI results in much tighter constraints on all the considered parameters. For all the probes considered the kinetic operator, $\gamma_0^{(1)}$, is found to be unconstrained.

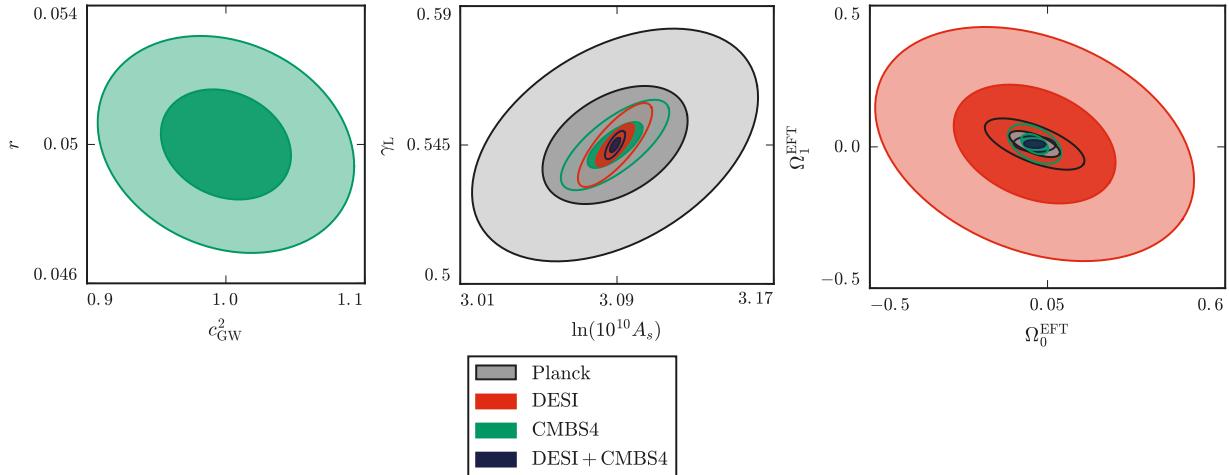


Figure 27. Forecasted marginalized constraints on different models. The left panel shows the joint constraints on the tensor to scalar ratio and the speed of gravitational waves. The central panel shows the joint constraints on the growth index and the amplitude of scalar perturbations. The right panel shows the joint constraints on relative variations of the gravitational constant at early times Ω_0^{EFT} and late times Ω_1^{EFT} . Different colors correspond to different experiments, as shown in legend.

A similar picture also emerges from the forecast constraints on Horndeski couplings, with CMB-S4 providing the tightest bounds, as we can see from Figure 30 and Table 4-2. When considering the effective Planck mass \tilde{M}_0 and the tensor speed excess, α_0^T , DESI and *Planck* sensitivities are comparable while CMB-S4 is a factor 1.5 and 2.5 stronger, respectively. CMB measurements, on the other hand, are the most powerful at constraining the braiding coefficient α_0^B and we can notice that *Planck* measurements are slightly stronger than CMB-S4 ones leveraging on the constraining power of large angular scales. As expected, combining CMB-S4 to DESI, results in a significant improvement with respect to the single probes alone. As in the constant EFT case the scalar field kineticity α_0^K is found to be unconstrained.

When considering all the EFT couplings having different values at early and late times we found that early times changes are constrained by physical viability requirements and late time values have comparable bounds with respect to the constant case considered before. The only EFT coupling that does not display this behavior is the conformal one, $\Omega(a)$, and the corresponding forecast constraints are shown in the right panel of Figure 27 and Table 4-3. For these parameters we find that marginalization slightly degrades the forecast bounds with respect to the constant case. Moreover we find that data are more sensitive to the sum of these two parameters rather than their difference but when CMB-S4 is combined to DESI this degeneracy in parameter space is relieved.

Horndeski couplings in turn display a qualitatively different picture, as we can see from Figure 28 and Table 4-3. Not surprisingly CMB measurements are generally stronger than LSS surveys at constraining the early time values of these functions, i.e. \tilde{M}_{early} , α_{early}^T and α_{early}^B . However CMB-S4, leveraging on both the early and late time constraining power of CMB and CMB lensing measurements is sensitive to both early and late time values, to unmatched accuracy.

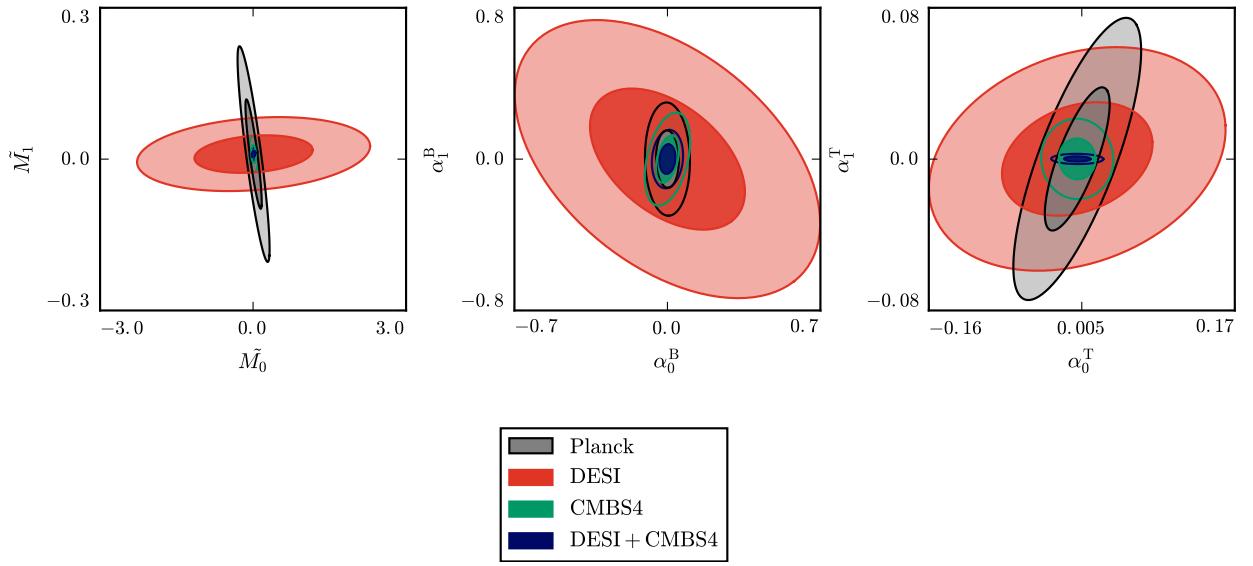


Figure 28. Forecasted marginalized constraints on Horndeski couplings. The left panel shows the joint constraints on the effective Planck mass at early times \tilde{M}_0 and late times \tilde{M}_1 . The central panel shows the joint constraints on the Horndeski braiding coefficient at early times α_0^B and late times α_1^B . The right panel shows the joint constraints on the Horndeski tensor speed excess coefficient at early times α_0^T and late times α_1^T . The Horndeski kineticity coefficient is unconstrained by all the experimental combinations considered. Different colors correspond to different experiments, as shown in legend.

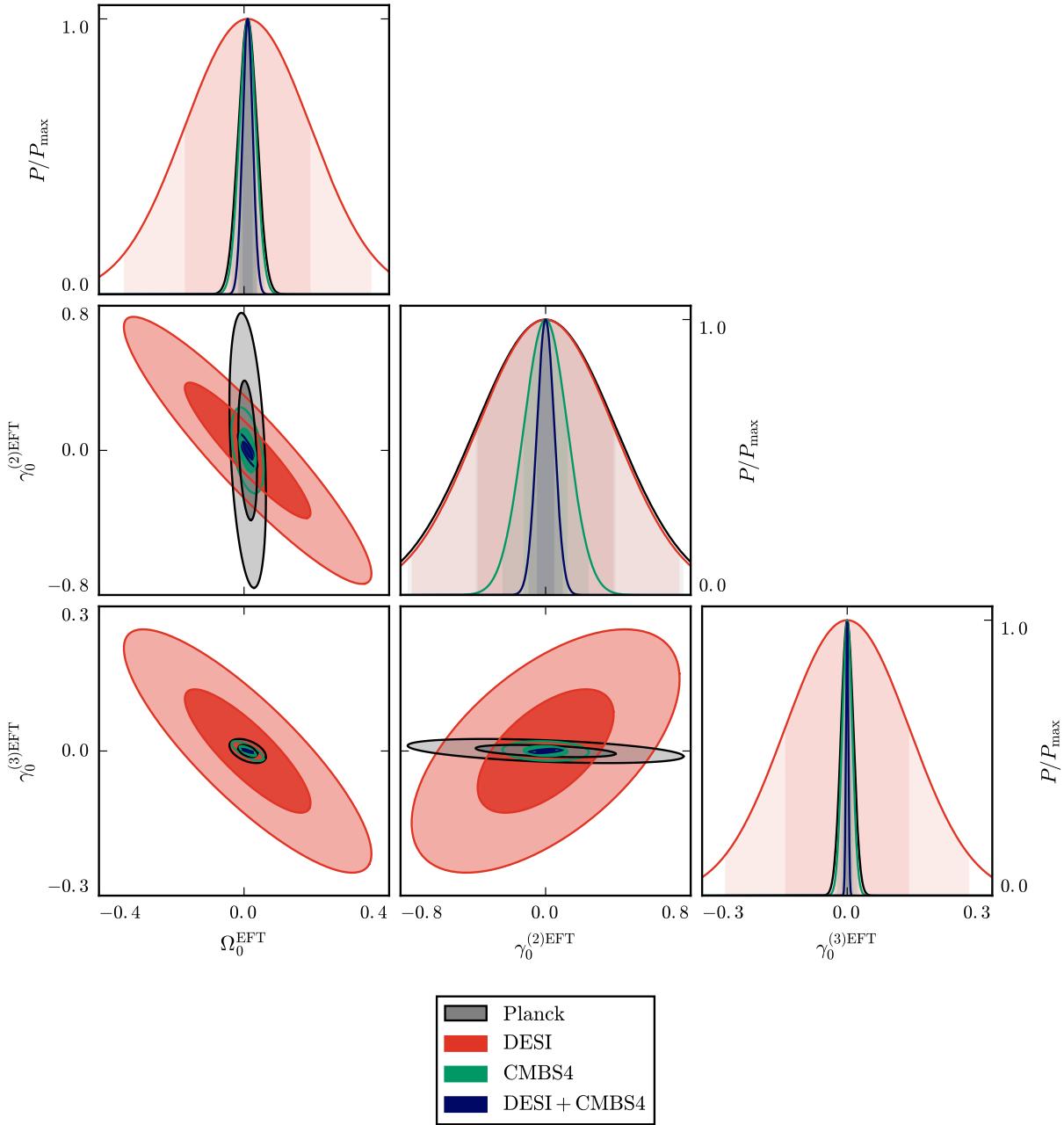


Figure 29. Forecasted marginalized constraints on constant EFT couplings. Different colors correspond to different experiments, as shown in legend.

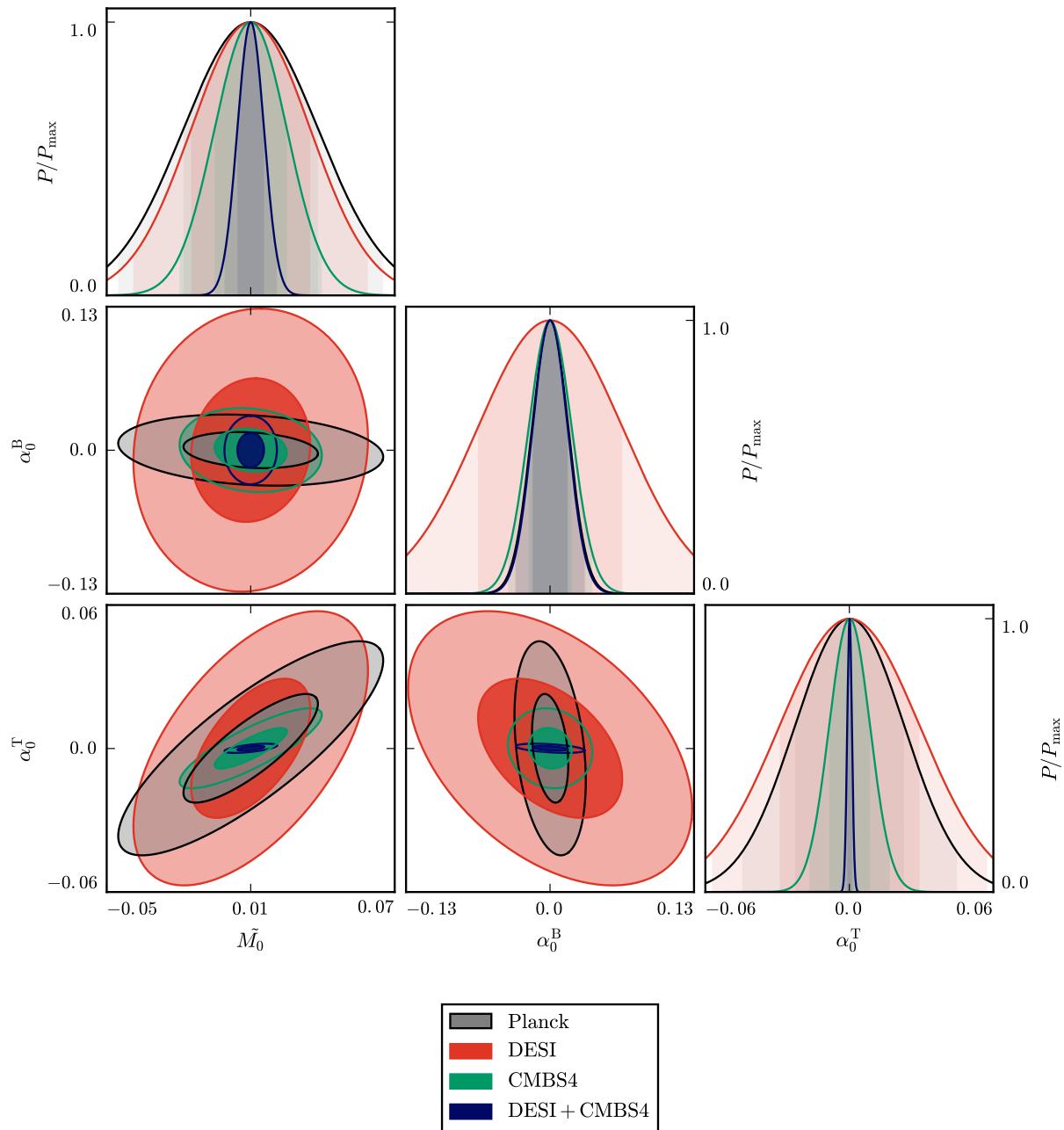


Figure 30. Forecasted marginalized constraints on constant Horndeski couplings. Different colors correspond to different experiments, as shown in legend.

4.2 Dark Matter

We have learned from many different experiments (weak and strong lensing, studies of the Bullet Cluster, the Cosmic Microwave Background, etc) that approximately 84% of all the matter in the universe is composed of dark matter, which is not accounted for by the Standard Model of particles. However, the particle nature of dark matter remains unknown.

There are various types of experiments trying to shed light on this question: direct detection experiments, indirect detection experiments, and collider experiments. An alternative observable where Dark Matter interactions can modify the Standard Model prediction is the CMB power spectrum. CMB-Stage IV will have the sensitivity to detect new cosmological signatures originating from various types of dark matter interaction. In the sections below, we will discuss some of these possible scenarios.

4.2.1 Dark Matter Annihilation

One of the leading candidates for dark matter are the Weakly Interactive Massive Particles (WIMPs). If dark matter consists of WIMPs, we would expect these particles to self-annihilate. The annihilation of dark matter produces a shower of very energetic particles, that injects energy into the universe, ionizing the matter in it.

This extra source of ionization has distinctive effects on the Cosmic Microwave Background (CMB): it suppresses the CMB temperature and polarization fluctuations at small angular scales, and it enhances the CMB polarization fluctuations at large angular scales due to the extra scattering of photons off free electrons [393, 394]. CMB temperature and polarization spectra can constrain the parameter $p_{\text{ann}} = f\langle\sigma v\rangle/m_{\text{DM}}$, where f is the fraction of energy deposited into the plasma, $\langle\sigma v\rangle$ is the velocity-weighted cross section, and m_{DM} is the mass of the DM particle. Current constraints coming from WMAP 9-year data, Planck, ACT, SPT, BAO, HST and SN data excluded Dark Matter masses below 26 GeV at the 2σ level, assuming that all the energy is deposited in the plasma [395]. CMB-Stage IV is expected to tighten these constraints by a factor of 10 [225]. Ref. [225] found that the main factor that improves the limit in m_{DM} is the sky coverage f_{sky} . This is because the constraints are mostly sample variance limited. Fig. 31 shows the dependence on f_{sky} , and the small dependence on detector number and beam size.

Dark-matter annihilation also leads to growing ionization fraction perturbations and amplified small-scale cosmological perturbations, leaving an imprint on the CMB bispectrum [396].

4.2.2 Non-standard Dark Matter Interactions

Near the epoch of CMB last scattering, dark matter accounts for about 65% of the energy budget of the Universe, hence making the CMB a particularly good probe of potential new physics in the dark matter sector. Of particular relevance to CMB-S4 studies, the presence of new dark matter interactions with light degrees of freedom [397, 398, 399, 400, 401, 402, 403, 404, 405, 406, 407, 290, 408, 409, 410, 411, 412, 413, 291, 414, 415, 416, 417, 292, 418, 419, 420, 421, 422, 423, 424, 425, 426, 427, 428, 429, 430] can leave subtle imprints on the temperature and polarization CMB power spectra. The introduction of such non-minimal dark matter models has been primarily (but not exclusively) motivated in the literature by potential shortcomings of the standard cold dark matter scenario at small sub-galactic scales [431, 432, 433, 434, 435, 436, 437, 438, 439,

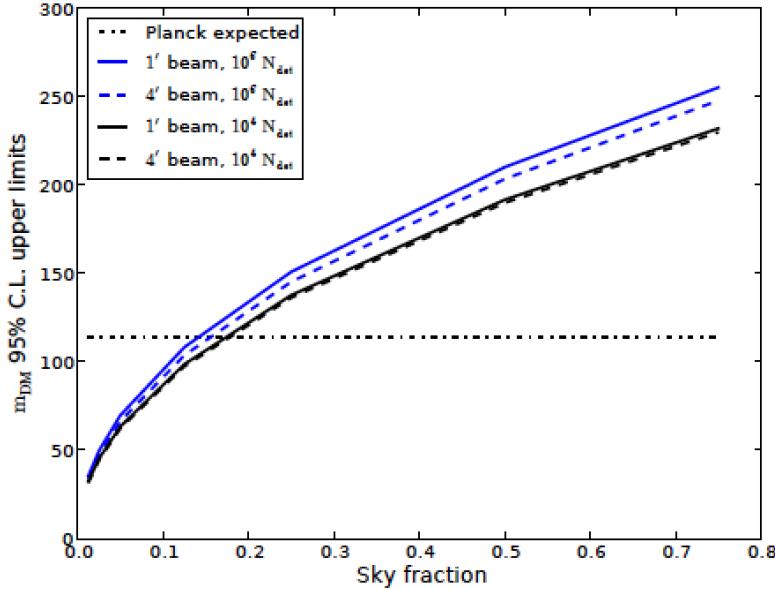


Figure 31. 95 % CL upper limit on m_{DM} in GeV as a function of sky coverage, f_{sky} . The blue/black lines correspond to $10^6/10^4$ detectors. The solid/dashed lines correspond to 1'/4' beams. The dashed/dotted lines show the limit from Planck for a thermal cross section and 100% of the energy absorbed by the plasma (it was expected from Fisher forecasts and then confirmed by the Planck collaboration measurements).

440, 441, 442]. While these issues are far from settled, they motivate the search for other non-minimal dark matter signatures in complementary data sets (such as the CMB) that could indicate whether or not dark matter can be part of the solution.

4.2.2.1 Dark Matter-Baryon Scattering

A possible non-standard dark matter scenario is that in which dark matter scatters off baryons in the early universe. In this scenario, there is a drag force produced by the baryons on the dark matter fluid, which affects the CMB temperature and polarization power spectra and the matter power spectrum. Ref. [423] has done a model-independent analysis on the Dark Matter-Baryon interactions using CMB temperature data from the *Planck satellite*, and the Lyman- α forest data from the *Sloan Digital Sky Survey*, as a tracer of the matter fluctuations. This analysis suggests that the constraints could become significantly better with better temperature data on small scales, and additional polarization data at large and small scales. Therefore, an experiment such as CMB-S4 would have a large impact on these constraints.

4.2.2.2 Dark Matter-Dark Radiation Interaction

Dark matter interacting with light (or massless) dark radiation has been put forward [429, 430] as a potential solution to the small discrepancy between the amplitude of matter fluctuations inferred from CMB measurements and those inferred from cluster number counts and weak lensing measurements. CMB-S4

measurements of the lensing power spectrum have the potential to significantly improve constraints on dark matter interacting with light degrees of freedom in the early Universe.

The key equations governing the evolution of cosmological fluctuations for this broad class of non-minimal dark matter models are presented in Ref. [443]. Essentially, the new dark matter physics enters entirely through the introduction of dark matter and dark radiation opacities, which, similarly to the photon-baryon case, prohibit dark radiation free-streaming at early times and provides a pressure term that opposes the gravitational growth of dark matter density fluctuations. The impact of this new physics on CMB fluctuations has been studied in detail in Ref. [420] and we briefly review it here. First, the presence of extra dark radiation mimics the presence of extra neutrino species and affects the expansion history of the Universe, possibly modifying the epoch of matter-radiation equality, the CMB Silk damping tail, and the early Integrated Sachs-Wolfe effect. However, unlike standard free-streaming neutrinos, the dark radiation forms a tightly-coupled fluid at early times, leading to distinct signatures on CMB fluctuations which include a phase and amplitude shift of the acoustic peaks (see e.g. Ref. [301, 444, 302]). Second, the dark radiation pressure prohibits the growth of interacting dark matter fluctuations on length scales entering the causal horizon before the epoch of dark matter kinematic decoupling. This weakens the depth of gravitational potential fluctuations on these scales, hence affecting the source term of CMB temperature fluctuations. Finally, the modified matter clustering in the Universe due to nonstandard dark matter properties will affect the lensing of the CMB as it travel from the last-scattering surface to us. For interacting dark matter models that are still allowed by the current Planck data, this latter effect is where CMB-S4 can significantly improve the constraints on these non-minimal theories.

Given the large array of possible dark matter theories to constrain, we use the Effective Theory Of Structure formation (ETHOS) [443] to systematically parametrize the deviations from standard cold dark matter. Within ETHOS, the impact of having all or a fraction of dark matter interacting with dark radiation can be captured with a handful of “effective” parameters which entirely determine the structure of the linear matter power spectrum. The most relevant parameters are [443]

$$\Xi_{\text{ETHOS}} = \left\{ \omega_{\text{DR}}, f_{\text{int}}, \{a_n, \alpha_l\} \right\}, \quad (4.3)$$

where $\omega_{\text{DR}} = \Omega_{\text{DR}} h^2$ is the physical energy density in dark radiation in units of the critical density of the Universe, f_{int} is the fraction of the total dark matter density that interacts with dark radiation, and where a_n and α_l are parameters describing the size of the interaction cross section and its angular dependence, respectively. The coefficients a_n enter directly into the calculation of the dark matter drag opacity κ_χ as:

$$\dot{\kappa}_\chi = -\omega_{\text{DR}} \sum_n \left(\frac{2+n}{3} \right) a_n \frac{(1+z)^{n+1}}{z_{\text{D}}^n}, \quad (4.4)$$

where z_{D} is a normalization scale. Choosing the latter to correspond to the dark matter decoupling redshift ensures that the a_n coefficients are of order unity. The index n is directly related to the nature of the physical process coupling dark matter and dark radiation: a non-vanishing a_n coefficient implies a scattering process characterized by the matrix element $|\mathcal{M}|^2 \propto (p_{\text{DR}}/m_\chi)^{n-2}$, where p_{DR} is the incoming momentum of the dark radiation and m_χ is the dark matter mass. Since the decoupling of dark matter from dark radiation is given by the approximate criterion $-\dot{\kappa}_\chi = H$, the magnitude of the ETHOS coefficients a_n set the scale at which the CMB lensing power spectrum departs from its ΛCDM counterpart. We use the ETHOS parametrization to illustrate that CMB-S4 can provide competitive constraints on partially interacting dark matter theories.

We illustrate in Fig. 32 the impact of different interacting dark matter models on the CMB lensing power spectrum. In the top panel, we show four partially-interacting dark matter models parametrized by their ETHOS opacity coefficient and for which only 5% of the total amount of dark matter is interacting. We

display the fractional difference between the ETHOS models and a standard Λ CDM model with vanishing neutrino mass. For comparison, we also illustrate the difference for a standard massive neutrino Λ CDM model with $\sum m_\nu = 0.06$ meV. Interestingly, the damping of the lensing power spectrum has a different shape than that caused by massive neutrinos. Given the expected performance of CMB-S4 in measuring the lensing power spectrum, all the model illustrated there (which are currently allowed by Planck data) could be ruled out, significantly improving our knowledge about interacting dark matter. The lower panel of Fig. 32 is similar, but illustrates how the fractional difference in the CMB lensing power spectrum is affected as the fraction of interacting dark matter is varied from 5 to 2 percents. Again, this illustrates that CMB-S4 can provide very tight constraints on the fraction of interacting dark matter.

Since non-standard dark matter models affect primarily the large CMB lensing multipoles, the constraining power of CMB-S4 on interacting dark matter is largely independent of the specific choice of ℓ_{\min} . We foresee that the main difficulty in constraining non-standard dark matter theories with CMB-S4 will be the proper modeling of non-linearities in the matter power spectrum, which are quite important for $\ell > 500$. We note that recent progress has been made in this direction [445].

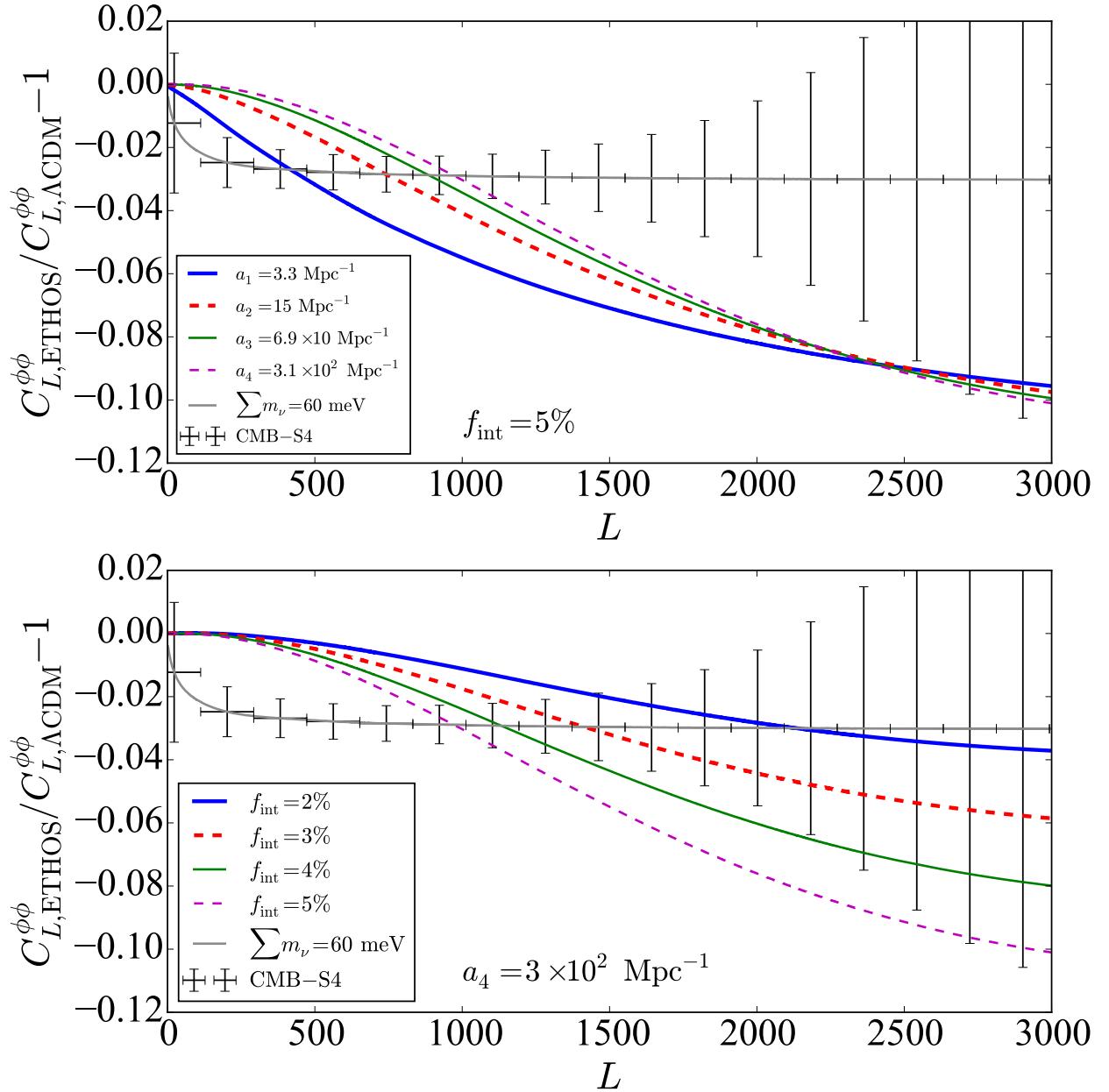


Figure 32. Top panel: Fractional difference of the CMB lensing spectrum between a standard ΛCDM model (with massless neutrinos) and four different ETHOS models with opacity coefficients a_n given in the legend. In all models shown, 5% of the dark matter is allowed to interact with dark radiation. For comparison, we also display a standard massive neutrino model with $\sum m_\nu = 0.06 \text{ meV}$. Lower panel: Similar to the top panel, but we now vary the fraction of dark matter that can interact with dark radiation, for a fixed opacity coefficient of $a_4 = 3 \times 10^2 \text{ Mpc}^{-1}$.

5

Physics Beyond The Standard Model

In addition to constraints on primordial parameters in the standard 6-parameter model, and a detection of (or upper limits on) the scalar-to-tensor ratio, CMB-S4 will yield unprecedented constraints on interesting physics beyond the standard picture.

5.1 Constraints on Axion-like particles

The QCD axion and other axion-like particles (ALPs), if stable on cosmological timescales, can contribute to the DM density. Along with thermal WIMPs, they are a well-motivated DM candidate (see Ref. [446] for a recent review).

Non-thermally produced ultralight axions (ULAs) with masses in the range $10^{-33} \text{ eV} \leq m_a \leq 10^{-20} \text{ eV}$ are well motivated by string theory, can contribute to either the dark matter or dark energy components of the Universe, depending on their masses, and are distinguishable from DE and CDM in cosmological observables. The current best constraints from the primary CMB TT power, and WiggleZ galaxy redshift survey were made in [447]. Our fiducial axion energy density is chosen to be consistent with these constraints.

5.1.1 Constraints on the axion energy density

The degeneracies of the axions with other cosmological parameters, such as N_{eff} or m_ν , vary depending on the axion mass (see Fig. 33, right panel). Dark energy-like axions with masses around 10^{-33} eV change the late-time expansion rate and therefore the sound horizon, changing the location of the acoustic peaks. This has degeneracies with the matter and curvature content. Heavier axions ($m_a \gtrsim 10^{-26} \text{ eV}$) affect the expansion rate in the radiation era and reduce the angular scale of the diffusion distance, leading to a boost in the higher acoustic peaks, which has a degeneracy with N_{eff} .

In both of these cases, improved errors on the temperature and polarization power spectrum, coupled with constraints on the Hubble constant (for the lightest axions) from Baryon Acoustic Oscillations, lead to improvements in the error on allowed axion energy density of a factor of three from these spectra alone.

In the matter power spectrum, and thus CMB lensing power, light axions suppress clustering power, suggesting a degeneracy with effects of massive neutrinos that must be broken to make an unambiguous measurement of neutrino mass using the CMB. The above-mentioned effects in the expansion rate break this degeneracy for some axion masses. There remains a significant degeneracy between axions and massive neutrinos $m_a = 3 \times 10^{-29} \text{ eV}$ and $\Sigma m_\nu = 60 \text{ meV}$. Effort should be made to break this degeneracy and distinguish the effects of non-thermal axions from massive neutrinos for an unambiguous detection of neutrino mass using the CMB.

We show the forecasted constraints on the axion energy density from S4 including lensing in the left panel of Figure 33 (for fixed neutrino mass of $\Sigma m_\nu = 0.06 \text{ eV}$). Adding in information from the lensing reconstruction using S4 will improve constraints on axion DM significantly. A percent-level measurement of the lensing

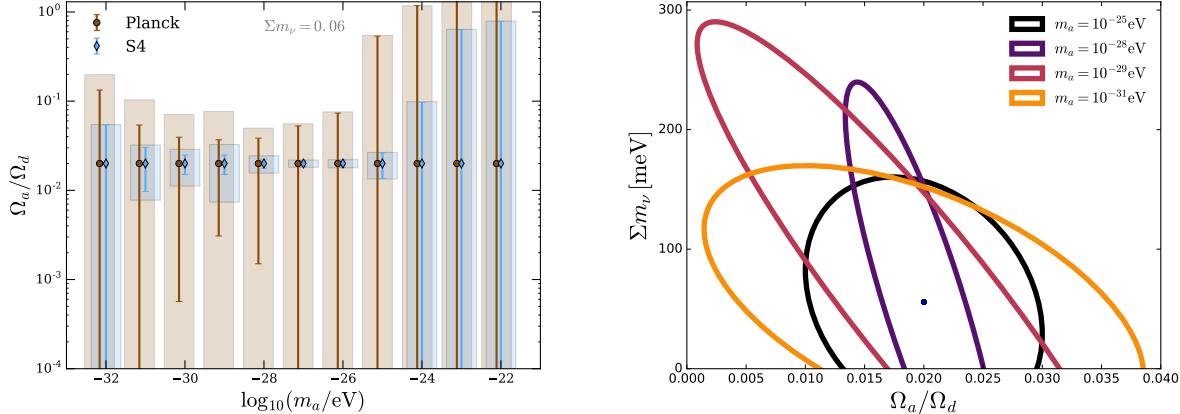


Figure 33. *Left:* Constraints on the axion energy density as a function of axion mass at fixed neutrino mass $\Sigma m_\nu = 0.06$ eV (errorbars) and marginalising over the neutrino mass (shaded error bars), for Planck Blue Book constraints (brown) and a CMB-S4-like survey (blue). Over the ‘fuzzy’ dark matter region ($-28 < \log(m_a/\text{eV}) < -25$), S4 allow for percent-level constraints on an axion component, improving significantly on current constraints. For the same range of masses, the degeneracy is weakest with massive neutrinos - the bands shown when marginalising over the neutrino error are not much larger than the case where the neutrino mass is fixed. *Right:* Degeneracy of axions with massive neutrinos. There is a significant degeneracy for $m_a = 3 \times 10^{-29}$ eV and $\Sigma m_\nu = 60$ meV. Figure from Hložek et al. (prep).

deflection power at multipoles $\ell > 1000$ leads to an improvement in the error on the axion energy density of a factor of eight relative to the current Planck constraints, for an axion mass of $m_a = 10^{-26}$ eV. This represents an ability to test the component nature of dark matter, and thus the CDM paradigm, at the percent level. Furthermore, since $\Omega_a \propto f_a^2$ this improves the expected constraint on the axion decay constant from 10^{17} GeV with Planck to 10^{16} GeV with S4, testing the predictions of the “string axiverse scenario” [305].

Planck is degenerate with CDM at $m_a = 10^{-24}$ eV, and only has weak constraints at $m_a = 10^{-25}$ eV. CMB-S4 introduces a 5σ constraint at $m_a = 10^{-25}$ eV, and improves the lower bound on DM particle mass to $m_a = 10^{-23}$ eV and fractions O(10%). However, realising this requires improved understanding of non-linear clustering in axions [448].

5.1.2 Axion Isocurvature

A key axion parameter is the symmetry breaking scale, f_a . If $H_I/2\pi < f_a$, the axion DM acquires *uncorrelated isocurvature perturbations* (e.g. Refs. [449, 450, 451]).¹ The uncorrelated CDM isocurvature amplitude is bounded by *Planck* to be $A_I/A_s < 0.038$ at 95% C.L. [188]. We performed a forecast for CMB-S4: the isocurvature limit will be improved by a factor of approximately five compared to *Planck*, allowing for detection of axion-type isocurvature at 2σ significance in the region $0.008 < A_I/A_s < 0.038$.

¹We ignore the case where $H_I/2\pi > f_a$, since no isocurvature initial conditions are excited. The limit $r_{0.05} < 0.12$ implies that isocurvature is produced if $f_a > 1.8 \times 10^{13}$ GeV. This accounts for the QCD axion in the “anthropic” window (roughly half of the allowed range of f_a on a logarithmic scale), axions with GUT scale decay constants (such as string axions [452, 305]) and axions with lower f_a in models of low-scale inflation.

The axion isocurvature amplitude is:

$$A_I = \left(\frac{\Omega_a}{\Omega_d} \right)^2 \frac{(H_I/M_{\text{Pl}})^2}{\pi^2(\phi_i/M_{\text{Pl}})^2}. \quad (5.1)$$

The initial axion displacement, ϕ_i , fixes the axion relic abundance such that $\Omega_a = \Omega_a(\phi_i, m_a)$ [453, 454, 455, 456, 457, 458]. Thus, if the relic density and mass can be measured by independent means, *a measurement of the axion isocurvature amplitude can be used to measure the energy scale of inflation, H_I*

If the QCD axion is all of the DM, axion direct detection experiments can be used in conjunction with CMB-S4 to probe H_I in the range

$$2.5 \times 10^6 \lesssim H_I/\text{GeV} \lesssim 4 \times 10^9 \text{ (QCD axion + direct detection)} \quad (5.2)$$

This is demonstrated in Fig. 34 (left panel) for the case of ADMX [459] (in operation), and CASPER [460] (proposed), where we have used the standard formulae relating the QCD axion mass and relic abundance to the decay constant (e.g. Ref. [450]).² *Combining axion DM direct detection with CMB-S4 isocurvature measurements allows a unique probe of low-scale inflation, inaccessible to searches for tensor modes.*

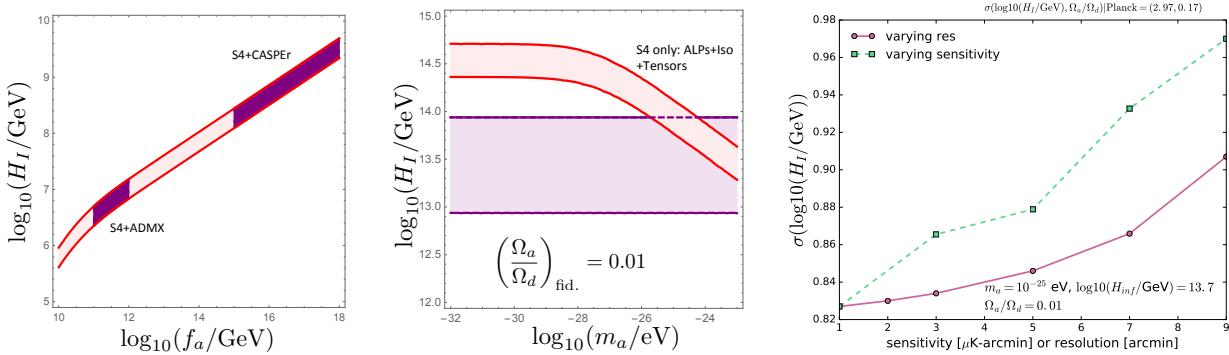


Figure 34. Axion dark matter isocurvature. Red bands in the left and middle panels show the isocurvature amplitude consistent with Planck and detectable with S4. **Left Panel:** The QCD axion: measuring the energy scale of inflation with S4+axion direct detection. Here we restrict axions to be all of the DM. The purple regions show the range of f_a accessible to axion direct detection experiments. Combining ADMX [459] (in operation), CASPER [460] (proposed), and S4 it is possible to measure $4 \times 10^5 \lesssim H_I/\text{GeV} \lesssim 4 \times 10^9$. **Middle Panel:** ALPs - a combination measurement using S4 alone. Assuming 1% of the total DM resides in an ultralight axion, the mass and axion density can be determined to high significance using, for example, the lensing power. The isocurvature amplitude can also be determined, allowing for an independent determination of H_I in the same regime as is accessible from tensor modes (purple band). **Right Panel:** We vary H_I directly around a fiducial model of $H_I/\text{GeV} = 10^{13.7}$ for axions of mass 10^{-25} eV making up 1% of the total dark content, for a range of possible S4 survey parameters. While the error degrades as the resolution and sensitivity are worsened, this degradation is small compared to the factor of three improvement in the error moving from Planck to CMB-S4.

We now consider isocurvature in ultralight ALPs (ULAs, see e.g. Refs. [466, 462]). ULA DM has a number of distinctive features in large scale structure and the CMB [447, 467]. For ULAs with $10^{-32} \lesssim m_a/\text{eV} \lesssim 10^{-23}$ a DM fraction of Ω_a/Ω_d in the range of 1% is consistent with Planck [447] and high- z galaxy formation [468, 469], yet can be distinguished from pure CDM using S4 lensing power at $> 2\sigma$ (depending on the ULA

²In simple models of inflation, the high- f_a QCD axion is incompatible with detection of tensor modes [450, 451, 461, 462, 461], although non-standard cosmic thermal histories of PQ breaking mechanisms can lift constraints, e.g. [463, 464, 465].

mass, Sec. ??). Fig. 34 (middle panel) shows isocurvature constraints possible with S4, compared to tensor constraints. We fix the fiducial ULA fraction to 1%, such that Ω_a and m_a can be separately measured using the S4 lensing power, and thus using Eq. (5.1) a measurement of A_I is a measurement of H_I .

In contrast to the QCD axion, there are masses, $m_a \lesssim 10^{-26}$ eV, for which tensor modes impose a stronger constraint on H_I than isocurvature (such that isocurvature in these ALPs would be undetectably small). However, there are also regions of overlap between possible tensor and isocurvature measurements. Using S4 in these regions, it is possible to make a combination measurement of isocurvature and axion parameters, giving an independent measurement of H_I :

$$2.5 \times 10^{13} \lesssim H_I/\text{GeV} \lesssim 10^{14} \text{ (ultralight ALPs, S4 alone)} . \quad (5.3)$$

This applies to ALPs in the mass range $10^{-26} \lesssim m_a/\text{eV} \lesssim 10^{-23}$, where effects on lensing of a 1% axion fraction can be distinguished from CDM. *Detecting isocurvature and lensing effects from ULAs using CMB-S4 can provide a measurement of H_I complementary to searches for tensor modes.*

5.1.3 Constraints on Axion couplings

Two very common features of models containing ALPs are that the ALPs are typically light (in many cases, $m \ll 1$ eV) and their interactions are suppressed by powers of the (typically large) decay constant f_a . These two features make ALPs a particularly compelling target for cosmology and N_{eff} specifically. Because of the small masses, they will often behave as relativistic species in the early universe. Furthermore, because their production rate will scale as T^{2n+1}/f_a^{2n} for some $n \geq 1$, they are likely to be thermalized at high temperatures. Given that $\Delta N_{\text{eff}} > 0.027$ under such circumstances, a CMB experiment with sensitivity at this level will be sensitive to a very wide range of ALP models. In the absence of a detection of ΔN_{eff} , we can place constraints on the axions couplings to the Standard model.

Two couplings of particular interest for axion phenomenology are the coupling to gluons and photons,

$$\frac{1}{4}g_{a\gamma\gamma}a\tilde{F}_{\mu\nu}F^{\mu\nu}, \quad \frac{1}{4}g_{agg}a\tilde{G}_{\mu\nu}G^{\mu\nu} . \quad (5.4)$$

These couplings typically appear as the consequence of chiral anomalies. The coupling of the axion to gluons is what makes the solution to the strong-CP problem possible. The coupling to photons is somewhat model dependent but typically arises in conjunction with the gluon coupling. In addition to or instead of these couplings, a variety of other possible couplings to matter may also be included.

The coupling of axions to matter has an additional feature that it can bring axions into thermal equilibrium at low temperatures. Specifically, the lowest dimension coupling of an axion to charged matter takes the form

$$\mathcal{L} = -\frac{\partial_\mu a}{\Lambda_\psi}\bar{\psi}_i(\gamma^\mu g_V^{ij} + g_A^{ij}\gamma^5)\psi_j \quad (5.5)$$

where ψ_i is any of the charged fermions of the Standard Model and i, j label the three generations of fermions of with the same charges. Above the scale of electroweak symmetry breaking (EWSB), this coupling leads to an abundance of axions with $\Delta N_{\text{eff}} = 0.027$. Through freeze-out, we are again very sensitive to g_V and g_A at levels that vastly exceed current limits. In addition, below the scale of EWSB, this coupling can bring the axions into thermal equilibrium at low temperatures (freeze-in) below the mass of the heaviest fermion, $T_F \lesssim m_3$. Freeze-in will produce $\Delta N_{\text{eff}} \approx 0.05$ and is therefore easier to detect. For reheating temperatures well above the electroweak scale, the sensitivity of freeze-out exceeds that of freeze-in, although both are far more sensitive than current limits on axion couplings to second and third generation fermions.

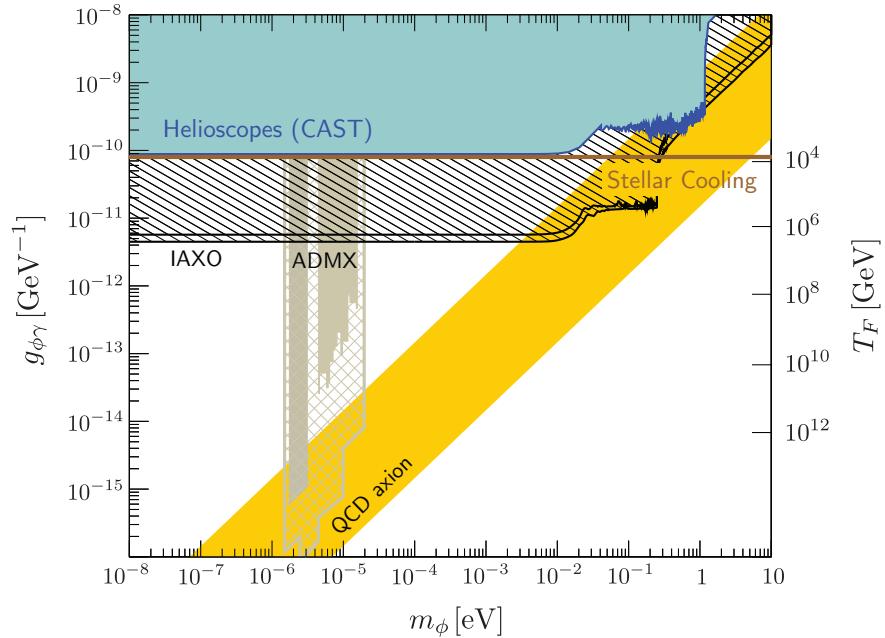


Figure 35. Relationship between freeze-out temperature (T_F) and axion-photon coupling ($g_{a\gamma\gamma}$). While freeze-out is independent of the mass, experimental probes of the coupling are strongly mass dependent. For $T_F > 10^4$ GeV the axion coupling would predict $\Delta N_{\text{eff}} = 0.027$. Sensitivity to ΔN_{eff} at that level would translate into sensitivity to axions with $T_F \leq T_{\text{reheat}}$. For plausible values of $T_{\text{reheat}} > 10^{10}$ GeV, cosmology is orders of magnitude more sensitive than existing bounds over a wide range of masses.

The coupling to matter is motivated also by the approximate $U(3)^5$ flavor symmetry of the Standard Model. It is natural for such couplings to arise if the axion is a goldstone boson that results from spontaneous breaking of this symmetry (or a sub-group). Given the non-abelian nature of the flavor symmetry, these scenarios can often lead to many axions (also known as familons). Under such circumstances, the contribution to freeze-out is given by

$$\Delta N_{\text{eff}} = N_a \times 0.027 \quad (5.6)$$

where N_a is the number of axions / number of broken generators of the symmetry group. It is easy to find scenarios where $N_a \sim \mathcal{O}(10)$ which is at the current level of sensitivity.

Status of current observations – Current constraints on ALPs arise from a combination of experimental [470], astrophysical [471], and cosmological [446] probes. Current cosmological constraints are driven by several effects that depend on the mass of the axion. For axion masses greater than 100 eV, stable thermal ALPs are easily excluded because they produce dark matter abundances inconsistent with observations. By including the free-streaming effects of thermal QCD-axions, Planck data [472] combined with local measurements provide the constrain $m_a < 0.525$ eV (95 % CL). At larger masses, ALPs become unstable and can be constrained by the change to N_{eff} from energy injection as well as from spectral distortions and changes to BBN [473, 302].

Implications for CMB Stange IV – Sensitivity to $\Delta N_{\text{eff}} = 0.027$ is sufficient to probe the entire mass range of ALPs down to $m_a = 0$ under the assumption that it thermalized in the early universe. Interpreting such bounds in terms of the couplings of axions is more complicated [326] and can depend on assumptions about the reheating temperature. For high (but plausible) reheat temperatures of 10^{10} GeV, CMB stage IV would

be sensitive to $g_{a\gamma\gamma}$, $g_{agg} < 10^{-13}\text{GeV}^{-1}$ [474] as illustrated in figures 35 and 36. These projected limits exceed current constraints and future probes for a range of possible axion masses (including the QCD axion).

The implications for the couplings to matter are similar for the contribution from freeze-out. However, the freeze-in contribution of $\Delta N_{\text{eff}} \gtrsim 0.05$ would be easier to exclude experimentally, but still produces the limits [474]

$$\Lambda_{\psi_i} > \begin{cases} 1.3 \times 10^8 \text{ GeV} \left(\frac{g_{*,i}}{g_{*,\tau}} \right)^{-1/4} \left(\frac{m_i}{m_\tau} \right)^{1/2} & i = \text{leptons}, \\ 2.1 \times 10^9 \text{ GeV} \left(\frac{g_{*,i}}{g_{*,t}} \right)^{-1/4} \left(\frac{m_i}{m_t} \right)^{1/2} & i = \text{quarks}. \end{cases} \quad (5.7)$$

where $g_{*,i}$ is the number of degrees of freedom at temperature $T = m_i$. For second and third generation fermions, these limits would exceed current bounds by several orders of magnitude.

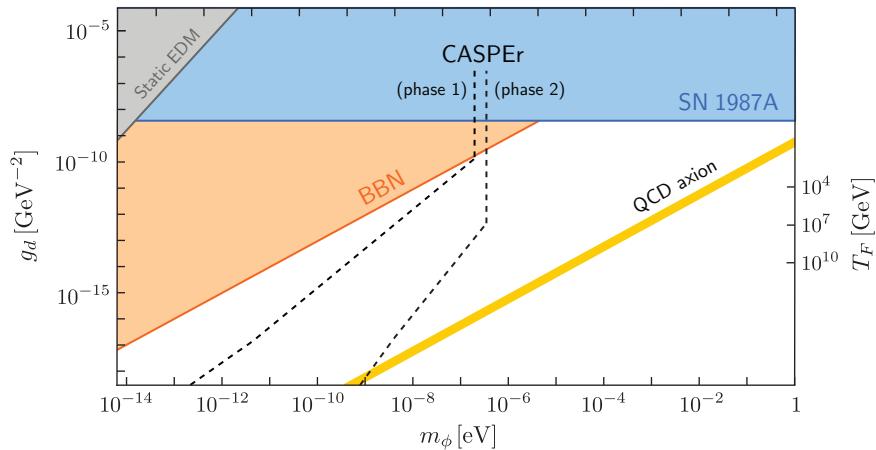


Figure 36. Relationship between freeze-out temperature (T_F) and axion-gluon coupling via the neutron dipole (g_d). While freeze-out is independent of the mass, experimental probes of the coupling are strongly mass dependent. We see that cosmology is more sensitive than current limits even for freeze-out temperatures below the electro-weak phase transition and therefore could be probed by $\Delta N_{\text{eff}} > 0.027$.

5.2 Cosmic Birefringence

The simplest dynamical way to model the accelerated expansion of the universe is to invoke a new slowly evolving scalar field that dominates its energy budget (the quintessence models for DE). Such a field generically couples to photons through the Chern-Simons term in the electromagnetic Lagrangian, causing linear polarization of photons propagating cosmological distances to rotate—the effect known as cosmic birefringence [475]. In the case of the CMB, such rotation converts the primordial E mode into B mode, producing characteristic TB and EB cross-correlations in the CMB maps [476, 477]. Even though there is no firm theoretical prediction for the size of this effect, if observed, it would be a clear smoking-gun evidence for physics beyond the standard model in the form of a new scalar field. Previous studies have used quadratic estimator formalism to constrain this effect [478], with the best current limit coming from sub-degree scale polarization measurements with POLARBEAR [479] ($< 0.33 \text{ deg}^2$ for the amplitude of a scale-invariant rotation-angle power spectrum). A promising way to pursue search for cosmic birefringence in the future

is measurement of the off-diagonal EB cross correlations on small angular scales, and the measurement of polarization anisotropy on a wide range of scales is going to be essential for achieving this.

Fig. 37 shows the current upper limit on the rotation-angle power spectrum from POLARBEAR and a projection for Planck, and a forecast for a Stage-IV experiment (with noise of $1.41 \mu\text{K}\text{-arcmin}$ in polarization, and a resolution of $1'$). The improvement from the current constraint at all multipoles is about two orders of magnitude. We assumed access to polarization modes from $\ell = 30$ to $\ell = 5000$.

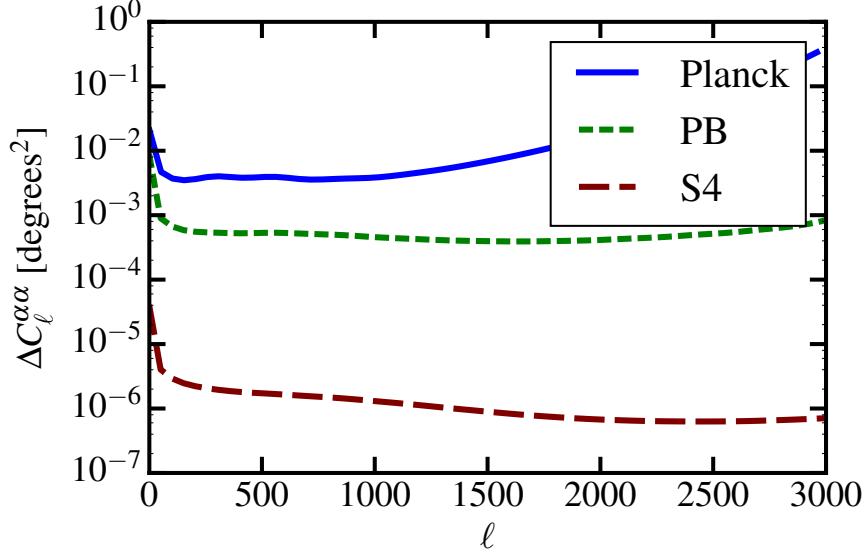


Figure 37. The current (from POLARBEAR, labeled as PB) and projected (for Planck and Stage-IV experiment) 1σ errorbars on the birefringent rotation-angle power spectrum are shown on the vertical axis. A Stage-IV has the potential to improve the current best constraint on anisotropic birefringence by more than two orders of magnitude at all multipoles. For the Stage-IV forecast, we assumed noise of $1.41 \mu\text{K}\text{-arcmin}$ (in polarization), a resolution of $1'$, and have considered polarization modes from $\ell = 30$ to $\ell = 5000$.

For a fixed integration time (and a varied noise level and sky coverage), large sky coverage optimizes sensitivity to low multipoles of the rotation angle and gives the best signal-to-noise ratio for rotation models that have power on large scales (such as, for example, a model with a scale-invariant power spectrum, which could result from fluctuations in a spectator scalar field present during inflation). Conversely, for models that have power on scales corresponding to multipoles above $\ell \sim 1000$, best signal-to-noise is achieved with deeper integration on small sky patches. For a measurement of the magnitude of the quadrupole of the rotation angle, reducing the resolution from $1'$ to $9'$ produces a factor of a few increase in the projected errorbar (for all other parameters fixed). Increasing the noise from 1.41 to $12.7 \mu\text{K}\text{-arcmin}$ produces a factor of about 20 increase in the errorbar. Access to polarization modes down to $\ell = 2$ does not significantly affect the forecasts.

5.3 Primordial Magnetic Fields

The origin of the microgauss (μG) strength magnetic fields in galaxies and galaxy clusters is one of the long standing puzzles in astrophysics [480]. It is challenging to explain such fields based solely on the dynamo mechanism, without there being some initial seed field. However, if magnetic fields were present in the early

universe, they would remain frozen in the cosmic plasma and collapse with the rest of the matter to form the galactic fields [481], or at least provide the seeds for the dynamo. A primordial magnetic field (PMF) could be produced in the aftermath of cosmic phase transitions [482] or in specially designed inflationary scenarios [483, 484]. Detecting their signatures in the CMB temperature and polarization would decisively prove their primordial origin. Aside from explaining the galactic fields, bounds on PMF have profound implications for our understanding of the early universe. They help constrain theories of inflation [485], models of the QCD and electroweak phase transitions [486] and baryogenesis [487].

A stochastic PMF affects CMB in several ways. Magnetic stress-energy induces scalar, vector and tensor mode perturbations in the metric, and the Lorentz force generates vorticity in the photon-baryon fluid [488, 489, 490, 491, 492]. Dissipation of PMF on small scales dumps energy into the plasma, which produces spectral distortions and affects the recombination history [493]. Finally, Faraday Rotation (FR) of CMB polarization converts some of the E -modes into B -modes [494, 495].

Stochastic PMF has two potentially observable frequency independent contributions to the B -mode spectrum [491]. One comes from the passive, or uncompensated tensor mode, which is generated by the PMF before neutrino decoupling. For nearly scale-invariant PMF, the spectrum of this component is indistinguishable from the inflationary gravity wave signal. The amplitude of this tensor contribution is proportional to $B_{1\text{Mpc}}^4 [\ln(a_\nu/a_{\text{PMF}})]^2$, where $B_{1\text{Mpc}}$ is the PMF strength smoothed over 1Mpc, a_ν is the scale factor at neutrino decoupling and a_{PMF} is the scale factor at which PMF was generated. The other is the PMF vector mode which peaks at $l \sim 2000$, with the precise peak position dependent on the PMF spectrum. The vector-mode contribution is independent of a_{PMF} .

Planck data limits the magnetic field strength to $B_{1\text{Mpc}} < 4.4$ nanogauss (nG) at the 95% confidence level [496]. Similar bounds were recently obtained by POLARBEAR [479] based on their B-mode spectrum alone.

A Stage-IV experiment can improve the 95% bound to 0.6 nG based on the PMF vector mode contribution to the B-mode spectrum. Comparable bounds can be obtained from the mode-coupling correlations induced by Faraday Rotation. The mode-coupling is the same as in the case of birefringence discussed in Sec. 5.2, except for the Faraday Rotation being frequency dependent.

5.4 Cosmic Strings

Model	Planck	CMB-S4 (1' resolution)
Fixed α_{str} :	$\sigma(f_{10}) = 0.015$	$\sigma(f_{10}) = 1.06 \times 10^{-3}$
Varying α_{str} :	$\sigma(f_{10}) = 0.017$	$\sigma(f_{10}) = 1.85 \times 10^{-3}$
	$\sigma(\alpha_{\text{str}}) = 5.55$	$\sigma(\alpha_{\text{str}}) = 0.64$

Table 5-1. Forecast constraints on the fraction of power in cosmic strings at $\ell = 10$ and on the ‘wiggleness’ of the string network, α . Planck forecast is based on Blue Book values, with $f_{\text{sky}} = 0.75$. CMB-S4 will yield an order of magnitude improvement in constraints on cosmic string parameters.

Cosmic strings can at most contribute $O(1\%)$ to the total CMB temperature anisotropy [98, 497, 498], however, they can still generate observable B-modes. As shown in [499], the bounds on cosmic strings obtained solely from the POLARBEAR [500] and BICEP2 [501] B-mode spectra are comparable to those from temperature spectra. We forecast the predicted constraints on cosmic strings using the StringFast code [502], based on the CMBACT simulations [503] of a general string network, which allows for the correlation length of the strings, the ‘wiggleness’ (which controls the small-scale structure of the string network) and the

string rms velocity. StringFast allows for fast computation of the relevant string spectra, and includes the contribution to the string spectrum from scalar, vector and tensor modes, which are most relevant for the string B-modes [502].

In keeping with the methodology of recent results, we compute the string spectrum with a value of the string tension ($G_\mu/c^2 = 1.97 \times 10^{-6}$) that allows strings to make up all the TT power at $\ell = 10$, and then use the fraction of the spectrum at that multipol f_{10} as the forecast parameter. f_{10} scales with the string tension as $f_{10} \propto G_\mu^2$. The Fisher projections for Planck around a fiducial model of $f_{10} = 0.01$ are $f_{10} < 0.032$, which maps to $G_\mu/c^2 < 3.5 \times 10^{-7}$ 95%CL, is consistent with the Planck constraints on the AH-mimic spectra. We assume a fiducial model for the ‘wiggleness’ α_{str} , string velocity v_{str} and correlation length ξ_{str} of 1.05, 0.4 and 0.35 respectively, in keeping with the model assumed in [502]. We consider models where only the string fraction is varied, and the additional model where the small-scale structure of the string network is varied. The constraints are summarised in Table 5.4 for a baseline resolution of 1 arcminute. The error on the string fraction for a 2 arcminute beam is only degraded by a few percent relative to the nominal case. In addition, the constraints are not strongly improved with the addition of BAO data, or with a more improved measurement of τ .

6

CMB Lensing

(send feedback on this chapter to s4-lensing@cosmo.uchicago.edu)

6.1 Introduction to CMB Lensing

As CMB photons travel from the last scattering surface to Earth, their travel paths are bent by interactions with intervening matter in a process known as *gravitational lensing*. This process distorts the observed pattern of CMB anisotropies, which has two important consequences:

- CMB lensing encodes a wealth of statistical information about the entire large-scale structure (LSS) mass distribution, which is sensitive to the properties of neutrinos and dark energy.
- CMB lensing distortions obscure our view of the primordial Universe, limiting our power to constrain inflationary signals; removing this lensing noise more cleanly brings the early Universe and any inflationary signatures into sharper focus.

Gravitational lensing of the CMB can be measured by relying on the fact that the statistical properties of the primordial CMB are well known. The primordial (un-lensed) CMB anisotropies are statistically isotropic. Gravitational lensing shifts the apparent arrival direction of CMB photons, which breaks the primordial statistical isotropy; lensing thus correlates previously independent Fourier modes of the CMB temperature and polarization fields. These correlations can be used to make maps of the LSS projected along the line-of-sight; see the discussion in Section 6.2.1.

A CMB-S4 experiment will make radical improvements in CMB lensing science: high sensitivity will enable lensing maps that have much higher signal to noise; the high polarization sensitivity will allow lensing maps that are much less sensitive to foreground contamination; multi-frequency coverage will greatly reduce foreground contamination in the temperature-based lensing estimates, allowing lensing maps with higher resolution; large area coverage will provide maps for cross-correlation with large scale structure for next generation surveys, including Euclid and LSST.

The information contained in lensing mass maps can be accessed and used in several ways. First, the power spectrum of the lensing deflection map is sensitive to any physics that modifies how structure grows, such as dark energy, modified gravity, and the masses of neutrinos. In Section 6.2.2, we discuss how the lensing power spectrum is measured, and in Section 6.6, we give parameter forecasts combining primordial CMB power spectrum and CMB lensing power spectrum measurements. Second, lensing mass maps can be compared to other tracers of LSS at lower redshifts such as the distribution of galaxies and optical weak lensing shear maps. By cross correlating, for example, CMB lensing and optical shear mass maps, which are each derived from lensed sources at widely differing redshifts, one can enhance dark energy constraints and improve the calibration of systematic effects. Cross-correlation science with CMB lensing maps is discussed in Section 6.3. Finally, lensing distortions partially obscure potential signatures of cosmic inflation in the primordial

B-mode polarization signal. With precise measurements, this lensing-induced noise can be characterized and removed in a procedure known as “delensing.” Because B-mode polarization measurements from CMB-S4 are expected to be lensing-noise dominated, delensing will be critical to maximize the information we can infer about cosmic inflation; see the discussion in Section 6.4.

We discuss systematics from astrophysical and instrumental effects that can impact the lensing signal as well as ways to mitigate them in Section 6.5. Section 6.6 describes forecasted parameter constraints when including CMB lensing measurements as well as the instrument requirements for CMB-S4 to maximize the science gain from CMB lensing.

6.2 Measuring CMB Lensing

6.2.1 Constructing a Lensing Map

A map of the CMB lensing deflection field is a direct probe of the projected matter distribution that exists in the observable Universe. This lensing map is a fundamental object for nearly all areas of CMB lensing science: it is used to measure the lensing power spectrum, measure cross correlations between CMB lensing and external data sets, and to de-lens maps of the B-mode polarization.

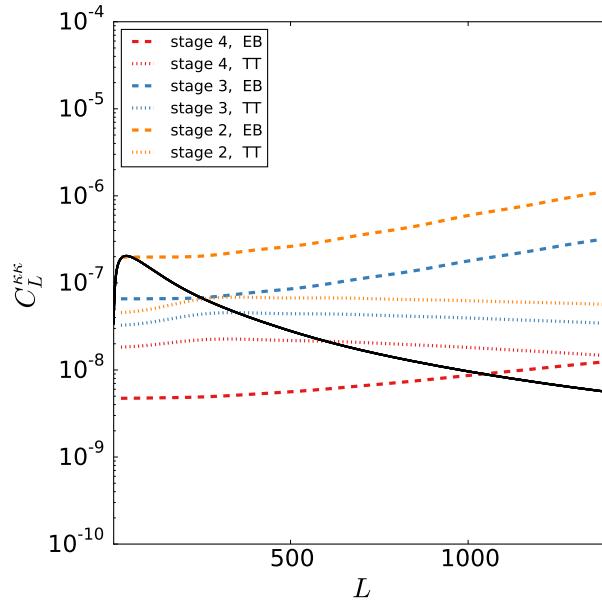


Figure 38. Signal and noise-per-mode curves for three experiments. “Stage 2” is meant to represent a current-generation survey like SPTpol or ACTPol and has $\Delta_T = 9\mu K^-1$; “Stage 3” is an imminent survey like SPT-3G or AdvACT, with $\Delta_T = 5\mu K^-1$; and “Stage 4” has a nominal noise level of $\Delta_T = 1\mu K^-1$. These noise-per-mode curves do not depend on the area of sky surveyed. All experiments assume a $1.4'$ beam.

To date, all maps of the lensing field have been constructed using the quadratic estimator by Hu & Okamoto 2002. This estimator uses information about the off-diagonal mode-coupling in spherical harmonic space that lensing induces to reconstruct the deflection field. An estimate for the amount of lensing on a given scale is obtained by averaging over pairs of CMB modes in harmonic space separated by this scale. CMB-S4

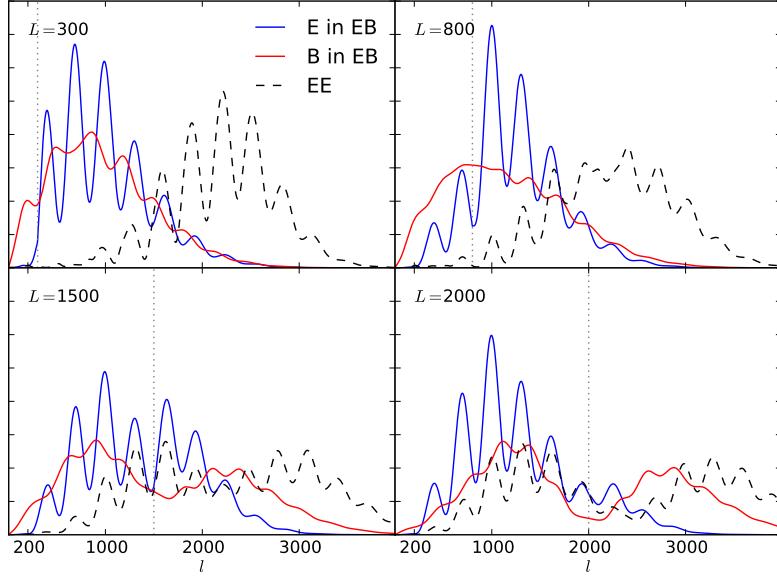


Figure 39. Contributions from CMB scales (ℓ) to lensing reconstruction on four lensing scales (L). The EB estimator is expected to be the main channel for lensing science with CMB-S4. On degree and sub-degree scales, $L = 300$ and 800 , the estimator uses B and E modes at $\ell \sim 1000$. On scales of several arcmin, $L = 1500$ and 2000 , the estimator uses B modes on significantly smaller scales. Figure taken from Pearson et al. 2014

will greatly improve over existing measurements by having high angular resolution at high sensitivity to both temperature and polarization.

One way that the high angular resolution and sensitivity of CMB-S4 improves upon the Planck measurement is simply by increasing the number of CMB modes imaged on scales smaller than the Planck beam. Imaging CMB modes between $l = 2000$ and 4000 , which can be achieved with CMB-S4 yields considerable gain in the accuracy of the lensing power spectrum measurement.

However, the primary reason for the increased power of CMB-S4 lensing measurements is this experiment's ability to measure CMB polarization with unprecedented sensitivity. To date, all CMB lensing results have had their signal-to-noise dominated by lensing reconstructions based on CMB temperature data (cite). Such lensing measurements in temperature are limited for two reasons. First, they are limited by systematic biases from astrophysical foregrounds and atmospheric noise. Second, the signal-to-noise on lensing measurements from temperature is intrinsically limited by the cosmic variance of the unlensed CMB temperature field. Due to the unprecedented sensitivity of CMB-S4, the bulk of the lensing signal-to-noise will now be derived from CMB polarization data; polarization lensing reconstruction will allow CMB-S4 to overcome both of these limitations. First, the challenges of astrophysical emission and atmospheric noise are much reduced in polarization data. Second, low-noise polarization lensing measurements are not limited by primordial CMB cosmic variance, because they make use of measurements of the B -mode polarization, which contains no primordial signal on small scales. To fully exploit the lack of limiting primordial signal in the B -mode polarization, maximum likelihood lensing reconstruction algorithms can be used, which use iteration to surpass the quadratic estimator. This iterative lensing reconstruction procedure is discussed in more detail in Section 6.4.

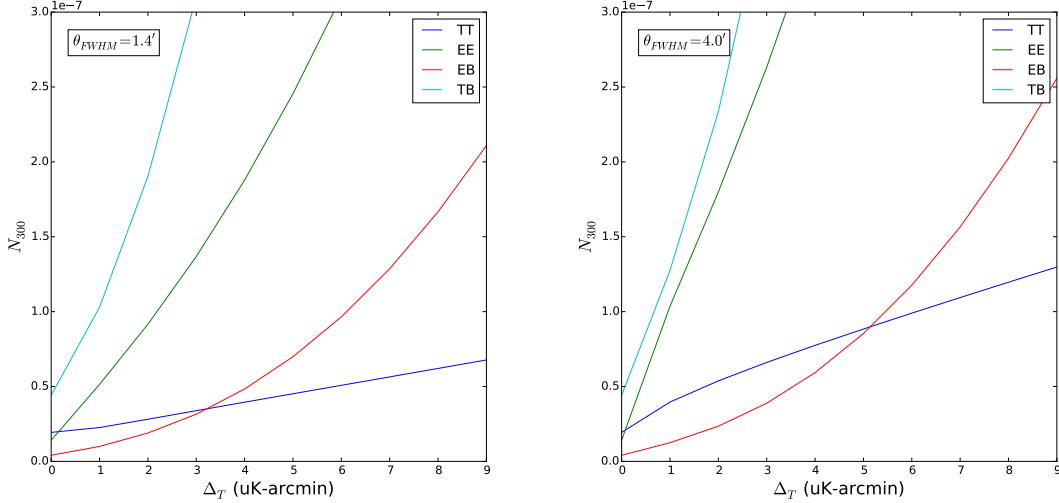


Figure 40. Noise per mode in the lensing field for different lensing estimators at $L = 300$. Left panel is for 1.4 arcmin resolution, and right panel is for 4 arcmin resolution. For a 1.4 and 4 arcmin resolution experiment, the EB polarization estimator yields lower noise than the temperature estimator, below 3 uK-arcmin and 5 uK-arcmin noise in temperature respectively.

6.2.2 Lensing Power Spectrum

The power spectrum of reconstructed CMB lensing maps is a measure of the matter power spectrum integrated over redshift. The lensing power spectrum has a broad redshift response kernel, with most of the contribution coming from $z \sim 1 - 5$, with a peak at $z \sim 2$ (see Figure 43). Most of the scales probed by the lensing power spectrum are on sufficiently large scales that they are mainly in the linear regime. As such, the lensing power spectrum is sensitive to physics which affects the growth of structure on large scales and at high redshift, such as the mass of the neutrinos.

The latest measurements of the CMB lensing autospectrum, as of early 2016, are shown in Figure 41. The first detections were obtained by the Atacama Cosmology Telescope (ACT; Das+ 2011) and South Pole Telescope (SPT; van Engelen+ 2012) teams, who analyzed maps of several hundreds of square degrees yielding precisions on the lensing power spectrum of approximately 25% and 18% respectively. The Planck collaboration has since provided all-sky lensing maps whose precision on the power spectrum amplitude is approximately 4% in the 2013 data release and 2.5% in the 2015 data release. The first detections of the lensing autospectrum using CMB polarization, which is ultimately a more sensitive measure of lensing for low-noise maps, have also been obtained (Story+2013, Polarbear 2014).

There has been rapid improvement in these measurements over the period of just a few years. Early detections of the CMB lensing autospectrum were not sample variance limited over a broad range in L and were only covering a relatively small sky area; the power spectrum of the noise in the CMB lensing reconstruction in the 2015 Planck data release is approximately equal to the lensing power spectrum only at its peak of $L \sim 40$, but smaller scales are noise-dominated and therefore not limited only by sample variance. Lensing reconstructions from current ground-based surveys (like SPTPol, ACTPol, PolarBear) are strongly signal-dominated below $L \sim 200$ and noise-dominated on smaller scales. However, they have been obtained over relatively small sky areas of several hundreds of degrees. A ground-based survey such as CMB-S4, with wide

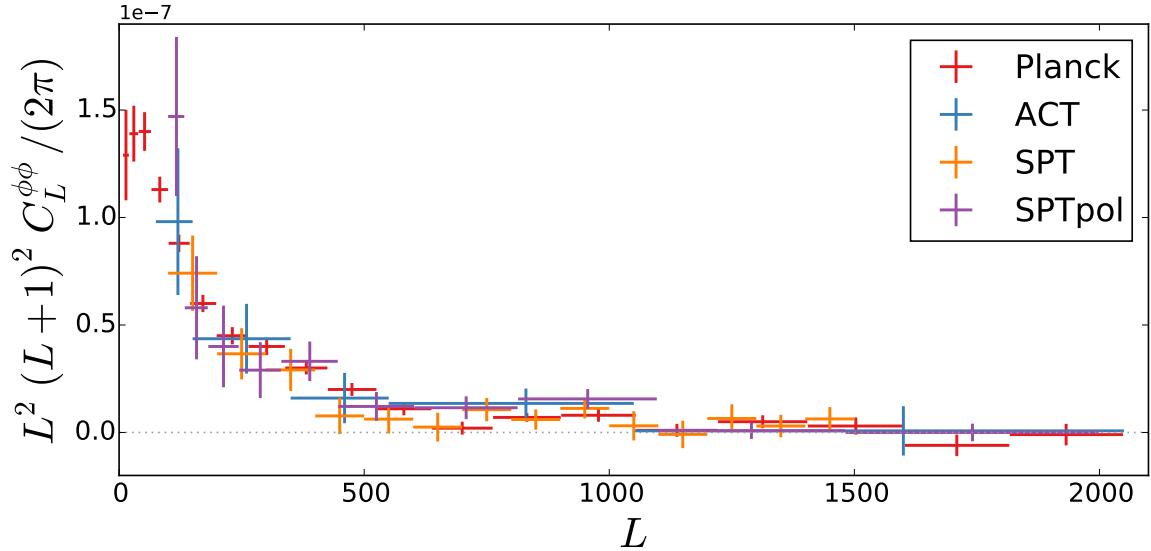


Figure 41. Compendium of lensing power spectrum measurements, since first discovery in 2011.

sky coverage, low-noise, and high resolution, will provide a sample-variance-limited measurement to scales below $L \sim 1000$ (see Figure 42) over a wide area.

Such a measurement holds the promise to qualitatively improve our understanding of cosmology. While the cosmological parameters describing the standard Lambda-Cold Dark Matter model have been precisely measured, extensions to this model can be constrained by including growth or geometrical information at a new redshift. From the redshifts probed by CMB lensing, extensions to the standard model such as a non-minimal mass for the sum of the neutrinos, a dark energy equation of state deviating from the vacuum expectation, and a non-zero curvature of the Universe can all be probed to much higher precision than with the primordial CMB alone.

6.3 Cross Correlations with CMB Lensing

Cross-correlating CMB lensing maps with other probes of large-scale structure provides a powerful source of information inaccessible to either measurement alone. Because the CMB last-scattering surface is extremely distant, the CMB lensing potential includes contributions from a wide range of intervening distances extending to high redshift; as a result, many other cosmic observables trace some of the same LSS that lenses the CMB. These cross-correlations can yield high-significance detections, are generally less prone to systematic effects, and given the generally lower redshift distribution of other tracers, are probing LSS in exactly the redshift range relevant for dark energy studies (see Figure 43). With CMB-S4, cross-correlations will transition from detections to powerful cosmological probes.

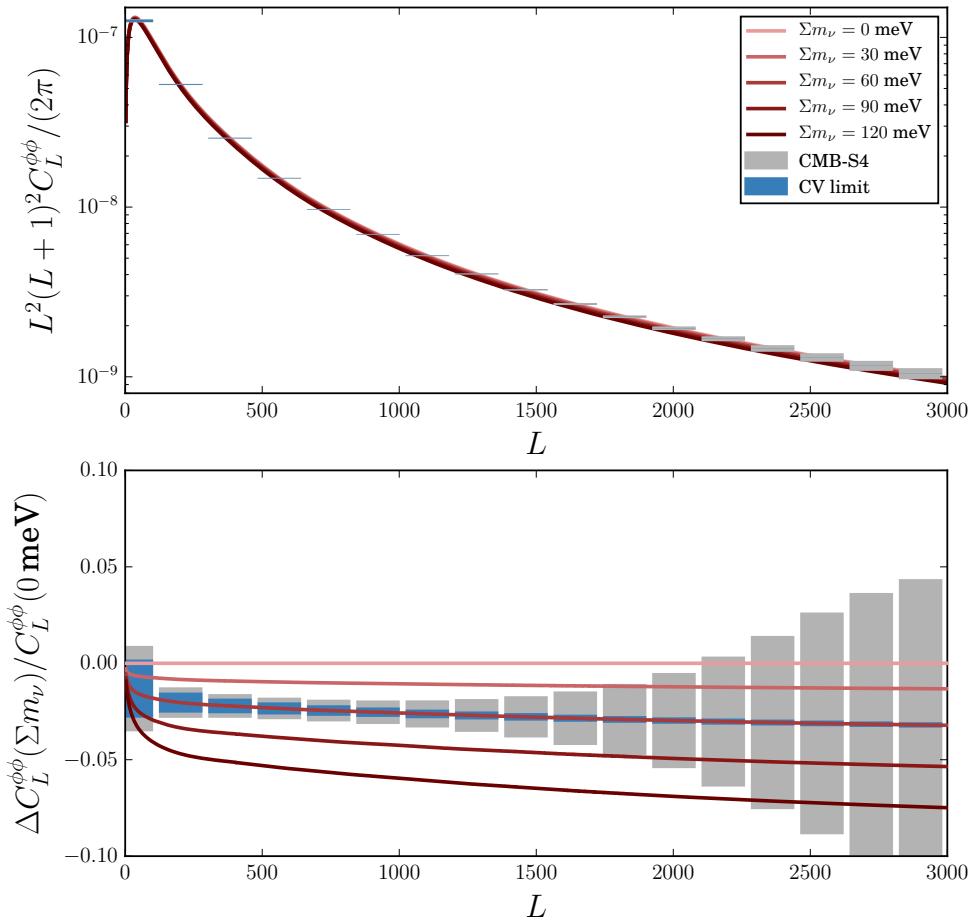


Figure 42. Constraining neutrino mass with CMB-S4. Top: lensing power spectra for multiple neutrino masses (curves) together with forecasted errors for S4. Bottom: residual from curve at zero neutrino mass. Error boxes are shown centered at the minimal value of 60 meV. S4 will be targeted to resolve differences in neutrino mass of 20 meV.

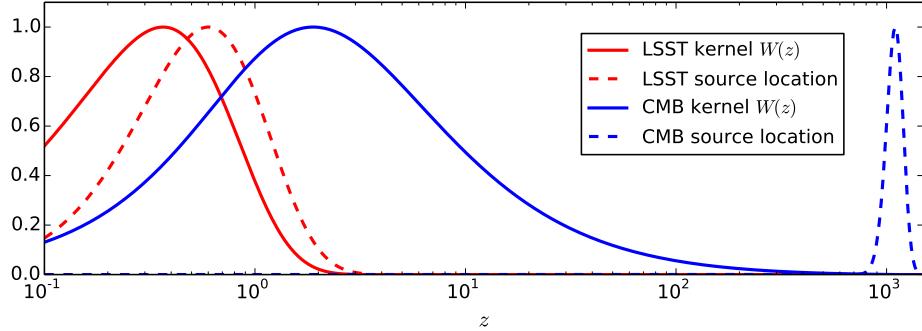


Figure 43. Redshift kernel for CMB lensing (blue solid) and for cosmic shear with LSST (red solid), together with the expected redshift distribution of LSST galaxies (red dashed) and the CMB source redshift (blue dashed).

6.3.1 CMB Lensing Cross Galaxy Density

Galaxies form in the peaks of the cosmic density field; thus the distribution of galaxies traces the underlying dark matter structure – which contributes to the CMB lensing potential. Cross-correlating galaxy density distributions with CMB lensing is highly complementary to galaxy clustering measurements. Galaxy surveys measure luminous matter while CMB lensing maps directly probe the underlying dark matter structure. Thus these correlations provide a clean measurement of the relation between luminous matter and dark matter. Cross-correlations between independent surveys are more robust against details of selection functions or spatially inhomogeneous noise that could add spurious power to auto-correlations. Additionally, while CMB lensing maps are projected along the line-of-sight, galaxy redshift surveys provide information about the line-of-sight distance; thus correlating redshift slices of galaxy populations allow for tomographic analysis of the CMB lensing signal (see, e.g., SPT/DES 2015). These properties can lead to improved constraints on cosmology: for example, with LSST galaxies, it has been shown that including cross-correlation with CMB lensing can substantially improve constraints on neutrino masses (Pearson & Zahn 2013).

CMB lensing was first detected using such a cross-correlation (Smith+ 2007, Hirata+ 2008). Since these first detections, cross-correlation analyses have been performed with tracers at many wavelengths, including optically-selected sources (Bleem+ 2012, Sherwin+ 2012, Planck 2013 XVII, SPT/DES 2015, Pullen+ 2014), infrared-selected sources (Bleem+ 2012, Geach+ 2013, DiPompeo 2015), X-ray-selected galaxy clusters (Planck 2013 XVIII), sub-mm-selected galaxies (Bianchini+ 2014, 2015) and maps of flux from unresolved dusty star-forming galaxies (Holder+ 2013, Hanson+ 2013, Planck 2013 XVIII, van Engelen+ 2015).

These cross-correlations between CMB lensing and galaxy clustering have already been used to test key predictions of general relativity, such as the growth of structure (SPT/DES 2015) as a function of cosmic time, and the relation between curvature fluctuations and velocity perturbations (Pullen+ 2015). Cross-correlations using CMB-S4 lensing data will enable percent level tests of general relativity on cosmological scales.

On the timescale of the S4 experiment, a number of large surveys are expected to be complete, including DESI, Euclid, and LSST. Due to the high number density of objects detected, wide areal coverage, and accurate redshifts, the precision of cross-correlation measurements with these surveys will be much higher than those performed to date. For example, the amplitude of cross-correlation between the S4 convergence map and the galaxy distribution from LSST is expected to be measured to XXX%.

6.3.2 CMB Lensing Cross Galaxy Shear

There have been several recent detections of the cross-correlation between lensing of the CMB and cosmic shear (Hand+ 2015, Liu+ 2015, Kirk+2016), demonstrating the emergence of a new cosmological tool.

Cosmic shear tomography is the measurement of cosmic shear of distant galaxies as a function of redshift, providing the ability to reconstruct the 3D mass distribution. CMB lensing offers similar signal-to-noise as cosmic shear surveys but provides the most distant source possible, allowing this 3D reconstruction to extend to the edge of the observable universe and providing a high-redshift anchor for dark energy studies.

CMB lensing, as a probe that is highly complementary to galaxy shear, can also be used as an external calibration for cosmic shear studies. It has been shown (Vallinotto 2012, Vallinotto 2013, Das+ 2013) that CMB lensing, galaxy clustering, and cosmic shear taken together can in principle cross-calibrate each other while still providing precise constraints on cosmological parameters. This has been successfully applied to existing surveys (Liu+ 2015, Baxter+ 2016) as a proof of principle; CMB-S4 will be a powerful tool for precision cosmology.

6.3.3 CMB Halo Lensing

In addition to constructing CMB lens maps of matter fluctuations on relatively large scales ($>\sim 5$ arcmin) as discussed in the preceding sections, one can also make CMB lens maps capturing arcminute-scale matter distributions. Such small-scale measurements capture lensing of the CMB by individual dark matter halos, as opposed to lensing by larger scale structure represented by the clustering of halos. This small-scale lensing signature, called CMB halo lensing, allows one to obtain measurements of the mass of these halos.

Using CMB halo lensing, CMB-S4 will be sensitive to halo masses in the range of $10^{13} M_{\odot}$ to $10^{15} M_{\odot}$. This corresponds to halos belonging to galaxy groups and galaxy clusters. As discussed in Chapter 3, measuring the abundance of galaxy clusters as a function of mass and redshift provides a direct handle on the growth of matter perturbations and consequently, on the equation of state of Dark Energy. Galaxy clusters can be identified internally in CMB maps through their wavelength-dependent imprint caused by the thermal Sunyaev-Zeldovich (tSZ) effect. However, the scaling between the tSZ observable, which is sensitive to baryonic physics, and the cluster mass, which is dominated by dark matter is not precisely constrained. Calibration of this mass scaling and scatter is currently the dominant systematic for extracting Dark Energy constraints from cluster abundance measurements.

Weak lensing of galaxies behind the galaxy cluster is a promising method for mass calibration since it is directly sensitive to the total matter content of the cluster. Reconstructing the mass profiles of clusters using measurements of the shapes of distant galaxies in deep photometric surveys is an active research program; however, it is often limited by the poor accuracy of source redshifts and the availability of sufficient galaxies behind the cluster, especially for very high-redshift clusters. CMB halo lensing has an advantage here since the CMB is a source of light which is behind every cluster, has a well defined source redshift, and well understood statistical properties.

A general approach for obtaining the average mass of a sample of clusters using CMB halo lensing is to reconstruct the lensing deflection field using a variation of the standard quadratic estimator, stack the reconstructed lens maps at the positions of the clusters, and fit the resulting signal to a cluster profile (e.g NFW). A modified quadratic estimator is used to reconstruct small-scale lensing signals since the standard estimator tends to underestimate the signal from massive clusters (Hu, Dedeo, Vale 2007). This modified

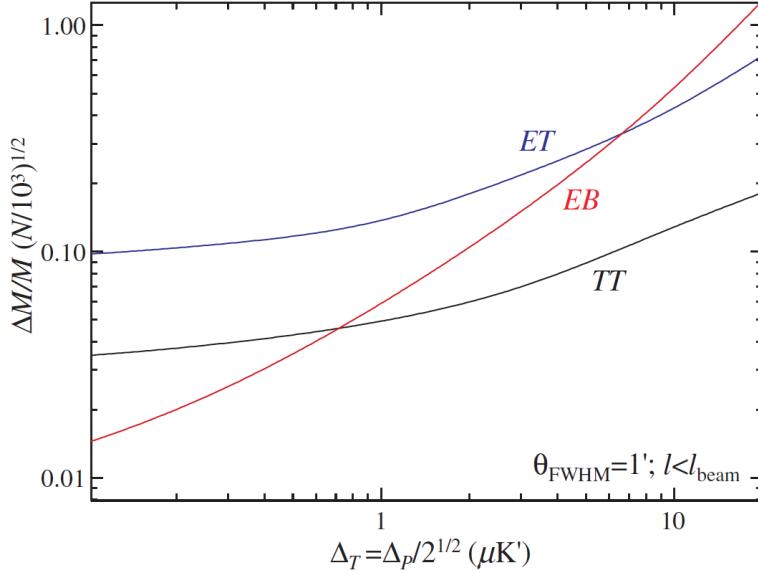


Figure 44. Lensing sensitivity to halos of a given mass M as a function of instrumental noise. Figure from Hu, DeDeo, Vale 2007.

estimator makes use of the fact that halo lensing induces a dipole pattern in the CMB that is aligned with the background gradient of the primordial CMB. The halo lensing signal can be measured with both temperature and polarization estimators, which can be used to cross check each other and reduce systematics.

CMB experiments have only very recently reached the sensitivity required to detect the lensing signal on scales of dark matter halos. The first detections were reported in 2015 by ACTPol (Madhavacheril+2015), SPT (Baxter+2015), and Planck (Planck 2015 XXIV). CMB-S4 will be capable of providing precision mass calibration for thousands of clusters which will be an independent cross check of galaxy shear mass estimates and will be indispensable for high-redshift clusters. Figure 44 shows that an arcminute resolution experiment with a sensitivity of around $1\mu\text{K}\text{-arcmin}$ can determine the mass of 1000 stacked clusters to $\sim 5\%$ using temperature maps and independently to $\sim 5\%$ using polarization maps. The primary systematic in temperature maps is contamination from the thermal SZ effect and radio and infrared galaxies coincident with the halos. This systematic can be mitigated using multi-frequency information due to the spectral dependence of the thermal SZ effect and galaxy emission, a procedure that requires the high sensitivity at multiple frequencies of CMB-S4. Halo lensing from polarization maps is relatively free of these systematics, and ultimately may be the cleanest way to measure halo masses, which requires the high polarization sensitivity provided by CMB-S4.

6.4 Delensing

To probe an inflationary gravity wave signal it is important to have low-noise B-mode polarization maps. However, for instrumental noise levels below $\Delta_P \simeq 5\mu\text{K}\text{-arcmin}$ in polarization, the dominant source of noise is no longer instrumental, but instead is from the generation of B-mode polarization by lensing of E-mode polarization from recombination (see Figure 45). This B-mode lensing signal has a well-understood amplitude, but the sample variance in these modes in the CMB maps leads to increased noise in estimates of

the inflationary B-modes. Unlike other sources of astrophysical B-mode fluctuations in the map, it cannot be removed with multifrequency data. Fortunately, this signal can be removed using map-level estimates of both the primordial E -mode map and the CMB lensing potential ϕ with a technique called delensing. However, this procedure requires precise maps of both the E-modes and of the gravitational lensing potential (which can be obtained from the CMB-S4 data itself).

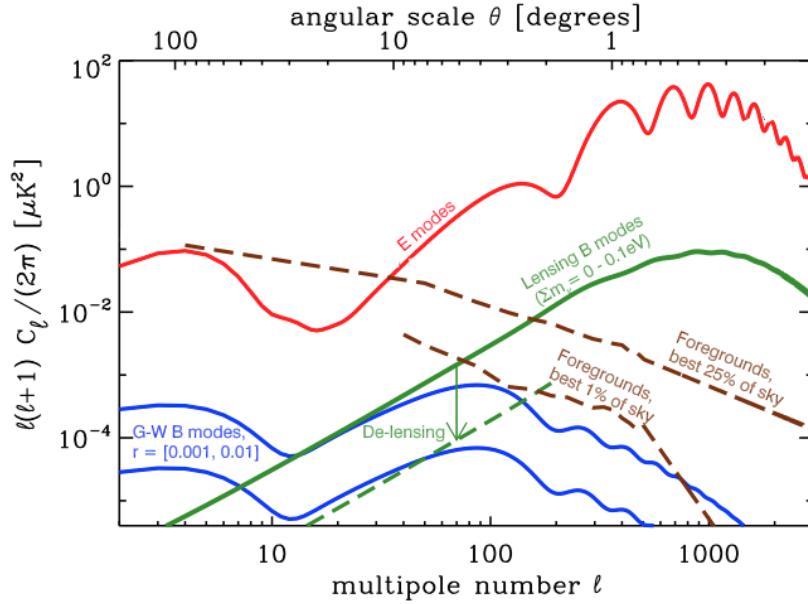


Figure 45. The green curve is the power spectrum of lens-induced E -to- B mixing. Delensing can reduce the amplitude of this effect by large factors (green dashed curve) yielding lower effective noise in B -mode maps.

Moreover, delensing will be a crucial part of the reconstruction of the CMB lensing field for CMB-S4, even for science goals like measuring the neutrino mass. This is because at low noise levels the standard quadratic reconstruction of lensing using the EB estimator (Hu & Okamoto 2001) can be improved upon by cleaning the B-mode CMB maps of the lens-induced B -mode fluctuations and then performing lens reconstruction again. This procedure can be repeated until CMB maps cleaned of the lensing signals are produced (See Figure 46).

Delensing in principle can be a perfect procedure: in the limit of no instrumental noise or primordial B modes, the lensing potential and the primordial E -mode map can be perfectly imaged (Hirata/Seljak 2003). However, the finite noise in a CMB-S4 survey will lead to residual lensing B modes which cannot be removed and will act as a noise floor for studying primordial B-modes from tensors. In particular, as shown in Figure 39, it is important to have high-angular resolution maps in order to obtain the small-scale E and B fluctuations needed for the EB quadratic lensing estimator.

Potential systematic biases with the delensing procedure are similar to those for measuring the lensing power spectrum. The impact of polarized dust and polarized synchrotron emission from the Galaxy as well as the impact of polarized extragalactic emission on the reconstructed lensing field are discussed in Section 6.5.1 as well as ways to mitigate them.

Additionally, rather than using an estimate of the CMB lensing field obtained internally from the CMB itself, it is also possible to use other tracers of large-scale structure which are correlated with CMB lensing

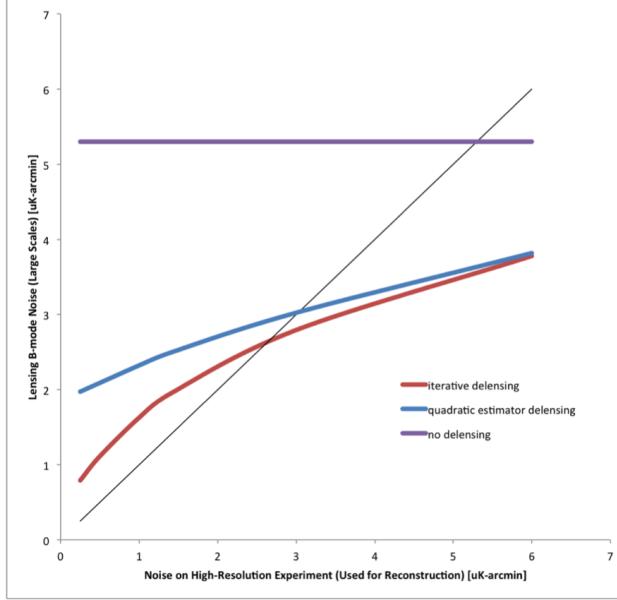


Figure 46. The B -mode noise on large scales as a function of the noise level used in EB -based lens reconstruction. The purple line is for no delensing and shows that lens-induced E to B mixing manifests as an effective 5 uK-arcmin white noise level. The blue curve shows the improvement possible when using a lens reconstruction to remove this source of effective noise. The red curve shows further improvement when the delensing is performed in an iterative fashion.

(Smith+ 2010). In particular the dusty, star-forming galaxies that comprise the cosmic infrared background (CIB) are strongly correlated with CMB lensing due to their redshift distribution which peaks near $z \sim 2$ (Sherwin+ 2014; Simard+ 2014). The level of correlation can be as high as 80% (Planck 2013 XVIII) and can in principle be improved using multifrequency maps of the CIB which select different emission redshifts (Sherwin+ 2014). However, as shown in (Smith+2010), the gain from delensing with external galaxy tracers is modest, and delensing internally with CMB maps holds far more promise.

6.5 Systematic Effects and Mitigation

The quadratic estimators used for lens reconstruction search for departures from statistical isotropy. The lens effect locally changes the CMB power spectrum via shear and dilation effects (e.g. Bucher+2011). Other sources of deviation from statistical isotropy can thus be confused with lensing effects; these can be of both instrumental and astrophysical origin.

6.5.1 Astrophysical Systematics

Extragalactic sources and tSZ clusters in temperature maps can be troublesome for lensing estimates in two ways: they tend to cluster more strongly in overdense regions (i.e., are non-Gaussian), an effect which lensing estimators can mistakenly attribute to lensing, while individual sources show up as strong local deviations from statistical isotropy.

Planck (2013) removed the effect of Poisson sources in the CMB four-point function from their lensing autospectrum which left untreated would have shifted their measured lensing power spectrum amplitude by 4%, a 1σ shift. For an experiment with lower map noise level and smaller beam, such as CMB-S4, sources can be found and removed to much fainter flux thresholds, making this a much smaller effect. The largest sources of bias thus come from the three-point and four-point correlation functions of the non-Gaussian clustering of sources and non-Gaussian clustering between the sources and the lensing field. These biases can be as large as several percent (van Engelen+2013, Osborne+2013) and their amplitude is highly model-dependent in temperature-based CMB lensing estimates. However, the extragalactic sources and tSZ clusters that can cause large sources of bias in temperature-based CMB lensing estimates are expected to be nearly unpolarized and therefore not a concern for polarization-based lensing estimates. In addition, sensitive multi-frequency temperature measurements should be able to spectrally remove these foregrounds through their unique frequency signatures. In addition, a robust campaign to measure these non-Gaussianities in the CMB data should allow a careful empirical understanding of these effects, an approach known as “bias-hardening” (Osborne+ 2013).

Observed levels of the polarization fraction of the diffuse Galactic emission at intermediate and high latitudes, reaching 10% or more, have been shown to impact non-negligibly on the quadratic estimators for achieving lensing extraction (Fantaye et al. 2012). This is due to leakage of the dominating long wavelength modes of the foreground signal onto the scales at which the lensing pattern is reconstructed. Therefore, as was the case for the Planck data analysis (Planck 2015), lensing extraction has to be validated on foreground cleaned maps output of a Component Separation process (see Section 6.4).

6.5.2 Instrumental and Modeling Systematics

Given the unprecedented precision targeted by CMB-S4 lensing measurements, the effects of instrumental systematic errors must be investigated and well-controlled. Since lensing results in a remapping or distortion of the sky, beam systematics are a particular concern.

The main beam systematics that affect CMB measurements are commonly described by differential gain, differential beamwidth, differential ellipticity, as well as differential pointing and rotation. In Smith et al. 2003, the impact of all these beam systematics on lensing measurements and hence on r and $\sum m_\nu$ was investigated using a Fisher matrix formalism. It was found that for an S4-type experiment, with $1\mu\text{K}$ -arcmin noise and a ~ 3 arcmin beam, the beam characterization from planets or other point sources will be sufficiently accurate that the biases arising from differential gain, differential beamwidth and differential ellipticity are less than one tenth of the one-sigma error on key parameters. Differential pointing and rotation must be controlled to within 0.02 arcmin and 0.02 degrees respectively in order to be similarly negligible.

While ideally the instrument can be designed or shown using measurements to have systematic errors that are negligible, one can also estimate residual beam systematics directly from the data, in a manner analogous to bias hardening. Many beam systematics result in a known mode-coupling (Yadav/Su/Zaldarriaga); their levels can hence be estimated by quadratic estimators and projected out, though complications due to the scan strategy must be accounted for. This method of beam-hardening was first demonstrated in (Planck 2013).

Another challenge in making high precision lensing power spectrum measurements is, given a set of cosmological parameters, predicting the actual observed lensing power spectrum. One example for such a challenge is the presence of higher-order biases, which have been previously neglected. For instance, often the Gaussianity of the lensing potential is assumed; however, if the full large scale structure non-linearity is taken into account, biases result that can affect measurements at the one-sigma level over many bandpowers.

In this respect, the modeling of ray tracing through N-body simulations (Calabrese et al. 2015) represents a resource for accounting for the complexity of this non-linearity.

In addition, the lensing power spectrum itself may not be exactly known, due to astrophysical or baryonic effects which modify the mass distribution. While this is a challenge for optical weak lensing measurements, investigations with simulations have found that such baryonic effects can be neglected for CMB lensing, at least at the precision achievable by CMB-S4 (cite paper incl. Battaglia).

6.6 Parameter Forecasts

Since CMB lensing is a sensitive probe of the matter power spectrum, CMB lensing measurements added to measurements of the primordial CMB power spectrum serve to significantly tighten parameter constraints. In particular, measurements of the CMB lensing autospectrum yield tight constraints on the sum of the masses of the neutrinos ($\sum m_\nu$). Cross correlations of CMB lensing maps with maps of galaxy density and galaxy shear can provide tight constraints on the dark energy equation of state (w) and modified gravity. Delensing B-mode polarization maps can also give strong constraints on the tensor-to-scalar ratio (r) from inflationary primordial gravity waves, as well as tighten constraints on the number of neutrino species (N_{eff}).

Below we forecast these parameter constraints including CMB lensing or delensing. Table 6-1 shows the expected error on the neutrino mass with just the primordial CMB alone and after including CMB lensing. We also show how Baryon Acoustic Oscillation (BAO) measurements from DESI tighten neutrino mass constraints further.

Table 6-1. Constraints on neutrino mass from CMB-S4 primordial CMB measurements plus CMB lensing and BAO measurements from DESI. Inputs to these forecasts are discussed in Section 6.6. These constraints are robust to modest variations in instrument resolution and sensitivity.

$\sigma(\sum m_\nu)$ meV	S4 Primordial	S4 Primordial +Lens	S4 Primordial +Lens+DESI
1' , 1 μ K'	324	55	18
3' , 1 μ K'	336	56	19
1' , 5 μ K'	378	60	20
3' , 5 μ K'	395	61	20

The assumptions input into these forecasts are that we assume no foregrounds and only white instrumental noise. Since most of the lensing signal-to-noise is coming from the EB lensing estimator (see Figure 40), white noise may be a reasonable assumption if leakage of atmospheric noise in temperature maps into polarization maps can be prevented via a half-wave plate or some alternative. We also assume CMB-S4 will observe 40% of the sky, and we include Planck primordial CMB data in the non-overlapping region of the sky ($65\% - 40\% = 25\%$ of sky). For both CMB-S4 and Planck, we use temperature modes between $l = 50 - 3000$, and polarization modes between $l = 50 - 5000$. We also include Planck low-ell data between modes $l = 2 - 50$. Lensing modes between $L = 40 - 3000$ were used in the calculations below. Given that there is covariance between the CMB power spectrum (2-point function) and the CMB lensing power spectrum (4-point function) because lensing induces peak-smearing in the former, we forecast Table I assuming unlensed CMB power spectra plus the CMB lensing 4-point function.

In Table 6-1, we see that 1-sigma errors on the sum of the neutrino masses of 60 meV is possible combining CMB-S4 lensing measurements with measurements of the primordial CMB from CMB-S4 and Planck. Table

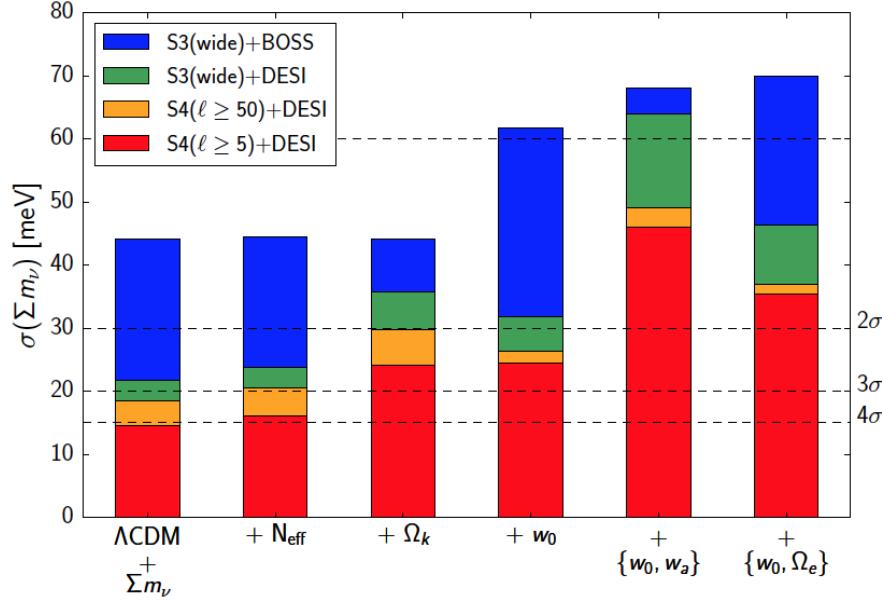


Figure 47. Constraints on total neutrino mass from various CMB surveys and with and without large-scale structure data included. In the minimal seven-parameter model (left bar) with all data included (red), the minimal neutrino mass of 60 meV can be detected at 4 σ ; the plot shows how this number degrades when including less data or when freeing up additional cosmological parameters. Figure from Allison et al 2015.

6-1 also shows the large improvement in neutrino mass constraint that CMB lensing measurements offer. When BAO data from DESI is added to this, neutrino mass errors of about 20 meV are achievable, which would yield a 4-sigma detection of the neutrino mass sum in the minimal mass scenario (also see Figure 47).

We also vary the resolution and sensitivity of CMB-S4 and explore the impact on neutrino mass constraints. We find that, given the assumptions above, the neutrino mass constraints are robust to modest variation in resolution (between 1 and 3 arcmin) and sensitivity (between 1uK-arcmin to 5uK-arcmin in temperature). However, we point out that a higher-resolution of 1 arcmin, as opposed to 3 arcmin, would make high-ell foreground removal more effective. High resolution on the scale of an arcminute is also critical for CMB halo lensing science as discussed in Section 6.3.3. Additionally, increased sensitivity gives more weight to the EB lensing estimator (see Figure 40), which is relatively free from foreground systematics and atmospheric noise.

Simulations and Data Analysis

(send any feedback on this chapter to tcrawfor@kicp.uchicago.edu and jdborrill@lbl.gov)

7.1 Introduction

Extracting science from a CMB dataset is a complex, iterative process requiring expertise in both physical and computational sciences. In this chapter we start with an overview of the data analysis pipeline before diving more deeply into its subsets - time-ordered data processing, component separation and the estimation of statistics and parameters. We then discuss the drivers for, and corresponding structure of, the simulation pipeline, and describe in detail its sky modeling and data simulation subsets. Finally we assemble the full simulation and data analysis pipeline, noting its inherently iterative structure, and discuss its critical uses in forecasting and validation and verification, before concluding with a discussion of some key implementation issues. Throughout our goal is to describe the current state of the art, note the particular challenges posed by CMB-S4, and describe how these challenges might be addressed.

7.2 Data Analysis Overview

The reduction of a CMB data set typically proceeds in a sequence of steps:

Pre-processing: The raw time-ordered detector data are calibrated and gross time-domain systematics are either removed (typically by template subtraction, filtering or marginalization) or flagged. The goal here is to make the real data match a model that will underpin all subsequent analyses.

Map-making: At each observing frequency, estimates of the intensity I and the Stokes Q- and U-polarizations of the sky signal are extracted from the cleaned time-ordered data based on their spatial stationarity, typically using some degree of knowledge of the instrument's noise properties.

Component separation: If a sufficient number of frequency maps are available, the CMB can be separated from the various foreground sky components based on its unique spectral invariance (in CMB units).

Power spectrum estimation: The six auto- and cross-angular power spectra of the CMB temperature T and E- and B-mode polarizations are estimated from the CMB and/or frequency maps, and corrected for E- to B-mode lensing.

Parameter estimation: The best-fit parameters for any cosmological model are derived by comparing the theoretical TT, TE, EE and BB CMB power spectra predicted by the model with the data.

This reduction essentially consists of a series of changes of basis for the data, from time samples to map pixels to spectral multipoles to cosmological parameters, with each basis-change reducing the data volume, increasing the signal-to-noise, and exposing a different class of systematic effects for mitigation.

Note however that the data can only remain a sufficient statistic at each step in the reduction if we also propagate its full covariance matrix. Since this is an $\mathcal{N}_b \times \mathcal{N}_b$ matrix in the dimension of the basis, its construction, manipulation and reduction pose the greatest computational challenge to this analysis. In particular the full pixel-domain data covariance matrix is generally dense and unstructured, requiring $O(\mathcal{N}_p^3)$ operations to build and $O(\mathcal{N}_p^2)$ bytes to store. All the major drivers of CMB science - polarization sensitivity, higher resolution, larger sky coverage - push us towards larger pixel counts, with an instrument mapping a fraction of the sky f_{sky} with a beam of b arcminutes covering $O(10^9 f_{\text{sky}}/b^2)$ pixels per IQU-component. For the last decade or more the computational intractability of the resulting pixel-domain matrices has forced us to replace explicit covariance propagation with Monte Carlo methods in all but a limited set of small sky fraction/low resolution cases.

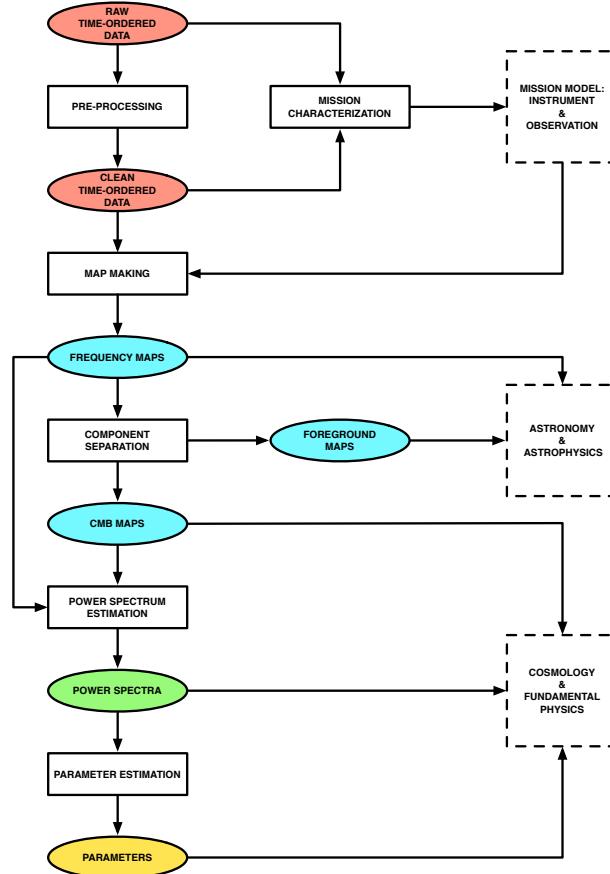


Figure 48. The CMB data analysis pipeline

Beyond this basic data reduction, the full analysis pipeline (Figure 48)¹ also includes mission characterization and science exploitation branches. Time domain data are extensively used to build a model of the mission, comprising the instrument and the observation. For the instrument this modeling can include such steps as determining beam profiles and estimating noise properties (including cross-correlations); for the observation, it includes reconstructing the detector pointing and polarization orientation from telescope sensor data, and incorporating atmosphere records in the data-flagging. The resulting mission model then feeds

¹In all pipeline figures ovals represent data objects in the various domain (red for time, blue for pixel, green for multipole and yellow for parameter) and rectangles represent data processing steps.

back into all of the ensuing data reduction and interpretation. The primary science exploitation derives cosmology and fundamental physics results from the various correlation functions of the CMB maps, from the power spectra's energy scale of inflation and neutrino mass to the higher-order statistics' measures of non-Gaussianity and the lensing potential. In addition the frequency maps represent important astronomical and astrophysical observations, particularly when the frequency sampling is sufficient to isolate individual foreground components along with the CMB.

7.3 Time-Ordered Data Processing

This section discusses the first part of a CMB data analysis pipeline, in which the raw time-ordered data from every detector is pre-processed then added into the estimate of temperature and polarization on the sky. These pre-processing and map-making steps, and the input and output data products associated with them, are shown in Figure 49. We discuss those steps—and some of the challenges we expect to face in implementing those steps in the CMB-S4 era—below.

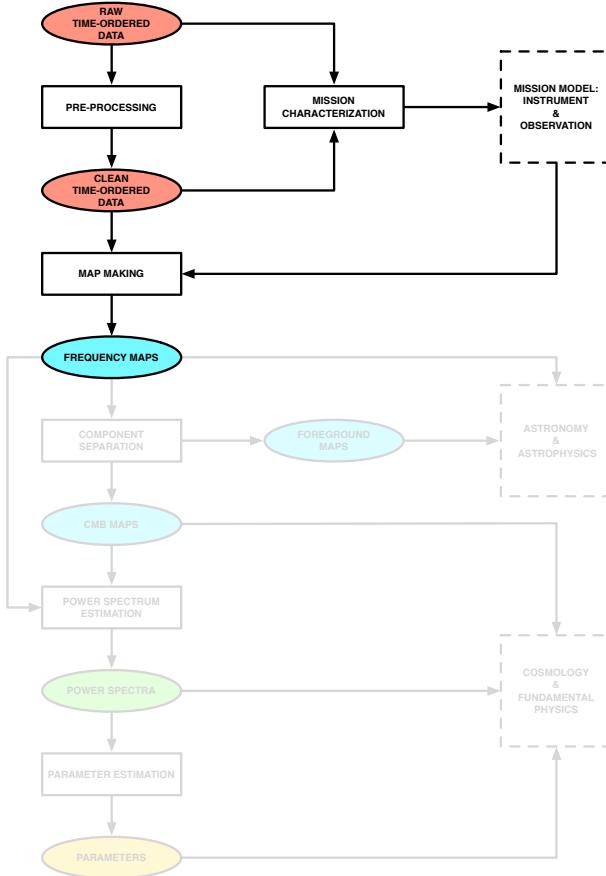


Figure 49. The time-ordered data processing subset of the CMB data analysis pipeline

7.3.1 Pre-Processing and Mission Characterization

The first stage of analysis in typical CMB experiments involves pre-processing of the raw time-ordered data in an attempt to clean the data of time-domain systematics and make the real data match a model that will underpin all subsequent analyses. Typical steps in this pre-processing are finding and removing cosmic-ray hits (“glitches”) on individual detectors and narrow-band filtering of spectral-line-like contamination to the time-ordered data (often from detector sensitivity to a mechanical apparatus such as the cryocooler). A challenge in the CMB-S4 era will be to properly account for these steps—which can involve the data from

the entire set of detectors over a long observing period—in the time-ordered data simulation pipeline, in order to characterize their effects on the final science results. The data volume is sufficiently large at this step that multiple full simulations may be unfeasible.

The pre-processing phase is also where many of the inputs to the mission model (a key piece of the data simulation pipeline) are measured. These inputs include pointing reconstruction, beam profiles, bandpasses, and noise, though in ground-based experiments some of these are more easily measured using dedicated observations of bright sources or off-sky characterization.

7.3.2 Map-Making

Map-making is the stage of the analysis when the major compression of the time-ordered data happens and some estimate of the sky signal is produced at each observing frequency. It is usually a linear operation, characterized by some operator, \mathbf{L} , which transforms the input time-ordered data, \mathbf{d} , into a pixel domain map, \mathbf{m} , e.g., [504],

$$\mathbf{m} = \mathbf{L}\mathbf{d}, \quad (7.1)$$

typically under the condition that the estimator is unbiased over the statistical ensemble of instrumental noise realizations, i.e.,

$$\langle \mathbf{m} - \mathbf{s} \rangle = 0, \quad (7.2)$$

where \mathbf{s} is the underlying pixelized sky signal. Given the usual model for the time-ordered data as the sum of sky-synchronous signal and time-varying noise,

$$\mathbf{d} = \mathbf{A}\mathbf{s} + \mathbf{n}, \quad (7.3)$$

for a pointing matrix \mathbf{A} , this condition leads to,

$$\langle \mathbf{m} - \mathbf{s} \rangle = (\mathbf{L}\mathbf{A} - \mathbf{1})\mathbf{s} + \langle \mathbf{n} \rangle = (\mathbf{L}\mathbf{A} - \mathbf{1})\mathbf{s}, \quad (7.4)$$

as the average noise is assumed to vanish. Hence,

$$\mathbf{L}\mathbf{A} = \mathbf{1}, \quad (7.5)$$

which is solved by,

$$\mathbf{L} = (\mathbf{A}^T \mathbf{W} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{W}. \quad (7.6)$$

Here the matrix \mathbf{W} is an arbitrary positive definite weight matrix, and different choices of \mathbf{W} lead to different estimates of the sky signal.

- If \mathbf{W} is taken to be the inverse of the time-domain noise covariance, i.e., $\mathbf{W} = \mathbf{N}^{-1}$, then the sky signal estimate, \mathbf{m} , will correspond to the **maximum likelihood** and **minimum variance** solution.
- If \mathbf{W} is taken to be proportional to some diagonal matrix minus some low-rank correction, i.e. $\mathbf{W} \propto \mathbf{1} - \mathbf{T}\mathbf{T}^T$, with \mathbf{T} assumed to be column-orthogonal, then the modes defined by its columns are marginalized over, effectively removing them from the solution. This approach includes as a special case so-called **desstriping** map-making, e.g., [505, 506], which has gained recognition thanks to its successful applications to the Planck data, e.g., [507, 508, 509, 510], and is therefore of potential interest to any experiments aiming to cover a large fraction of the sky. More generally, however \mathbf{T} can be constructed to remove any unwanted modes present in the time domain data, e.g., [511, 512, 513].

- If \mathbf{W} is taken to be diagonal, then the map-making solution corresponds to **binning**, i.e. the weighted co-addition of the samples falling within each pixel.

If the instrument beams display complex, non-axially symmetric structure, the proper estimation of the sky signal may require correcting for their effects at the map level, leading to the so-called **deconvolution** map-making [514, 515, 516]. However, further work is needed to demonstrate the effectiveness of such an approach in general.

If map-making is used primarily as a data compression operation on the way to deriving constraints on the statistical properties of the sky signal (such as its power spectra), one may choose to relax the condition in Eq. (7.2) in favor of the more computationally tractable, albeit potentially biased, sky estimate,

$$\mathbf{m} = (\mathbf{A}^T \text{diag}(\mathbf{W}) \mathbf{A})^{-1} \mathbf{A}^T \mathbf{W} \mathbf{d}, \quad (7.7)$$

where $\text{diag}(\mathbf{W})$ denotes the diagonal part of \mathbf{W} . In this approach any bias is then corrected at the next level of the data processing, e.g., [517]. This approach has been proven to be very effective, at least in the context of experiments with small sky coverage, e.g., [518, 519, 500, 501].

Formally the linearity of the mapmaking operation permits the propagation of the uncertainty due to the instrumental noise from time- to pixel-domain as

$$\hat{\mathbf{N}} = \mathbf{L} \mathbf{N} \mathbf{L}^T, \quad (7.8)$$

which leads to a particularly simple expression for maximum likelihood estimators

$$\hat{\mathbf{N}} = (\mathbf{A}^T \mathbf{N}^{-1} \mathbf{A})^{-1}. \quad (7.9)$$

However, as noted above, due its size the computational cost involved in computing such pixel-domain noise correlations make them impractical for all but special cases today, and the uncertainty is either carried over to the next stages of the data processing in implicit form or the final uncertainty is estimated using Monte Carlo simulations.

7.4 Component Separation

This section discusses the algorithms and methods for disentangling different sources of sky emission in multi-frequency maps. We first present the motivations and the general ideas of existing approaches. We then give some specifics of parametric and blind methods. Finally, we summarize several questions which might be answered by follow-up studies.

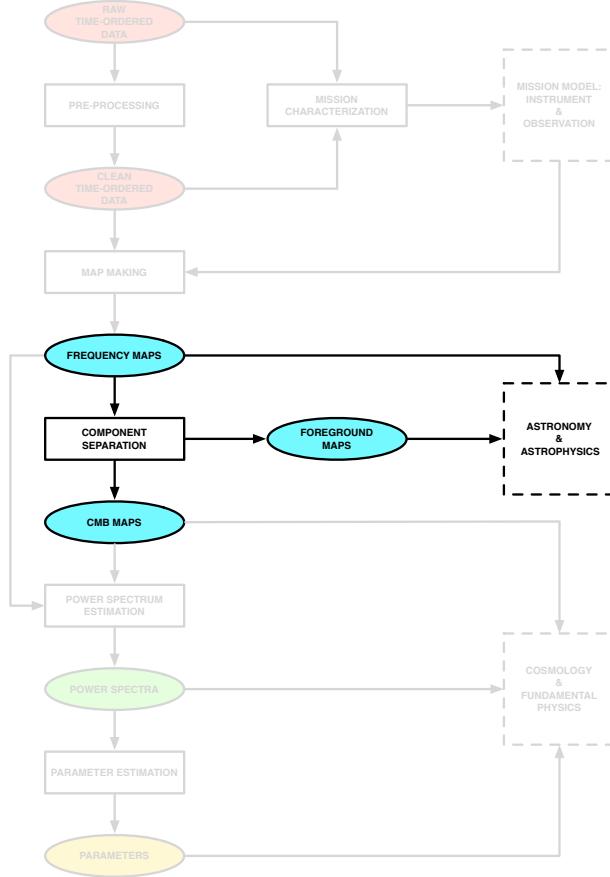


Figure 50. The component separation subset of the CMB data analysis pipeline

7.4.1 Introduction

7.4.1.1 Motivations

Recent measurements by BICEP2/Keck/Planck [50] confirm that on degree scales, where CMB-S4 is expected to search for the imprint of B modes from primordial GWs, the contamination from polarized foreground emission is comparable to or higher than the cosmological signal at 150 GHz even in one of the cleaner patches of the sky. Given that 150 GHz is expected to be close to the minimum of foreground contamination

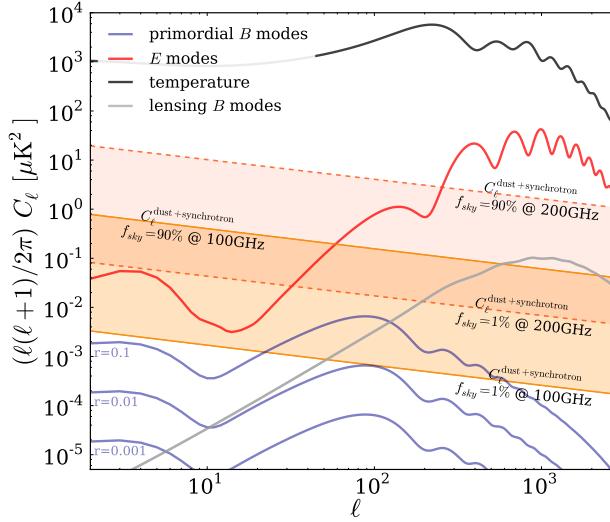


Figure 51. Angular power spectra showing primordial B modes, lensing B modes, total intensity, and E modes, as well as the total contribution of polarized B -mode foregrounds (dust plus synchrotron), expected on the cleanest 1 – 90% of the sky, at 100 and 200 GHz. Note that, as these results are derived from Planck data at intermediate and high Galactic latitudes, sensitive primarily to the large scale foreground pattern in polarization and are not therefore optimized for high-resolution, small scale instruments, there is potential for discovery of small patches of sky (e.g., $f_{\text{sky}} \leq 5\%$) with a signal differing than those indicated here. From [222].

vs. CMB signal, this is likely to be the case at all frequencies and all but the smallest fractions of the sky. Given the power law behavior in ℓ found on larger scales by Planck and WMAP [520, 521], foregrounds are expected to be even more relevant at larger angular scales. Foregrounds are expected to be subdominant with respect to the B -mode lensing signal on the scale of a few arcminutes (see Figure 51); nevertheless, dust polarization fractions around 10% (comparable to observed levels) have been shown to have non-negligible impact on the 4-point function used for achieving lensing extraction [522]. Therefore, component separation is a necessary and important step in gaining insight into the amplitude of primordial GWs, as well as the neutrino masses and dark energy abundances through CMB lensing studies.

7.4.1.2 Definition of component separation

Broadly defined, the process of component separation would generally

- include any data processing that characterizes and exploits correlations between observations at multiple frequencies
- use external constraints and physical modeling
- aim at distinguishing between different physical sources of emission.

The general data modeling reads

$$d_p = \sum_{\text{comp}, p} a_p^{\text{comp}} s_p^{\text{comp}} + n_p \equiv \mathbf{A} s_p + n_p \quad (7.10)$$

where the vector d_p contains the measured signal in each observing band, \mathbf{A} is the so-called mixing matrix which encapsulates the emission law a_p^{comp} of each component, s_p is a vector containing the unknown CMB and foregrounds amplitude and n_p is a vector containing the noise level at each observing band. The index p refers to sky pixels (θ, ϕ) , or modes of a spherical harmonic decomposition (ℓ, m) , or a set of Fourier modes (k_x, k_y) , etc. Note that this modeling assumes spatial templates s_p that are the same in all observing bands.

Component separation aims at inverting Eq. 7.10, to estimate the foregrounds-disentangled CMB signal encapsulated in s_p , as well as the foreground templates which are relevant for testing and updating our knowledge of astrophysical processes (and hence improve the sky model), as illustrated in Fig. 50. The estimate \tilde{s}_p of the true sky templates s_p — given \mathbf{A} , d_p and the statistical properties of the noise — minimizes the following χ^2 :

$$\chi^2 \equiv \sum_p |s_p - \tilde{s}_p|^2 \quad (7.11)$$

and can be taken to have the following general form

$$\tilde{s}_p = \mathbf{W} d_p \quad (7.12)$$

where the weighting operator \mathbf{W} is chosen to optimize some criterion regarding \tilde{s}_p and s_p (variance of the cleaned map, unbiasedness, etc.) while keeping statistical consistency and robustness. In particular, a common requirement for all component separation algorithms is the ability of propagating errors due to foreground subtraction, while having the flexibility of including foreground modeling and external constraints in a transparent way.

Component separation is then defined as a method of estimating the mixing matrix \mathbf{A} and finding the weighting \mathbf{W} that provides closest possible estimate \tilde{s}_p to the true sky signal.

For example, a solution to Eqs. 7.11 and 7.12 is obtained by taking $\mathbf{W} \equiv (\mathbf{A}^T \mathbf{N}^{-1} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{N}^{-1}$ with $\mathbf{N} \equiv \langle n_p^T n_p \rangle$, leading to an unbiased estimate of the sky. As mentioned below, this expression can be changed (see, e.g., [523]), depending on the desired level of generality and complexity and on the level of prior knowledge of the sky signal.

Studies have demonstrated the applicability of classes of component separation algorithms to certain simulated multi-frequency datasets, either balloon-borne or ground-based, and targeting limited frequency ranges and sky areas [524, 525, 522]. Results indicate that generally, for a frequency range extending from 90 to 250 GHz, polarized foregrounds may be removed effectively through a multi-frequency combination, at the price of enhancing the white noise contribution due to channel mixing; moreover, a possible bias may be introduced if, at the lowest frequency interval edge, the synchrotron component is not negligible: lower frequency templates/data are required to avoid such a contribution [12]. The most comprehensive application of component separation to data, in terms of completeness of algorithms and frequency range, is represented by Planck [520], although the targeted CMB components in that analysis (total intensity and E -mode polarization) are not the same as in the CMB-S4 case.

7.4.2 Description of methods

The CMB extraction may be achieved essentially through two basic concepts: the fitting of foreground unknowns along with CMB, or the minimization of the variance of a linear combination of the data, constrained to have the frequency scaling of a black body. The first class of algorithms, known as “parametric”, makes the maximum use of prior knowledge of foreground emission. By contrast, the second class, known as “blind”,

makes the minimum set of assumptions. These two broad classes and other possibilities are discussed in turn below.

- **Parametric** – The overall idea of these methods boils down to two steps: 1) the estimation of the mixing matrix, \mathbf{A} ; and 2) the inversion of Eq. 7.10 to recover an estimate of the sky signal, s_p . Parametric methods assumes that the mixing matrix, used in Eq. 7.10, has a functional form which is known and which can be parametrized by so-called "spectral" parameters β , i.e. $\mathbf{A} = \mathbf{A}(\beta)$. The functional form of \mathbf{A} being fixed, the estimation of the mixing matrix is therefore equivalent to an estimation of the parameters β . The parameters of the model are determined via a fitting procedure, often performed over sky pixels. This can be achieved by maximizing the following so-called "spectral" likelihood [526, 527]:

$$-2 \log \mathcal{L}(\beta) = - \sum_p (\mathbf{A}^T \mathbf{N}^{-1} d)^T (\mathbf{A}^T \mathbf{N}^{-1} \mathbf{A})^{-1} (\mathbf{A}^T \mathbf{N}^{-1} d). \quad (7.13)$$

Any deviation between the true mixing matrix \mathbf{A} and the estimated $\tilde{\mathbf{A}} \equiv A(\tilde{\beta})$ leads to the presence of foreground residuals in the reconstructed component maps.

- **Blind** – Under the assumption that sky components are statistically independent, blind methods aim to recover these components with an a priori unknown mixing matrix. Blind methods make minimal assumptions about the foregrounds and focus on reconstructing the CMB from its well known black body spectral energy distribution. The Internal Linear Combination (ILC, [528]) belongs to this class of methods. It only uses the CMB column of the mixing matrix elements (noted a hereafter) to perform the minimum variance reconstruction, cf. Eq. 7.12:

$$\tilde{s}_p = \sum_{i=0}^{i=m} w_i d_{p,i} \quad (7.14)$$

with $\sum_i w_i a_i = 1$, leading to the following solution:

$$w_i = a^T N^{-1} (a^T N^{-1} a)^{-1} \quad (7.15)$$

In this scheme, no attempt is made to design a foreground model. The decorrelation property between CMB and foregrounds alone is used to project out the contamination into a $m-1$ subspace (with m being the number of frequency maps).

The main caveat in this method is its well known bias ([529, 523], etc) which comes from empirical correlation between the CMB and the foregrounds. The ILC bias is proportional to the number of detectors m and inversely proportional to the number of pixels used to compute N . In order to reduce this effect, one could think of reducing the foreground subspace size by adding further constraints. The SEVEM template fitting method ([530], etc) follows this idea, by building some foreground templates with a combination of a subset of the input frequency maps.

The semi-blind SMICA method [531] also works at containing the foreground in a smaller dimension space, but in a more general way. The idea of Independent Component Analysis (ICA) is to blindly recover the full mixing matrix \mathbf{A} by using the independence property of the different components. As we know that they are spatial correlations between the foregrounds, the ICA principle is used to disentangle the CMB from the noise and the foregrounds taken as a whole.

The main advantage of such blind or semi-blind methods is their ability to face any unknown and/or complex foreground contamination, to reconstruct a clean CMB signal. This is a big advantage when

real data comes, one can then focus on instrumental effects, or data set combination issues at first, and leave the complex task of the foreground modeling and reconstruction for a future analysis step.

Moreover, in a framework like SMICA, the level of blindness can be adjusted via the plugin of any parametric component to its flexible engine as described in [531], allowing for a step by step fine grain design of the foreground model.

- **Template fitting** – In this variant, emission laws are not modeled, and the analysis is reduced to the maximisation of a likelihood over the CMB contribution and the amplitudes of each foreground component (see, e.g., [532]).

For all of the approaches discussed above, Eq. 7.12 can be implemented equivalently with any representations of the map—i.e. pixel, harmonic, wavelet, etc. The resulting component separation is independent of this choice as long as the linear data modeling (Eq. 7.10) holds. This complementarity, and the internal comparison of results through these pipelines has been proven to be relevant in actual analysis of Planck data [520]. That said, the difference between domain of application will lie in the computational needs: for high number of sky pixels, the implementation of Eq. 7.12 might be significantly more efficient in harmonic space.

7.4.3 Questions to be addressed during follow-up studies

- **E/B or Q/U basis of analysis** – Component separation between CMB radiation and its foregrounds can be performed either dealing with Stokes parameters Q and U maps of the sky in real space or Fourier space, and either before or after the separation between the E and B modes. Several approaches have been followed by CMB experiments so far [533, 534, 50], and each of them has some advantages and some caveats. For example, processing Q/U data in the map domain simplifies the treatment of foreground components that have non-Gaussian and/or non-stationary spatial distributions. However, in the Q and U basis, the CMB E and B modes are mixed and the CMB E modes will be the dominant contribution to the variance at intermediate and small scales in the CMB observing frequencies, limiting the accuracy of the separation. To overcome this limitation, E and B observables can be constructed in Fourier space. The separation of the B -mode components (primordial CMB, Galactic foregrounds, lensing, etc.) can then be done in the angular power spectrum domain (where the final accuracy might be limited by the cosmic variance associated to foregrounds), in the two-dimensional, phase-full Fourier domain (where the treatment of non-stationary components will be complicated) or in the map domain (where the final accuracy might be limited by ringing of the foregrounds due to the non-local transformation). Although these different approaches are currently giving satisfactory results on simulated data, these effects will become crucial at the sensitivity of CMB-S4 and merit a dedicated study.
- **Combining data from multiple instruments** – Ground-based instruments heavily filter time streams because of atmosphere contamination, ground emission, etc. In particular, large angular scales are usually suppressed anisotropically, and this suppression is corrected in the power spectrum estimate. Component separation using observations from different platforms will be made more straightforward if all maps are derived from common filters.
As stressed already, the first attempt at component separation or foreground cleaning for B modes on multi-platform data was recently implemented in [50], using data from BICEP2, Keck Array and Planck. A template fitting analysis was implemented with the primary objective of minimizing the

variance in the CMB solution. The simultaneous analysis of combined data sets required an additional layer in the analysis, namely the simulated scans of the Planck data through the filtering by the ground observatories, along with validation through simulations of the whole procedure. A simpler approach for CMB-S4 would be to have a single pipeline reducing and combining different datasets. With a common filtering implemented from scratch in a multi-site experiment, the combination would be built-in, thus avoiding the extra layer and increasing confidence and robustness of results.

- **Various resolutions** – Under the approximation that the mixing matrix does not significantly vary as a function of resolution, the impact of different beam sizes can be propagated to the noise level of the final CMB map by incorporating the beam for each frequency channel in the expression of the noise covariance matrix

$$\mathbf{N}(i) \equiv \mathbf{N}(i)_\ell = (\sigma_i)^2 \exp \left[\frac{\ell(\ell+1)\theta_{\text{FWHM}}^2}{8 \log(2)} \right] \quad (7.16)$$

where i is a frequency channel and σ_i is the noise level in the corresponding map. The noise variance in the reconstructed CMB map, i.e. after component separation, would then be given by

$$N_\ell^{\text{post comp sep}} = \left[\left(\mathbf{A}^T (\mathbf{N}_\ell)^{-1} \mathbf{A} \right)^{-1} \right]_{\text{CMBxCMB}} \quad (7.17)$$

- **Atmosphere residuals** – Atmosphere residuals appear at large scales in ground-based CMB observations, and they scale with frequency in a similar way as dust, $\propto \nu^\beta$ [535]. Having redundant frequencies among the different observatories could help mitigate the atmospheric and astrophysical foregrounds. Furthermore, the small intrinsic polarization of the atmosphere [536, 535] will limit the contamination to component separation of the polarized signals. Still, this effect will have to be investigated quantitatively with realistic simulations.

7.5 Statistics and Parameters

In this section, we discuss the process of going from sky maps at different frequencies—or, in light of the previous section, foreground-cleaned CMB maps and an estimate of foreground residuals—to post-map products such as angular power spectra, estimates of lensing potential ϕ , and finally cosmological parameters, as well as covariance estimates for all of these quantities. We briefly describe the current practice for this process, then we address specific challenges anticipated in the CMB-S4 era.

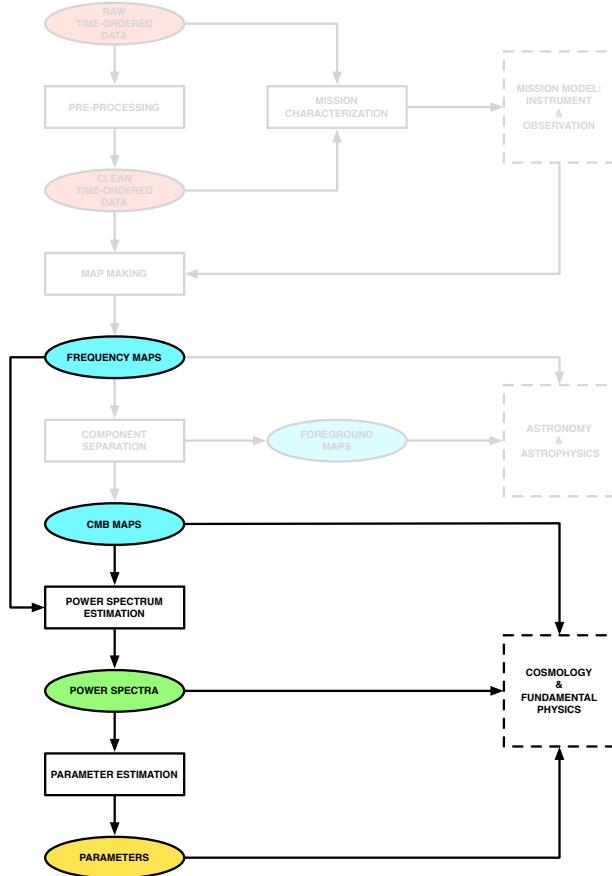


Figure 52. The statistics and parameters subset of the CMB data analysis pipeline

7.5.1 Current practice

Early measurements of CMB temperature anisotropy, with comparatively few map pixels or angular modes measured, often used maximum-likelihood methods to produce maps of the sky (e.g., [537]) and either a direct evaluation of the full likelihood or a quadratic approximation to that likelihood (e.g., [538]) to go from maps to angular power spectra (the first step in Figure 52). With the advent of the WMAP and Planck space missions, which would map the entire sky at sub-degree resolution, it became apparent that computing resources could not compete with the $\mathcal{O}(\mathcal{N}^3)$ scaling of the full-likelihood approach (e.g., [539]).

The solution for power spectrum analysis that has been adopted by most current CMB experiments is a Monte-Carlo-based approach advocated in [517]. In this approach, a biased estimate of the angular power spectrum of the data is obtained by simply binning and averaging the square of the spherical harmonic transform of the sky map. That estimate (known as the “pseudo- C_ℓ spectrum”) is related to the unbiased estimate that would be obtained in a maximum-likelihood procedure through the combined effect of noise bias, sky windowing, and any filtering applied to the data before or after mapmaking (including the effects of instrument beam and pixelization). These effects are estimated by “observing” and analyzing simulated data and constructing a matrix describing their net influence on simulated data. This matrix is inverted, and the inverse matrix is applied to the pseudo- C_ℓ s to produce the final data product. Some version of this Monte-Carlo treatment is likely to be adopted for CMB-S4.

Pseudo- C_ℓ methods are also now commonly used in analysis of CMB polarization anisotropy [534, 540, 541]. An added complication in polarization analyses is that pseudo- C_ℓ methods do not cleanly separate E and B modes (e.g., [542]). “Pure” B -mode estimators can be constructed that suppress the spurious B -mode contribution from estimating E and B on a cut sky with pseudo- C_ℓ methods [543]), but other analysis steps (such as particular choices of filtering) can produce spurious B modes that are immune to the pure estimators [11]. These can also be dealt with using Monte-Carlo methods, either by estimating the statistical bias to the final B -mode spectrum or by constructing a matrix representing the effect of any analysis steps on the true sky [501]. The latter approach involves constructing an $\mathcal{N}_{\text{pixel}}$ -by- $\mathcal{N}_{\text{pixel}}$ matrix, equal in size to the full pixel-pixel covariance, and will not be feasible for high-resolution CMB-S4 data but could be used in analyzing lower-resolution data.

In addition to the two-point function of CMB maps, higher-order statistics of the maps have recently been of great interest to the community. In particular, the four-point function encodes the effect of gravitational lensing, and estimators can be constructed to go from CMB temperature and polarization maps to estimates of CMB lensing ϕ and the associated covariance (e.g., [215, 544]). These quadratic estimators are the first step in an iterative estimation of the true likelihood, and in the weak-lensing limit they are nearly optimal; as a result, they remain the state of the art for estimating the large-scale ϕ from CMB lensing (e.g., [545]). For CMB-S4 sensitivity levels, it is possible that further gains can be made with more iterations (see Section 6.4). Even with multiple steps, the computational burden involved in this step of the analysis is unlikely to be significantly greater for CMB-S4 than for Planck.

The final step in the analysis of a CMB data set is the estimation of cosmological parameters from the various post-map statistics discussed above (the second step in Figure 52). This involves estimating the likelihood of the data given a model parameterized by the standard six Λ CDM parameters, possible extensions of the cosmological model, and any nuisance parameters involving the instrument, foregrounds, and other sources of systematic uncertainty. The current industry standard for this part of the analysis are Monte-Carlo Markov-Chain (MCMC) methods, in particular the implementation in CosmoMC [546], and it is expected that CMB-S4 will use similar methods. There will be several aspects of the CMB-S4 dataset, however, that will necessitate going beyond what past analyses have done at this step. First of all, the data from several different telescopes and cameras will need to be combined in as lossless a fashion as possible—such that combining at the parameters stage may be sub-optimal. Further, as shown by [50], foregrounds cannot be ignored in the estimation of the B -mode power spectrum, even in the cleanest parts of the sky and in the least contaminated observing bands. Foreground modeling will be used to mitigate the contamination, but there will be foreground residuals (both from noise and imperfect modeling), and these need to be properly characterized and accounted for in parameter extraction. Similarly, algorithms to separate the contributions to the B -mode power spectrum from a background of gravitational waves and from lensing of E modes (so-called “de-lensing”, see the Section 6.4 for details) will leave an uncertain level of lensing residuals in the primordial B -mode spectrum, and this residual will need to be treated properly. Finally, for information

from angular power spectra and lensing potential ϕ to be properly combined, the covariance between the two-point and four-point functions of the CMB needs to be taken into account.

7.5.2 Challenges

As discussed in the previous section, some of the avenues in the analysis that need to be re-addressed for an experiment such as CMB-S4 are:

- The combination of data from different telescopes and cameras (with different heritage and observation/analysis techniques) without significant loss of information.
- The impact of uncertainties in foreground modeling on cosmological parameters, particularly the tensor-to-scalar ratio r .
- The covariance between different observables (for example the lensed CMB power spectrum and the reconstructed lensing potential power spectrum).
- The impact of delensing—the separation of the gravitational lensing signal and the primordial B -mode signal, lowering the effective lensing background—and lensing residuals on cosmological parameters.

We treat each of these challenges individually in the sections below.

7.5.2.1 Combining different data sets

At what stage in the analysis does it make the most sense to combine data from different experimental platforms? One possibility is to estimate angular power spectra or even cosmological parameters from every data set individually and combine them at that stage. This would be computationally efficient but sub-optimal from a sensitivity standpoint unless every experiment covered fully independent patches of the sky. For any overlap between data sets, combining at the map or time-ordered data stage (adding before squaring) will lead to lower final uncertainties than combining at the power spectrum stage (squaring before adding). Of course, the earlier in the analysis we choose to combine data, the more work it will be to standardize the data between experimental platforms—the time-ordered data is generally quite instrument-specific, the maps less so, etc. The trade-off between maximizing constraining power and possibly placing undue burdens on the individual pipelines will need to be balanced in answering this question.

7.5.2.2 Foreground-related uncertainty on cosmological parameters

To separate the CMB signal from the contaminating signals of Galactic and extragalactic foregrounds, data from multiple bands will be combined, either in a cross-spectrum analysis or, as detailed in Section 7.4, by making linear combinations of maps in different bands to produce a “pure-CMB” map for power spectrum estimation. In either case, an underlying model of foreground behavior is assumed—even if that model is simply an assumption regarding the level to which the spectral behavior of foregrounds varies over the sky. Any model of foreground behavior is by definition imperfect, and the resulting component separation or frequency-cross-spectrum fit will have leakage between the foreground and CMB components. At the sensitivity levels attainable by CMB-S4, these residuals have the potential to dominate the error budget on

cosmological parameters and, more troublingly, to significantly bias the best-fit parameter values if they are not properly taken into account.

Section 7.7 discusses the baseline plan for, and challenges involved in, modeling Galactic and extragalactic foregrounds. It is likely that more information will be needed—from Stage-3 experiments, or from a possible dedicated, balloon-borne CMB foreground mission—before we can confidently assess the level to which foregrounds will limit the final parameter constraints from CMB-S4 and how flexible we will need to make the underlying foreground models that inform component separation and parameter extraction.

7.5.2.3 CMB lensing covariances for CMB S4

The measured lensing power spectrum is given by a 4-point function of the lensed CMB. This is not statistically independent from the lensed CMB 2-point function, because both depend on the same observed, lensed CMB maps. As a consequence, measured lensing power spectra and lensed CMB power spectra may be correlated. This correlation should be taken into account when combining these measurements to avoid spurious double counting of information. For the specific case of Planck this correlation is negligible [547]. However, the level of correlation depends on experiment specifications and the multipole range where power spectra have high signal-to-noise. The correlation should thus be included in analyses that combine 2- and 4-point measurements unless it is known to be negligible for a specific experiment.

For CMB-S4, the best lensing measurements are expected to come from the auto-power spectrum of EB reconstruction. Its covariance with the lensed EE and BB power spectra depends on six-point functions of the lensed CMB, e.g. $\langle EBEBEE \rangle$. Although many terms contribute, the dominant effect is expected from only a few contributions [547]:

- First, there are signal contributions to the covariance of the form

$$\text{cov}(\hat{C}_{l,\text{expt}}^{EE}, \hat{C}_L^{\phi_{EB}\phi_{EB}})_{\text{signal}} = \frac{\partial C_l^{EE}}{\partial C_L^{\phi\phi}} \frac{2}{2L+1} (C_L^{\phi\phi})^2, \quad (7.18)$$

$$\text{cov}(\hat{C}_{l,\text{expt}}^{BB}, \hat{C}_L^{\phi_{EB}\phi_{EB}})_{\text{signal}} = \frac{\partial C_l^{BB}}{\partial C_L^{\phi\phi}} \frac{2}{2L+1} (C_L^{\phi\phi})^2, \quad (7.19)$$

where \hat{C} are data power spectra, and C_{expt} are power spectra of observed (noisy, beam-deconvolved) CMB fluctuations $X \in \{E, B\}$:

$$\langle \hat{C}_{l,\text{expt}}^{XX} \rangle = C_{l,\text{expt}}^{XX} = C_l^{XX} + \left(\frac{\sigma_X}{T_{\text{CMB}}} \right)^2 e^{l(l+1)\sigma_{\text{FWHM}}^2/(8 \ln 2)}. \quad (7.20)$$

The signal covariance in Eqs. (7.18) and (7.19) arises because cosmic variance fluctuations of the true lensing potential (i.e. fluctuations of matter along the line of sight) modify the lensing reconstruction power as well as the lensed EE and BB power spectra. Formally, this follows from the connected part of the lensed CMB 6-point function.

- Second, a noise covariance follows from the disconnected 6-point function,

$$\text{cov}(\hat{C}_{l,\text{expt}}^{EE}, \hat{C}_L^{\phi_{EB}\phi_{EB}})_{\text{noise}} = \frac{2}{2l+1} (C_{l,\text{expt}}^{EE})^2 \frac{\partial(2\hat{N}_L^{(0)})}{\partial \hat{C}_{l,\text{expt}}^{EE}}, \quad (7.21)$$

and similarly for BB . This noise covariance arises because fluctuations of the CMB and instrumental noise change both the Gaussian reconstruction noise $N^{(0)}$ and the CMB power spectra. It is however

cancelled if the Gaussian $N^{(0)}$ reconstruction noise is subtracted in a realization-dependent way [548, 549, 550, 547]

$$\hat{C}_L^{\hat{\phi}\hat{\phi}} \rightarrow \hat{C}_L^{\hat{\phi}\hat{\phi}} - 2\hat{N}_L^{(0)} + N_L^{(0)}. \quad (7.22)$$

For the specific case of *EBEB* reconstruction, the realization-dependent $\hat{N}^{(0)}$ is

$$\hat{N}_L^{(0)} = \frac{|A_L^{EB}|^2}{2L+1} \sum_{l_1, l_2} |g_{l_1 l_2}^{EB}(L)|^2 \frac{1}{2} \left[\hat{C}_{l_1, \text{expt}}^{EE} C_{l_2, \text{expt}}^{BB} + C_{l_1, \text{expt}}^{EE} \hat{C}_{l_2, \text{expt}}^{BB} \right], \quad (7.23)$$

where A^{EB} is the estimator normalization and g^{EB} is the optimal weight given by [544]

$$g_{l_1 l_2}^{EB}(L) = -i \frac{C_{l_1}^{EE} {}_2F_{l_2 L l_1} - C_{l_2}^{BB} {}_2F_{l_1 L l_2}}{C_{l_1, \text{expt}}^{EE} C_{l_2, \text{expt}}^{BB}} \quad (7.24)$$

with lensed spectra in numerator and denominator [551, 549], and

$$\pm s F_{l_1 L l_2} = [-l_1(l_1+1) + L(L+1) + l_2(l_2+1)] \sqrt{\frac{(2l_1+1)(2L+1)(2l_2+1)}{16\pi}} \begin{pmatrix} l_1 & L & l_2 \\ \pm s & 0 & \mp s \end{pmatrix}. \quad (7.25)$$

In the square brackets in Eq. (7.23) one of the CMB power spectra is replaced by a data power spectrum. On average, $\langle \hat{N}_L^{(0)} \rangle = N_L^{(0)}$. This realization-dependent $\hat{N}^{(0)}$ subtraction also follows more formally from optimal trispectrum estimation (see Appendix B in [547] and Appendix D in [552]).

- A third covariance contribution arises from the connected trispectrum part of the CMB 6-point function. If realization-dependent $\hat{N}^{(0)}$ subtraction is used, the dominant remaining term is expected to be (at leading order in ϕ , see also Eq. (D4) of [547]; similarly for *BB*)

$$\text{cov}(\hat{C}_{l, \text{expt}}^{EE}, \hat{C}_L^{\hat{\phi}^{EB} \hat{\phi}^{EB}}) = 2 \frac{C_L^{\phi\phi}}{A_L^{EB}} \frac{\partial(2\hat{N}_L^{(0)})}{\partial \hat{C}_{l, \text{expt}}^{EE}} \frac{2}{2l+1} (C_{l, \text{expt}}^{EE})^2. \quad (7.26)$$

Similarly to avoiding the noise covariance with the realization-dependent $\hat{N}^{(0)}$ subtraction, the signal covariance could in principle also be avoided by delensing CMB power spectra with the estimated lensing reconstruction, e.g. by forming [547]

$$\hat{C}_{l, \text{expt}}^{EE} \rightarrow \hat{C}_{l, \text{expt}}^{EE} - \sum_L \frac{\partial C_l^{EE}}{\partial C_L^{\phi\phi}} \left(\frac{C_L^{\phi\phi}}{\langle \hat{C}_L^{\phi\phi} \rangle} \right)^2 (\hat{C}_L^{\phi\phi} - 2\hat{N}_L^{(0)}), \quad (7.27)$$

or by applying more advanced delensing methods. However this has not yet been tested in practice and only makes sense if lensing reconstructions have sufficient signal-to-noise. In general, forming linear combinations of measured lensing and CMB power spectra as in Eqs. (7.22) and (7.27) does not add any new information, but does simplify covariances.

Since more covariance contributions arise from other couplings of the CMB 6-point function, it should be tested against simulations if the above contributions are sufficient. In practice, it is then favorable to use analytical covariances because they are less noisy than those derived from simulations.

On top of the cross-covariance between 2-point CMB power spectra and 4-point lensing power spectra, both power spectra can also have non-trivial auto-covariances. Covariances between CMB power spectra have been computed in [553, 554, 555]. They contain similar building blocks as the covariances above [555]. Covariances between two 4-point lensing power spectra involve the lensed CMB 8-point function. While many covariance contributions are cancelled when using realization-dependent $\hat{N}^{(0)}$ subtraction [549], other contributions may be relevant for future experiments. Finally, the discussion above applies to the standard quadratic lensing reconstruction estimators and may change for maximum-likelihood lensing estimators [223].

7.5.2.4 Delensing

For noise levels below $\Delta_P \simeq 5\mu\text{K-arcmin}$, the dominant source of effective noise in B -mode maps is the fluctuation induced by the lensing of E -modes from recombination. This signal has a well-understood amplitude, and unlike many other sources of astrophysical fluctuation in the map, it cannot be removed with multifrequency data. Instead it must be removed either using map-level estimates of both the primordial E -mode maps and the CMB lensing potential ϕ or using a prediction for the lensing B -mode spectrum. The latter approach necessarily leaves some cosmic variance residual of the lensing signal after cleaning, while the former can in principle result in nearly perfect cleaning, so we will concentrate on that approach here.

Even in the map-level approach, the finite noise in the CMB-Stage IV survey will lead to residual lensing B -modes which cannot be removed and will act as a noise floor for studying B modes from tensors. The amplitude of these residual lensed B -modes are discussed in Section 6.4 as a function of the angular resolution and the noise level of the S4 survey; in particular, it is crucial to have high-angular resolution maps in order to measure the small-scale E - and B -modes fluctuations needed for the EB quadratic lensing estimator.

The concerns with the delensing procedure are similar to those for measuring the lensing power spectrum. The impact of polarized dust and synchrotron emission from the Galaxy, and the impact of polarized point sources on small scales on the lensing reconstruction are addressed in chapter IV. Left untreated the effects may be large; however the use of multi-frequency data together with the application of dedicated point-source estimators can mitigate these effects.

Additionally, rather than using an estimate of the CMB lensing field obtained from the CMB itself, it is also possible to use other tracers of large-scale structure which are correlated with CMB lensing [224]. In particular the dusty, star-forming galaxies that comprise the cosmic infrared background (CIB) are strongly correlated with CMB lensing, due to their redshift distribution which peaks near $z \sim 2$ [556, 557]. The level of correlation is approximately 80% [545] and can in principle be improved using multifrequency maps of the CIB which select different emission redshifts [556].

Delensing can also impact the measurement of features of the CMB spectrum on small scales, in particular the CMB damping scale and the precise location of the acoustic peaks in harmonic space. Effects that can change these observables include changes in the effective number of neutrino species, the primordial helium fraction, and running of the spectral index of fluctuations. Using completely unlensed CMB spectra, rather than lensed spectra, can improve constraints on these parameters [298]. While the delensing procedure will not completely recover the unlensed CMB fluctuations for the S4 experiment, the low noise levels will enable the primordial CMB fluctuations to be measured with good enough fidelity that delensing should have a non-negligible impact on these parameter constraints.

7.6 Simulation Overview

Simulations of a CMB mission's data play a number of critical roles; specifically they are required for

- Forecasting: informing the design and development of a mission to ensure that it is capable of meeting its science goals.
- Validation and verification: ensuring that all of our data analysis tools meet their requirements and specifications.
- Uncertainty quantification and debiasing: providing an alternative to the full data covariance matrix when this is computationally intractable.

As shown in Figure 53, given a mission model (both instrument and observation) and a sky model (both CMB and extra-galactic and galactic foregrounds) we can generate a simulation of the mission data in any of its domains. However, there is an inevitable trade-off between how representative the simulation is of real data and the complexity of the input models and computational cost of generating the simulation. The choice of the simulation data domain will then be determined by the balance between the realism requirements and the complexity/cost constraints for the particular task at hand.

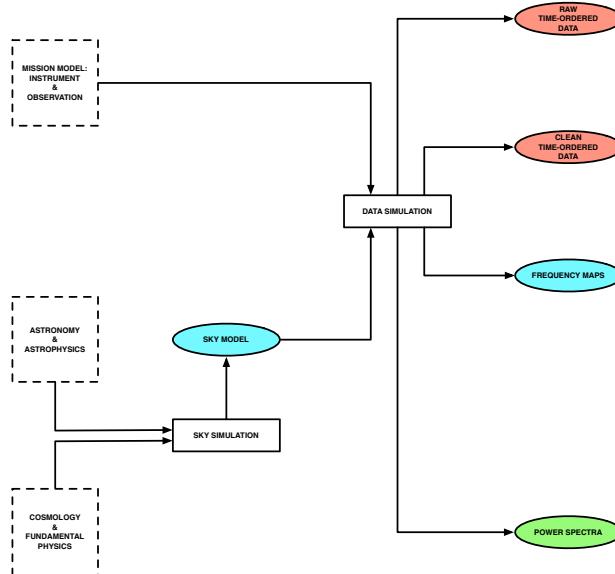


Figure 53. The CMB simulation pipeline

The generation of the input mission and sky models are themselves far from trivial tasks. The mission model is typically derived from pre-deployment measurements of the instrument properties refined by characterization from the data themselves, together with ancillary telescope and environmental data characterizing the observation; the sky model requires its own dedicated simulation capability which - since it is independent of the details of any single mission - can be a community-wide endeavor.

7.7 Sky Modeling

The capability of CMB-S4 to address its science program crucially depends on the possibility to separate the signals of interest from astrophysical emission originating from various astrophysical processes, and on the accuracy of the characterization of foreground residuals after such cleaning is performed.

The polarized CMB is mostly contaminated by diffuse emission from the interstellar medium of our own Milky Way. Both synchrotron and thermal dust are polarized—up to the level of tens of percents, depending on the observed region). Their integrated emission dominates over both the CMB E-modes and the CMB B-modes on large angular scales, and cannot be safely neglected at scales where B-modes from gravitational lensing dominate without robust analyses of their impact on lensing science. Other components, such as spinning dust, free-free emission, emission of molecular lines such as CO, could in principle be polarized at a lower level, of order 1-2 per cent or less, but measurements or upper limits are scarce, and not sufficient at this stage for robust predictions of the polarized amplitude of their emission over large patches of sky. For science on smaller angular scales, the presence of polarized extragalactic radio and infrared sources constitutes an additional source of contamination, which must be removed with a combination of masking or subtracting individual sources, and modeling residuals at the power-spectrum level.

Estimating more precisely the impact of foreground emission on the main science targets of CMB-S4 will require realistic simulations of the sky emission that can be used to test the effectiveness of component separation techniques and to assess any degradation of the error bars or possible biases due to residual foreground contamination. Figure 54 locates this modelling within the larger CMB simulation and data analysis pipeline.

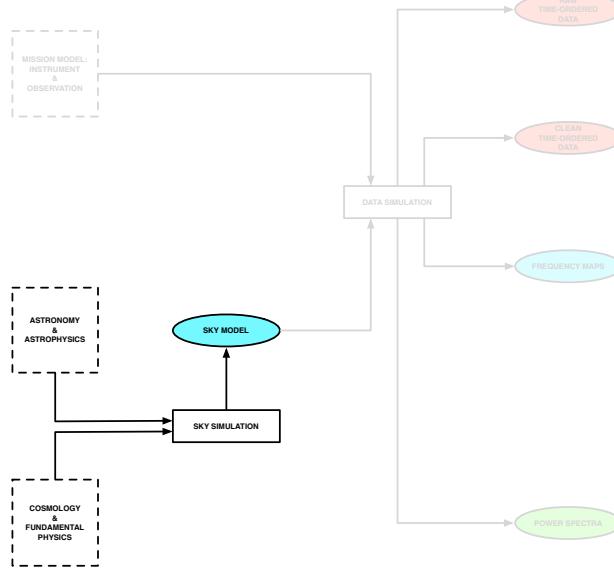


Figure 54. The sky modelling subset of the CMB simulation pipeline. The full simulation pipeline is shown in Figure 53.

The sky emission is naturally modeled as the sum of emission from different sources. These sources may be identified by their emission process (e.g. Galactic synchrotron, due to electrons spiralling in the Galactic magnetic field), or by their place of origin (e.g. emission of a particular extragalactic source). These emissions, as a function of sky pixel and frequency, must then be band-integrated and beam-integrated to

produce total emission maps as observed by the instrument. This latter part of the simulation pipeline is treated in Section 7.8; we concentrate in this section on the sky model.

A sky model is useful only as far as it captures adequately the characteristics of real sources of sky emission that play an essential role in the performance of cleaning techniques and on the amplitude and statistical properties of residual contamination after such cleaning is performed. The key characteristics of sky emission for foreground cleaning are:

- The coherence and decoherence of diffuse emission across observed frequencies, which is key to identifying foreground emission in the form of patterns that scale with an emission law different from that of the CMB;
- The existence or not of a simple parametric emission law for each component emission, such as power laws (for synchrotron) or modified blackbody emission (for dust components);
- The absolute level of foreground emission (in particular for those components that do not scale simply as a function of frequency, such as the superposition of many individual sources with a specific emission law each);
- Whether or not emissions for which the level of polarization is unknown or unclear (including possible surprises) are above the sensitivity objectives of CMB-S4, or can be safely neglected;
- The self-consistency of extragalactic emission (in particular CIB and SZ clusters) and CMB lensing. In particular, CIB as observed by Planck and by future instruments can be used to generate a proxy for the lensing potential, which can be used to partially de-lens the CMB B-modes.

The key challenges for constructing a sky model are hence:

- The reliability of models based on observations at angular resolution lower than that of CMB-S4, integrated in broad frequency bands, and with a sensitivity limit at least an order of magnitude worse than what will be achieved with CMB-S4; the complexity of sky emission below current sensitivity limits must be extrapolated on the basis of existing knowledge and theoretical models, taking into account past experience when orders of magnitude in sensitivity were gained;
- The self-consistency of CMB secondary anisotropies (lensing, SZ emission from hot intra-cluster gas and filaments, late ISW) and extra-Galactic foregrounds (CIB, radio and infrared sources) is crucial to both de-lensing, and to extragalactic science; generating reliable models over the entire Hubble volume is challenging, the evaluation of errors of such models even more so;
- The practical usability of the model (software engineering aspects for generating many simulations).

7.7.1 The Galactic interstellar medium

Strong evidence exists for variability of the physical properties of the interstellar medium of the Milky Way as a function of the line of sight. This variability implies that the properties vary across different regions of the Milky Way, with the total ISM emission in each line of sight being a superposition of emission from various regions. Even assuming that each such region has a simple parametric emission law, such as a power law or a modified blackbody, the superposition of such emission cannot be modeled with a single simple emission law. Modeling the Galactic ISM for future sensitive surveys such as CMB-S4 requires modeling this complexity

at the appropriate level. It seems reasonable to use a multi-layer approach, in which each ISM component is modeled as a superposition of several layers, with a simple (although pixel-dependent and polarization dependent) emission law for each such layer.

7.7.1.1 Synchrotron

The baseline Galactic synchrotron model we use here has a power law scaling with a modestly spatially varying spectral index. The emission templates are the Haslam 408 MHz data reprocessed by [558], and the WMAP 7-year 23 GHz Q/U maps [559] smoothed to 3 degree FWHM and with smaller scales added using the PSM code [560]. The spectral index map is derived using a combination of the Haslam 408 MHz data [561] and WMAP 23 GHz 7-year data [562]. The same scaling is used for intensity and polarization. This is the same prescription as used in the Planck Sky Model's v1.7.8 ‘power law’ option, but with the Haslam map updated to the version in [558].

Extensions to this model that we are exploring include a curved power law model with a single isotropic curvature index, and a polarization spectral index that steepens with Galactic latitude by $\delta\beta \sim 0.2$ from low to high latitude, as this is currently consistent with WMAP and Planck data.

7.7.1.2 Thermal dust

The baseline model we consider has thermal dust modelled as a single component modified blackbody. We use dust templates for emission at 545 GHz in intensity and 353 GHz in polarization from the Planck-2015 analysis, and scale these to different frequencies with a modified black-body spectrum using the spatially varying dust temperature and emissivity obtained from the Planck data using the Commander code [563]. This therefore assumes the same spectral index for polarization as for intensity. These templates are smoothed to degree scale.

Variations on this model that appear consistent with current data are a more strongly varying emissivity, e.g. up to $\sigma \sim 0.2$ dispersion on degree scales, in addition to different prescriptions for small-scale behaviour that account for turbulence in the magnetic field. A two (or more) component model for the dust, composed of the spatially varying sum of silicon and carbonaceous dust, each with a different emissivity, is also physically motivated.

7.7.1.3 Spinning dust

Spinning dust, or anomalous microwave emission, is nominally unpolarized. However, a fractional polarization of a few percent is physically possible and not excluded by current data. We construct a possible model for this polarization using the intensity templates for spinning dust from the Planck-2015 Commander fits [563], combined with the thermal dust polarization angles and an overall polarization fraction.

7.7.1.4 Other components

Other contributions to the intensity and polarization of the Milky Way at CMB-S4 frequencies, such as free-free emission and molecular line emission, are not expected to be at the same amplitude and degree of polarization as the components treated individually above. However, the full sky model will need to include

these components in at least the most pessimistic scenarios, unless further data is obtained that conclusively demonstrates they can be fully neglected.

7.7.2 CMB Secondary Anisotropies and Extragalactic Sources

The key goal for the extragalactic sky models of CMB-S4 is to provide fast and self-consistent simulations of CMB secondary anisotropies and extragalactic sources. These models will allow us to make more realistic forecasts. In our cosmological analyses they will allow us to Monte Carlo over the underlying astrophysical uncertainties of these secondaries and sources. Our plan to meet these challenges is modular and can be broken down as follows:

- We will use full hydrodynamical simulations of cosmological volume as the basis to parametrically model the complicated *gastrophysical* processes associated with extragalactic foregrounds.
- As the backbone of our model we will require fast simulations of growth of structure that generate halo catalogs for a large set of cosmological parameters.
- To have self-consistent maps we will have a flexible pipeline that generates simulated all sky maps which applies the parametric models from the hydrodynamical simulations to our backbone large-scale structure simulations and halo catalogs.

Hydrodynamical simulations of cosmological volumes are currently available which we can already use to model extragalactic foregrounds. These simulations will be used for the development and testing phases of the simulation pipeline. However, they are limited in their size and sub-grid modeling accuracy, and thus will not meet our accuracy requirements of CMB-S4. We will develop new full hydrodynamical simulations of cosmological volumes that include a variety of physical processes. An essential requirement of these simulations will be to capture growth and evolution of galaxies to cluster-size halos throughout cosmic time at a sufficient spatial resolution. Hydrodynamic simulations of this size and scale are already computationally feasible, the challenges will be the appropriate modeling of radiative cooling, star formation, and feedback processes in order to capture the global stellar and gas contents of these halos.

There are many different approaches already developed to provide us with the underlying large-scale structure simulations that will we build our extragalactic model upon. They vary in speed which tend to inversely scale with accuracy. A benefit of our modular and flexible approach is that we do not need to limit ourselves to one approach. In fact we will compare the various approaches to see how they bias our answers. It is in these simulations where we will vary cosmological parameters assuming that they only affect the growth of structure and not the *gastrophysical* properties of extragalactic foregrounds.

Our final product will be all sky maps. They will be in HEALPIX format to seamlessly interface with galactic and CMB simulated maps. The map products will include:

- Optical galaxies that correspond to the various overlapping surveys including LSST.
- Radio and dusty star-forming galaxy point sources.
- Unresolved CIB.
- Projected density maps (both total and gas) of the large-scale structure.

- Thermal and kinetic SZ maps.

We will explore the parameter space for each of the maps listed above and provide a sufficient number of realizations that we can marginalize over the many model uncertainties. For example, the lensing field can be constructed through a proper ray-tracing method from the projected density maps or via the Born approximation. Our self-consistent extragalactic sky model allows us to test various sources of contamination and systematic biases in our estimators. Additionally, any cross-correlation analyses can easily be checked and evaluated using these maps. **All the simulation products we create will become public.**

7.8 Data Simulation

The data simulation subset of the CMB simulation pipeline (Figure 55) takes the sky model and applies the mission model to it to generate a simulated data set for that mission. The mission model consists of two parts; the instrument model defines the data acquisition system (telescope, detectors, read-out), while the observation model defines its deployment (scanning strategy, environment). Depending on the degree of detail of the sky, instrument and observation model that we include, the resulting data set can be in any of the data domains - time-ordered (raw or clean), map, or spectral. Inevitably there is a trade-off between the realism of the simulation and the complexity and cost both of generating the model inputs and of performing the simulation, with the choice reflecting both the requirements of the subsequent analyses of the data set and the availability of computational resources.

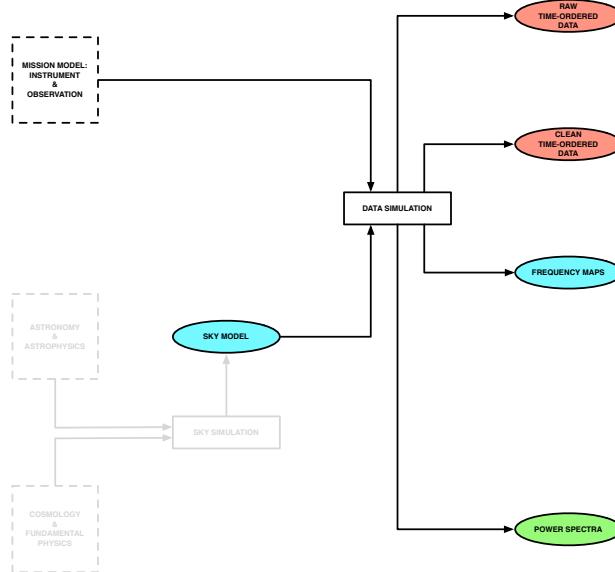


Figure 55. The data simulation subset of the CMB simulation pipeline

At the most detailed level, the observation model includes the telescope pointing (typically sampled more sparsely than the detectors), and its environment (comprising the atmosphere and surroundings for a ground-based telescope). Correspondingly, the instrument model includes each detector's polarized 4π -beam and bandpass (defining the optical power incident on the detector for a given pointing), and a model of its electronics and readout (defining the recorded output data resulting from that optical power).

7.8.1 Time Domain

TOD simulations are necessarily the most expensive to perform, but provide the most precise representation of the mission data. In particular they enable the injection of the full range of systematic effects into the data to assess strategies for their mitigation and to quantify any residuals. As such they are critical for the quantification of uncertainties due to inherently temporal data components such as noise. The TOD simulation is separated into signal and noise components, which are then added prior to the reduction of the total TOD.

For the signal simulation for a given detector, we first apply the detector’s bandpass to the sky model, component by component, to build up the total sky for that detector. We then reconstruct the detector pointing from the overall telescope pointing model and generate the astrophysical sky signal for each pointing by convolving the sky model map with the 4π beam. The astrophysical sky signal is added to additional simulated signals from atmospheric signal fluctuations and ground pickup (both of which will obviously induce correlated signals across the detectors), and the total signal is propagated through a simple model of the optics to include the polarization angle rotations and optical efficiencies of the optical stages. This results in the total millimeter-wave power incident on the detector. For simulating the clean TOD this is sufficient. However, for the raw TOD we now need to apply a physical model of the detector system and associated readout to convert the optical power into detector output. The details of the physical model depend on the detector technology, but as an example we consider a transition-edge superconducting (TES) bolometer read out with a multiplexed SQUID amplifier. The simulation would then need to model the flow of heat in the TES absorber and the flow of current and magnetic flux through the SQUID readout. Variations in ambient magnetic field could also be added at this stage. Such a simulation would also need to incorporate detector-detector correlations induced by crosstalk or thermal fluctuations. Additional filters applied by the readout electronics would also be included, including digitization with an analog to digital converter. For MKID or coherent receivers, the physical model would be different in detail, but would include a similarly detailed model.

For the noise simulation we can simply generate a white noise timestream and convolve it with the detector’s noise power spectral density (PSD), given in either analytic or numerical form. Cross-correlated noise can be included by simulating multiple noise timestreams each with their own PSD, with some being common to multiple detectors, while piecewise stationary noise simply requires us to use the appropriate PSD for each stationary interval.

7.8.2 Map Domain

The next level of abstraction from full TOD simulations is simulating the sky map that would be made from the TOD. The signal part of such simulations is straightforward: the various components of the sky model are bandpass-integrated and convolved with a beam and any filtering kernel that is applied to the actual data. Both of these operations can be performed on a per-detector basis (using the measured individual-detector beams and bandpasses), but this reduces much of the computational gain in going from TOD to map-level simulations, so it is more likely that maps would be simulated for large groups of detectors—possibly all detectors at a given observing frequency—at once. For experimental platforms that apply TOD filtering before mapmaking, it will be necessary to create a map-level (or two-dimensional Fourier-space) representation of the TOD filtering that results in simulated maps with the same modes missing or altered as the real map (or a map that has been constructed from full TOD simulations).

The simplest implementation for the noise part of map-level simulations is adding constant-amplitude white noise to every simulated map pixel. This ignores pixel-pixel correlations and incomplete coverage, both of which are naturally accounted for in full-TOD simulations. The exact nature of the correlated pixel-domain noise (or non-white Fourier-space noise) arises from a combination of non-white noise in the TOD and the scan strategy, and for some scan patterns can be analytically projected from the time domain to the pixel or 2d Fourier domain [564, 565, 566].

The effect of non-uniform coverage on the noise properties of the map is simple to simulate in the white-noise case: the uniform-coverage white-noise map is simply multiplied by the square root of the inverse of the “hit-count” map. The combined effect of non-white noise and non-uniform coverage—particularly if the

coverage map is not smooth on scales of the noise correlation—will be more difficult to simulate purely in the map domain.

7.8.3 Spectral Domain

Simulations at the one-dimensional power spectrum level are fast, computationally light, and can be used to explore large experimental and observing-strategy parameter spaces quickly and efficiently. As such, they will constitute the bulk of simulations used in forecasting for CMB-S4, particularly in the era in which the experimental design is not final, and reasonably fast communication and iteration between the experiment design and forecasting teams is crucial. Section 7.10.1.1 contains a detailed discussion of plans for implementing maximally realistic spectral-domain simulations.

7.9 The Simulation and Data Analysis Pipeline

The overall simulation and data analysis pipeline (Figure 56) can now be seen as both a top-down data reduction process and a wrap-around refinement of our mission and sky modeling. Typically the two phases are interleaved, with each new data reduction improving our mission and sky models, which are in turn fed back into an improved data reduction.

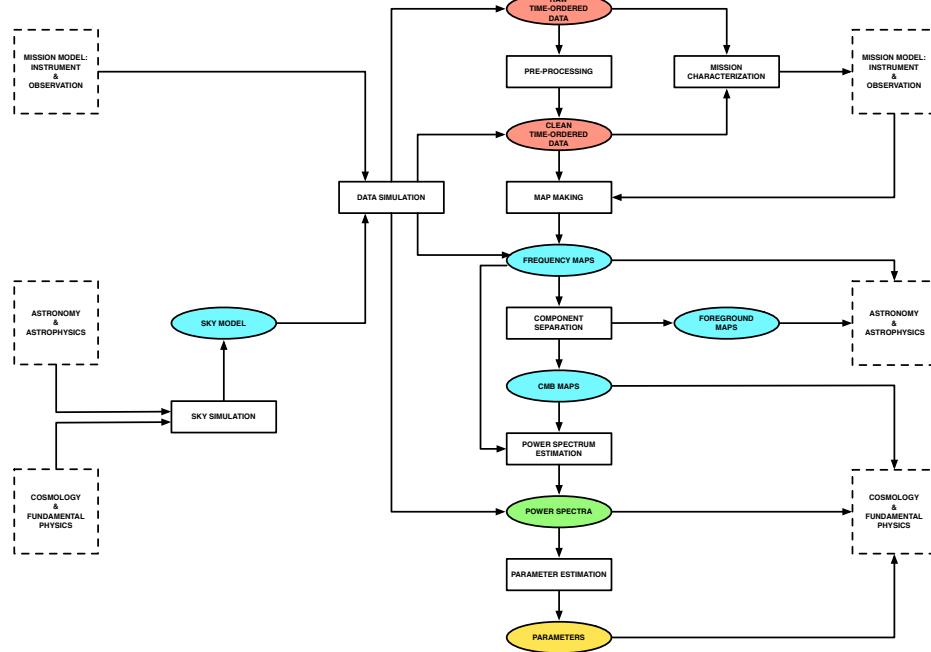


Figure 56. The full CMB simulation and data analysis pipeline

For CMB-S4 the most critical uses of this pipeline (or subsets thereof) are currently in forecasting and validation and verification. Further, the very discussion of such a pipeline provides a natural context in which to consider important implementation issues.

7.10 Forecasting

We emphasize the importance of accurate forecasting for CMB-S4. Forecasting efforts for Stage 1 and 2 experiments were hampered by lack of experience with previous deep polarization maps and little knowledge of high latitude Galactic foregrounds. Forecasting for CMB-S4 will be built on the solid foundation of map-derived evaluations of instrumental noise performance and astrophysical foreground levels from the Stage 2 experiments and the Planck satellite.

The forecasting approach will combine Fisher matrix-derived estimates of power spectrum errors with detailed map-level simulations. The spectral-domain projections are computationally easy, making them useful to explore the large parameter space of instrument and survey configurations. Map-domain simulations are used to ground the spectral-domain projections in reality and to challenge them with cases of real world astrophysical complexity. To gain the benefit of this complementarity, it is important that we maintain compatibility between these two forecasting approaches and establish agreement between them for simple questions before proceeding to more difficult tests.

A key input to the forecasting process are full-season noise maps from existing Stage 2 experiments, which encode actual noise performance and have been verified by null tests on real datasets. Performance of CMB-S4 can be estimated by rescaling these noise maps, which already contain reality factors such as detector yield, weather, observing efficiency, and filtering of sky modes. Systematic errors should be included in the projections, with unknown systematics allowed at a level that scales with the map noise used for jackknife null tests. Forecasting should also include our best knowledge of the astrophysical foregrounds and account for the impact of component separation on CMB-S4 science goals. The forecasting inputs will improve as we acquire data from Stage 3 experiments and possible complementary balloon-borne experiments, which will produce deeper noise maps and better assessments of foregrounds, as well as demonstrations of new techniques and technologies in development for Stage 4.

Here we describe the main approaches used by our community for forecasting the expected performance of CMB-S4. The central considerations for assessing the expected performance for large-scale B-modes are Galactic foregrounds, ability to delens the data, and a realistic assessment of instrument noise at large scales. For the smaller-scale polarization two-point functions (TE, EE) and the lensing four-point function ($\kappa\kappa$), extragalactic foregrounds and instrumental noise are the key considerations. To forecast the return of the thermal Sunyaev-Zel'dovich effects, an estimate of the expected cluster counts as a function of mass and redshift is the core statistic, combined with an estimate of how well the masses can be calibrated using overlaps with weak lensing surveys. For the kinetic Sunyaev-Zeldovich effect, extragalactic foregrounds and overlap with spectroscopic surveys must all be considered.

7.10.1 Forecasting CMB-S4 constraints on the tensor-to-scalar ratio

7.10.1.1 Spectrum-based domain forecasting

Power spectra are the primary tool used for CMB analysis. Forecasting the power spectrum uncertainty and resulting parameter constraints for CMB-S4 is an efficient and powerful tool to explore trade-offs in experiment design.

The bandpower covariance matrix describes the raw sensitivity of all auto and cross-spectra obtained between maps of T, E, and/or B modes at multiple observing frequencies, as well as the signal and noise correlations

that exist between these spectra. This covariance matrix includes contributions from the sample variance of signal fields (CMB and foreground) and instrumental noise, including signal \times noise terms. The signal variance depends on the assumed sky model, which can be modified to explore optimistic or pessimistic scenarios. As discussed above, estimates of the noise variance should be obtained by rescaling of noise levels that have actually been obtained by Stage 2 experiments (or Stage 3, when available). Only these scaled noise levels will include all the small “reality factors” that are incurred in operating a CMB experiment.

We will explicitly account for the impact of systematic errors by including them in the bandpower covariance matrix. For constraining tensor-to-scalar ratio, we are particularly concerned with effects that add B-mode power to maps. Even if the bandpowers are debiased using accurate simulations of such a systematic, it will still leave behind a noise floor due to its sample variance. Unknown and unforeseen sources of spurious signal in CMB-S4 will ultimately be constrained by jackknife null tests, which analyze a map constructed from the difference of two data subsets. The statistical power of the null test is set by the noise level of the maps, since any signal contributions should differ away. We can acknowledge this limitation by including in our projections an unknown systematic that adds B modes at the level of the null test uncertainty. Errors that are multiplicative in the signal, such as an absolute calibration error in the map, are best handled by adding nuisance parameters to the signal model.

Once we have a projection for the bandpower covariance matrix of CMB-S4, we can derive constraints on a parametrized model of cosmological and foreground signals via the Fisher information matrix. While we are most interested in parameter r , it is necessary to also consider the amplitude, spectrum, and spatial distribution of the dust and synchrotron foregrounds (see [50] for an example). The Fisher matrix formalism allows us to calculate the marginalized error on each parameter, with priors if desired, or to explore the degeneracies between parameters.

By compressing the data down to power spectra, it is feasible to use this technique to evaluate a wide range of survey designs. The parametrized signal model is also quite flexible and can include complications such as dust–synchrotron correlation or spatially varying foreground spectral indices. The limitation is that by considering the power spectrum only we are treating all signals as Gaussian, an approximation which must break down at some point for foregrounds. For this reason, it is important to have the ability to spot check the spectrum-based forecasts against map-based forecasts at specific choices of signal model.

7.10.1.2 Map-based domain forecasting

Foregrounds are intrinsically non-Gaussian, so it is beneficial to consider approaches directly in map space, to check the robustness of spectrum-based approaches, in particular in the case of pessimistic foregrounds where the spectral indices or dust emissivities have non-trivial spatial variation. Here one of the approaches our community uses is a Bayesian model fitting method, where the foregrounds are described parametrically using a physical model for each component.

Using this method, maps of the CMB plus Galactic foreground sky and expected noise are simulated at each of the CMB-S4 frequencies and integrated across the expected bandpasses, using Galactic models as described for example in Section 7.7. Simulations at ancillary frequencies that might be provided by other experiments, for example the Planck data, can also be included in the same way. A parameterized model is then fit to the simulated maps, for example fitting the CMB, thermal dust, and synchrotron in small pixels, and typically the synchrotron spectral index and dust emissivity and temperature in larger pixels of order degree-scale or larger. The BB power spectrum of the foreground-marginalized CMB map is then estimated using e.g., the MASTER [517] algorithm or a pixel-based likelihood, and converted into an estimate of r and its uncertainty. Our community has at least two such codes that can perform this procedure (Commander and BFoRe).

This method allows for an assessment of the expected bias on r if the model does not match the simulation, and shows how much the expected uncertainty on r would increase if more complicated foreground models are explored e.g. [567, 558]. It is more computationally expensive than spectral-domain forecasts though, so we limit this approach to a smaller subset of explorations.

7.10.2 Forecasting CMB-S4 constraints on parameters from TT/TE/EE/ $\kappa\kappa$

Throughout this Science Book we forecast the expected constraints on cosmological parameters from TT/TE/EE and $\kappa\kappa$ using Fisher matrix methods. This assumes that the resulting parameter distributions are close to Gaussian, which is sufficient for the majority of parameters we consider.

We assume that S4 data will be combined with existing Planck satellite data. We also assume that other non-CMB data will be available. In particular we consider measurements of Baryon Acoustic Oscillations from the DESI spectroscopic galaxy redshift survey. In some places we consider measurements of cosmic shear from the Large Synoptic Survey Telescope.

The codes we use either consider the unlensed maps and the lensing convergence map as the basic statistics, or the lensed power spectra of those maps together with the reconstructed $\kappa\kappa$ spectrum. For the power spectrum approach, to compute the Fisher matrix for the CMB we use the lensed power spectrum between each pair of fields X, Y :

$$\hat{C}_\ell^{XY} = \frac{1}{2\ell+1} \sum_{m=-\ell}^{m=\ell} x_{\ell m}^* y_{\ell m}. \quad (7.28)$$

The estimated power spectrum is Gaussian distribution to good approximation at small scales. In this case a full-sky survey has

$$-2 \ln \mathcal{L}(\boldsymbol{\theta}) = -2 \sum_\ell \ln p(\hat{C}_\ell | \boldsymbol{\theta}) = \sum_\ell \left[(\hat{C}_\ell - C_\ell(\boldsymbol{\theta}))^\top \mathbb{C}_\ell^{-1}(\boldsymbol{\theta}) (\hat{C}_\ell - C_\ell(\boldsymbol{\theta})) + \ln \det(2\pi \mathbb{C}_\ell(\boldsymbol{\theta})) \right] \quad (7.29)$$

where $\hat{C}_\ell = (\hat{C}_\ell^{TT}, \hat{C}_\ell^{TE}, \dots)$ contains auto- and cross-spectra and \mathbb{C}_ℓ is their covariance matrix. Discarding any parameter dependence in the power spectrum covariance matrix gives

$$F_{ij} = \sum_\ell \frac{\partial C_l^\top}{\partial \theta_i} \mathbb{C}_\ell^{-1} \frac{\partial C_l}{\partial \theta_j}. \quad (7.30)$$

Here the covariance matrix for the power spectra has elements

$$\mathbb{C}(\hat{C}_l^{\alpha\beta}, \hat{C}_l^{\gamma\delta}) = \frac{1}{(2l+1)f_{\text{sky}}} [(C_l^{\alpha\gamma} + N_l^{\alpha\gamma})(C_l^{\beta\delta} + N_l^{\beta\delta}) + (C_l^{\alpha\delta} + N_l^{\alpha\delta})(C_l^{\beta\gamma} + N_l^{\beta\gamma})], \quad (7.31)$$

where $\alpha, \beta, \gamma, \delta \in \{T, E, B, \kappa_c\}$ and f_{sky} is the effective fractional area of sky used.

Other codes construct the Fisher matrix using the unlensed temperature and polarization fields, and the lensing convergence field, rather than the suite of lensed two-point spectra and the lensing four-point function. Both approaches give consistent estimates.

The CMB lensing reconstruction noise is calculated using the [215] quadratic-estimator formalism. Our nominal approach is to neglect non-Gaussian terms in the power spectrum covariance. We also avoid including information from both lensed BB and the four-point $\kappa\kappa$, as they are covariant. The BB spectrum will not contribute as significantly to S4 constraints, compared to $\kappa\kappa$, and has a highly non-Gaussian covariance [555].

7.10.2.1 S4 specifications

The white noise part is given by

$$N_\ell^{\alpha\alpha} = (\Delta T)^2 \exp\left(\frac{\ell(\ell+1)\theta_{\text{FWHM}}^2}{8 \ln 2}\right) \quad (7.32)$$

for $\alpha \in \{T, E, B\}$, where ΔT (ΔP for polarization) is the map sensitivity in $\mu\text{K}\text{-arcmin}$ and θ_{FWHM} is the beam width.

We approximate the wide-field part of the S4 experiment as a 4-year survey using 250,000 detectors covering 40% of the sky in the lowest Galactic foreground region. We consider beam widths of both 1' and 2', and in some cases consider the effect of greater variation in the beam width. By scaling the simple map depths achieved by the BICEP2 experient, we estimate a white noise level of 1 $\mu\text{K}/\text{amin}$ in intensity, and $\sqrt{2}$ higher in polarization. For these smaller-scale forecasts we do not account for any possible mode filtering due to the mapping.

For polarization our nominal estimate is white noise, assuming that the tiny intrinsic polarization of the atmosphere, potentially combined with the use of polarization modulators, minimizes atmospheric contamination. In the longer term, these forecasts may be refined using scaled versions of noise spectra achieved in the field by experiments at the appropriate site. Eventually, full bandpower covariance matrices scaled from fielded experiments can also be used.

In these ‘non-r’ studies, we do not model the removal of Galactic foregrounds, instead assuming that all of the survey weight is focused at 150 GHz. In practice the survey would map the sky at a set of frequencies, and a component separation method used to estimate the CMB. Including these multiple frequencies will be the focus of future work; since Galactic foregrounds have a smaller effect on lensing and the CMB damping tail, we expect them to impact forecast constraints much less than for gravitational wave limits.

To address the issue of extragalactic foregrounds, we set as the default a maximum multipole for the recoverable information of $\ell_{\text{max}}^T = 3000$ and $\ell_{\text{max}}^P = 5000$ for CMB-S4, as foregrounds are expected to be limiting at smaller scales. We also set a minimum multipole due to the challenge of recovering large scales from the ground, and consider in general $\ell = 30$.

7.10.2.2 Non-S4 data specifications

We include Planck data at the scales $\ell < \ell_{\text{min}}$, nominally with $\ell_{\text{min}} = 30$, and we also add Planck data at all scales over the part of the sky not measured by S4 from Chile or the South Pole, approximated as covering an additional $f_{\text{sky}} = 0.2$.

For the noise levels of Planck, we assume that a data release including reliable polarization data will have happened before CMB-S4 data is taken, and forecast results that include TE and EE data and also large-scale temperature and polarization from HFI. This follows approaches in e.g. [568]. For the optical depth to reionization, we assume that Planck has reached currently published results, so impose a prior of $\tau = 0.06 \pm 0.01$.

In some cases we consider the addition of a cosmic-variance limited large-scale polarization measurement, as we might expect to get from a PIXIE or LiteBIRD satellite or potentially a high-altitude balloon.

Table 7-1. Forecasted LCDM parameters

	fiducial	Planck	S4+Planck
$100\Omega_bh^2$	2.22	± 0.017	± 0.003
Ω_ch^2	0.120	± 0.0014	± 0.0006
H_0	69.0	± 0.7	± 0.24
$10^9 A_s$	2.2	± 0.039	± 0.021
n_s	0.966	± 0.004	± 0.002
τ	0.06	± 0.01	± 0.006

To add information from Baryon Acoustic Oscillation (BAO) experiments, some of our codes add the BAO Fisher matrix

$$F_{ij}^{\text{BAO}} = \sum_k \frac{1}{\sigma_{f,k}^2} \frac{\partial f_k}{\partial \theta_i} \frac{\partial f_k}{\partial \theta_j} \quad (7.33)$$

where $f_k = r_s/d_V(z_k)$ is the sound horizon at photon-baryon decoupling r_s over the volume distance d_V to the source galaxies at redshift z_k . Other codes include the forecasted power spectra directly. We also follow standard approaches to including other low redshift probes.

7.10.2.3 Fisher code validation

We use six different Fisher matrix codes in the Science Book, but set up to use the same settings. We check that they all give consistent results for the Λ CDM model. These are shown in Table 7-1, which indicates the expected improvement of S4 over Planck for these parameters.

For forecasts quoted in this Science Book, we take the approach of adding just the individual parameters of interest to the basic LCDM set, unless stated.

7.10.3 Forecasting CMB-S4 constraints on parameters from tSZ/kSZ

As discussed in Section 4.1, some of the most important constraints on dark energy and tests of General Relativity will come from the thermal and kinematic Sunyaev-Zel'dovich effects. The information from the tSZ will mostly be in the form of the abundance and clustering of galaxy clusters, while the exact way in which kSZ information will be extracted is not fully determined, as this is a fairly new probe with rapidly developing analysis methods.

Forecasting constraints from cluster abundance is complicated by the fact that even current CMB experiments are not limited in their cluster-based constraints by raw sensitivity but rather by systematic uncertainties in the scaling relation between the tSZ observable and the cluster mass [376, 242]. Thus the cluster-based forecasting for CMB-S4 will likely be more focused on constraints on the observable-mass relation such as those that come from CMB-cluster lensing (see Section 4.1.2.1).

Forecasting constraints from kSZ will be an ongoing avenue of development. As early results become more mature, and the community explores new ways of measuring this signal [569, 570, 571, 572, 573, 574], the exact methods for forecasting will become more clear.

7.11 Implementation Issues

7.11.1 Time-Ordered Data Volume & High Performance Computing

The quest for ever-fainter signals in the CMB drives us to gather ever-larger time-ordered data (TOD) sets to obtain the necessary signal-to-noise to uncover them. As Figure 57 shows, the volumes of ground-based, balloon-borne and satellite CMB data sets have exhibited exponential growth over the last 20 years and are anticipated to do so again over the next. Moreover, for suborbital experiments the exponent exactly matches that of Moore's Law for the growth of computing capability, where we use as a proxy here the peak performance of the flagship high performance computing (HPC) system at the DOE's National Energy Research Scientific Computing (NERSC) Center at any epoch (reflecting the widespread use of NERSC for CMB data analyses over the last 20 years).

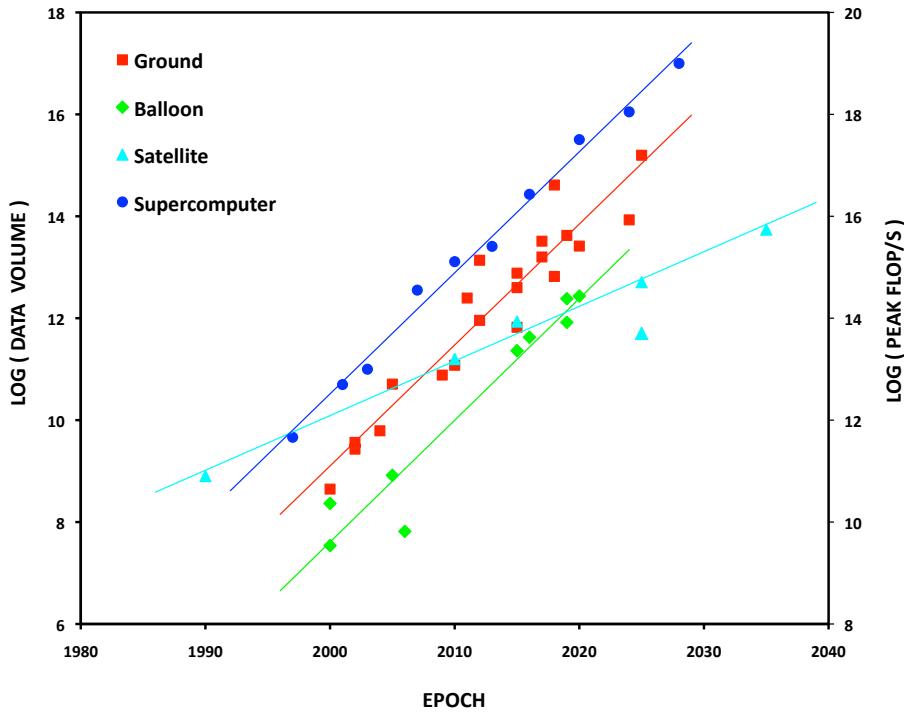


Figure 57. Exponential growth of CMB time-ordered data volume and HPC capability: 1990 – 2030.

Furthermore, in the absence of a full covariance matrix we rely on Monte Carlo methods for uncertainty quantification and debiasing, and to achieve the desired percent-level statistical uncertainty requires us to simulate and reduce 10^4 realizations of the data. Taken together, this implies that all TOD-processing steps (in simulation or analysis) (Figure 58) must employ algorithms that scale no worse than linearly in the number of samples, and that these algorithms must *collectively* be implemented efficiently on the largest high performance computing (HPC) platforms available to us.

The most massive Monte Carlo sets generated to date have been the Full Focal Plane (FFP) sets in support of the analysis of the Planck satellite data, with FFP8 comprising 10^4 realizations of the mission reduced to $O(10^6)$ maps. Key to achieving this scale has been an aggressive optimization of the software stack,

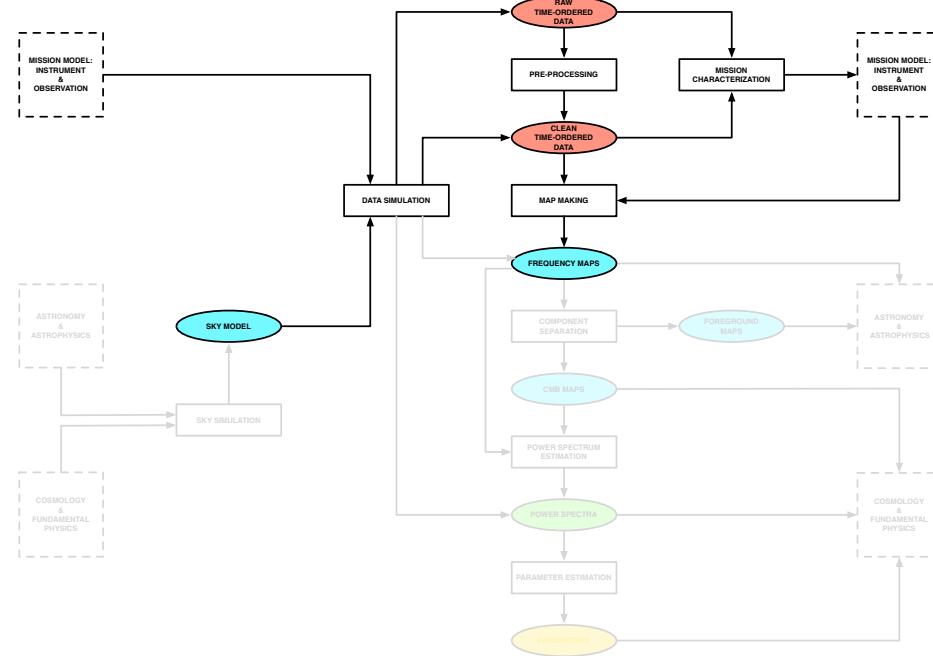


Figure 58. Time-ordered data processing elements of the CMB simulation and data analysis pipeline.

coupled with system-specific tuning over 6 generations of NERSC supercomputer. In particular wherever possible TOD input/output (IO) is removed from the pipeline so that, for example, instead of pre-computing the TOD and then pre-processing/mapping it, each realization is generated on demand and passed to the analysis pipeline in memory. While this necessitates the re-simulation of a realization should further analysis be required, it is still very substantially faster than writing it to disk and reading it back in. Similarly, inter-process communication is optimized by using a hybridized MPI/OpenMP implementation that employs explicit message passing for inter-node, and threading for intra-node, communication.

A critical challenge for CMB-S4 will be to develop this capability for a dataset 1000x the size of Planck's on the coming generations of energy-constrained HPC architectures, with their increased heterogeneity and deeper memory hierarchies. In the short term such systems will be based on either graphical programming unit (GPU) or many integrated core (MIC) technologies; in the longer term the path to Moore's Law is unclear.

Experimental Approach

8.1 Introduction

The rich science case for the CMB-S4 experiment has been articulated in the first seven chapters of this Science Book. This final chapter describes how the design of the CMB-S4 experiment flows from the science drivers. Simulations of CMB-S4 observations with CMB, foregrounds, and noise will determine the basic instrument parameters: the number and size of instrument platforms, observing bands, and the relative sensitivity in each band. Results from previous experiments will provide a foundation for detailed modeling of CMB-S4, which must incorporate systematic errors due to ground pickup, beam shape, pointing, and polarization errors. The earlier experiments will also inform the choice of scan patterns for CMB-S4 observations.

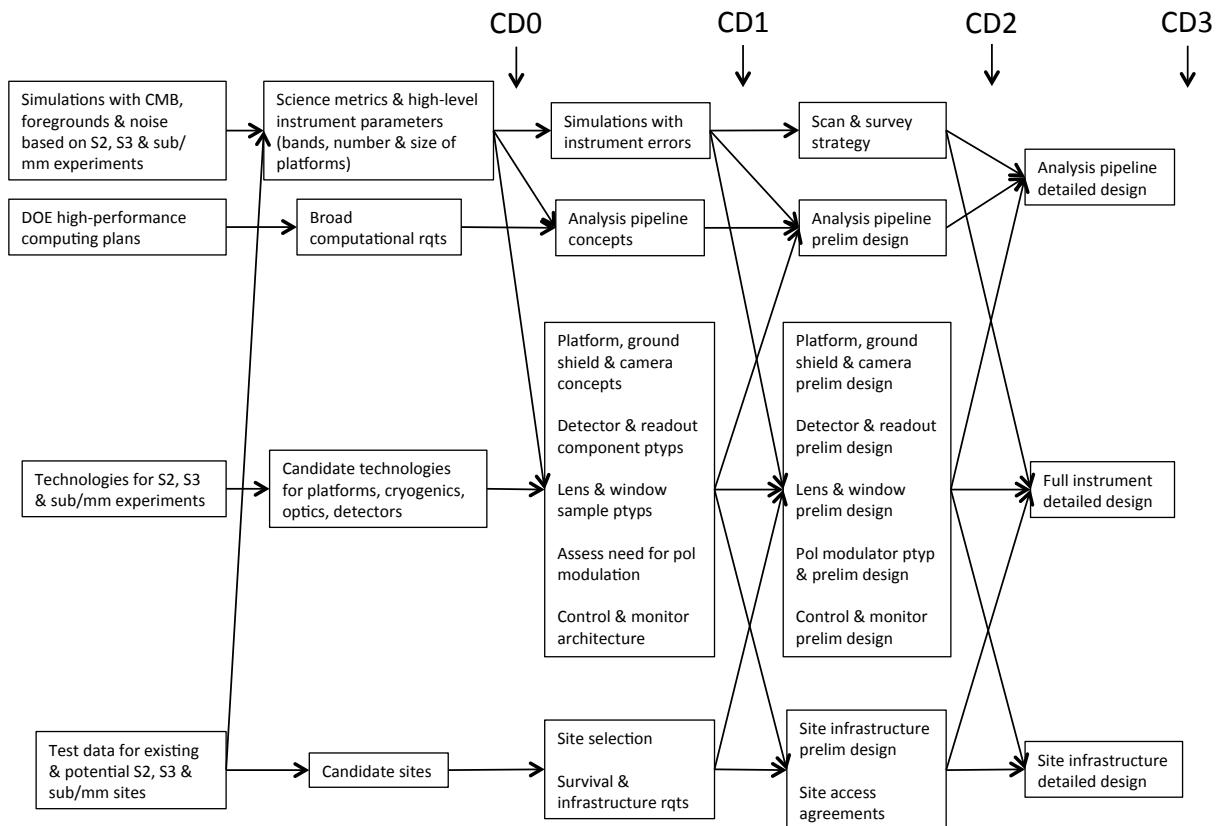


Figure 59. CMB-S4 Roadmap

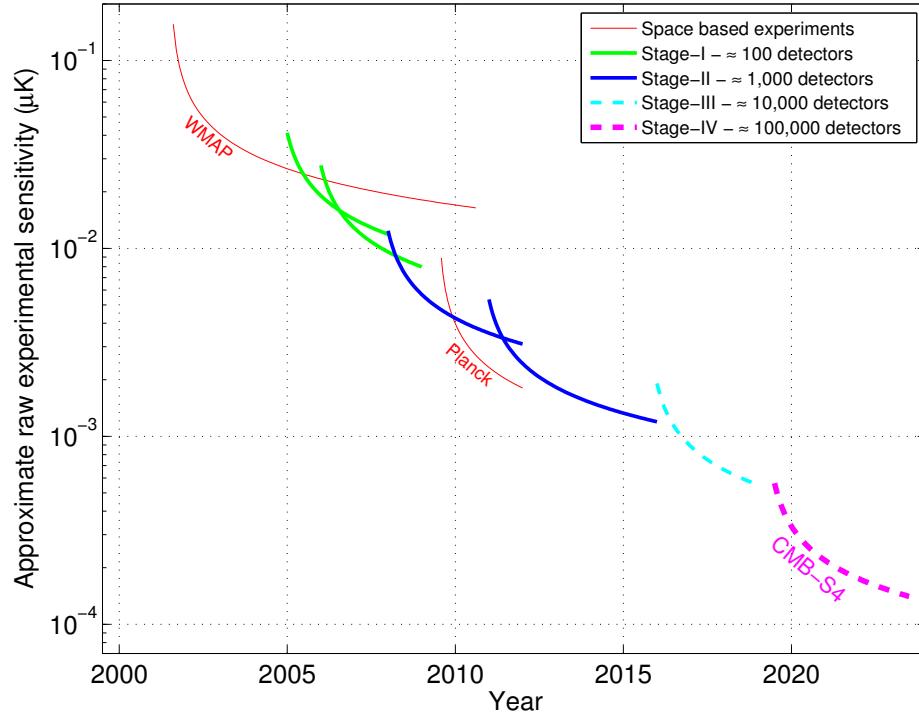


Figure 60. Plot illustrating the evolution of the raw sensitivity of CMB experiments, which scales as the total number of bolometers. Ground-based CMB experiments are classified into Stages with Stage II experiments having $O(1000)$ detectors, Stage III experiments having $O(10,000)$ detectors, and a Stage IV experiment (such as CMB-S4) having $O(100,000)$ detectors.

A summary of the technology roadmap for CMB-S4 is shown in Fig. 59. As mentioned, the initial choice of instrument parameters will be informed by our experience with Stage 2 and 3 CMB experiments. Technology experience from previous experiments also informs the list of candidate technologies for CMB-S4 and a rough costing for CD0. By CD1, conceptual design and critical technology development is complete, with instrument concepts feeding into the analysis pipeline design and site infrastructure design. The coupling between instrument, analysis pipeline, and site infrastructure design continues through preliminary design, which defines the baseline project at CD2, and detailed design, which delivers a construction-ready project at CD3.

8.2 Technical Development

8.2.1 Detector Arrays

State-of-the-art CMB detectors are sensitivity limited where the dominant noise in an individual detector element comes from shot noise arising from the arrival time of the photons. Thus, achieving the required CMB-S4 sensitivity requires increasing the number of detected modes, which is straightforward to achieve by increasing the number of detectors (see Fig. 60). CMB-S4 will have 500 times more detectors than the current

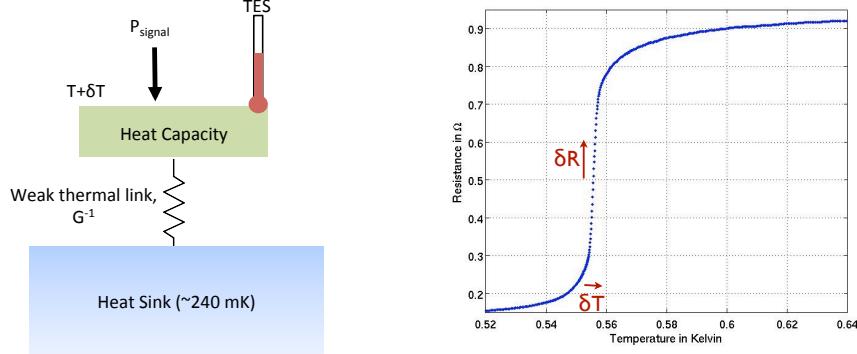


Figure 61. Left: Illustration of a thermal circuit for a typical Transition Edge Sensor (TES) detector highlighting the principles of signal detection. A weakly thermally sunk heat capacity absorbs power, P_{signal} , which is to be measured. Variations in the absorbed power change the heat capacity's temperature, which is measured by a TES operating under strong electro-thermal feedback. Right: Plot of resistance versus temperature for a typical TES illustrating the principles of negative electro-thermal feedback [575]. The TES is voltage biased into the middle of its superconducting-to-normal transition. Small changes in the TES temperature produce large changes in the TES resistance. Since the TES is voltage biased, an increase (or decrease) in the temperature produces an increase (or decrease) in the resistance leading to a decrease (or increase) in the Joule heating power supplied by the bias. This canceling effect corresponds to a strong negative electro-thermal feedback making the current through the TES nearly proportional to P_{signal} .

state-of-the-art Stage 2 experiments, 30 times more than planned Stage 3 experiments, making scaling the primary technical challenge of CMB-S4.

Towards this end, the baseline plan is that CMB-S4 will utilize Transition Edge Sensor (TES) bolometers as its detector technology, although we will follow the development of alternate technologies such as MKIDs. A TES is an ultra-sensitive thermometer consisting of a thin superconducting film weakly heat-sunk to a bath temperature much lower than the superconductor T_c (see Fig. 61, left). The principles of operation are simple to understand. By supplying electrical power to the TES, we can raise the temperature of the sensor so that the film is in the middle of its superconducting-to-normal transition (see Fig. 61, right). If the electrical power is supplied via a voltage bias, a negative feedback loop is established [575]. Small changes to the TES temperature, arising from thermal fluctuations (noise) or changes in the absorbed power from a source (signal), lead to large changes in the TES resistance. The change in resistance creates a canceling effect because increases (or decreases) in temperature produce decreases (or increases) in Joule heating power. This negative electro-thermal feedback is very strong because the transition is very sharp. It linearizes the detector response and expands the detector bandwidth.

The TES has a number of strengths making it an attractive technology to pursue for CMB-S4. First, TES detectors are fabricated via micro-machining of thin films deposited on silicon wafer substrates. As a consequence, the fundamental production unit for TES devices are arrays of detectors (see Fig. 62), an important attribute when considering the production of the 500,000 detectors required by CMB-S4. Second TES devices are low-impedance ($\leq 1 \Omega$) and can be multiplexed with modern-day Superconducting QUantum Interference Device (SQUID) multiplexers [576, 577, 578]. Multiplexed readouts are important for operating large detector arrays at sub-Kelvin temperatures and are essential for CMB-S4. Lastly, TES detectors have been successfully deployed as focal planes at the forefront of CMB measurements from 40 GHz to beyond the ~ 300 GHz requirement for CMB-S4.

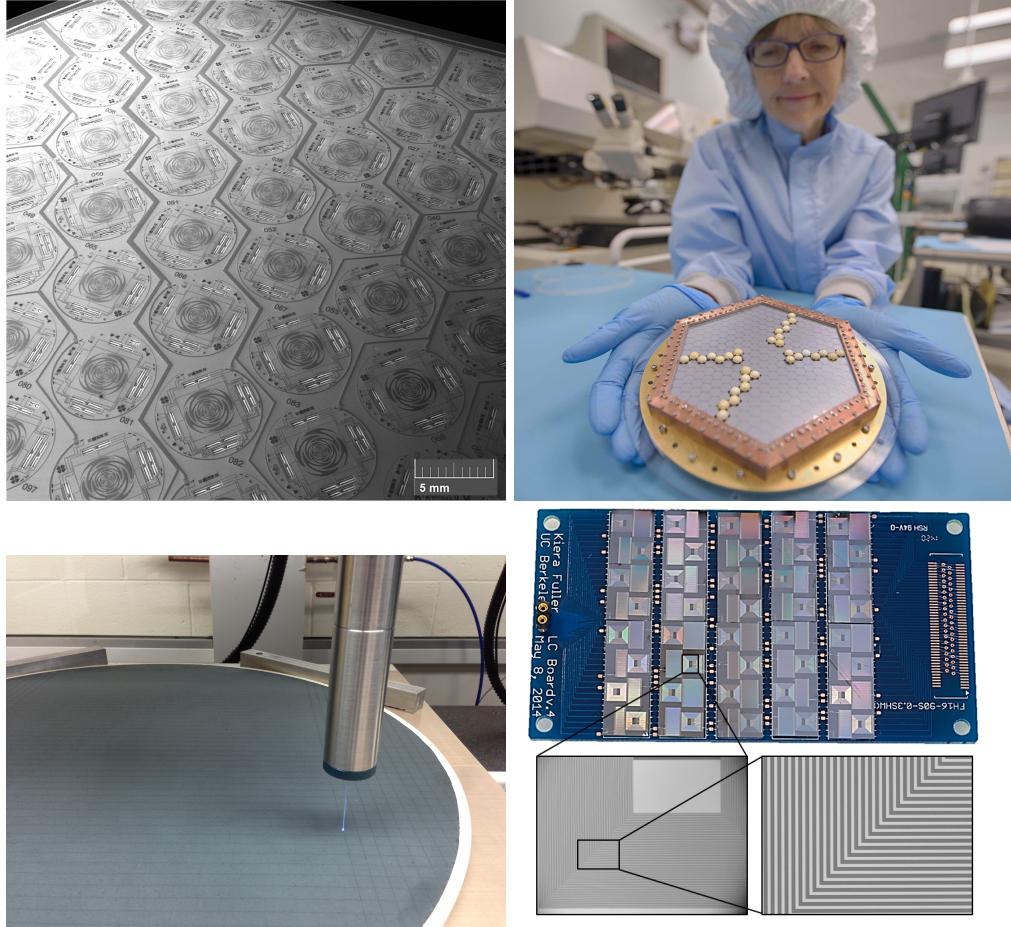


Figure 62. Illustrations of technology development at each of the four DOE labs working on CMB-S4. Clockwise from upper left: Multichroic bolometer array fabricated at ANL; Array module assembled at Fermilab; Superconductor resonators for bolometer readout fabricated at LBNL; and Laser-dicing of an optical filter at SLAC.

The TES was invented by HEP for detecting Dark Matter and neutrinos. Its subsequent integration into CMB focal planes has enabled kilo-pixel arrays realizing the Stage 2 CMB program and ushering in an era of unprecedented sensitivity. TES-based CMB detectors are the favored technology among Stage 2 and proposed Stage 3 experiments, and have a clear path to the sensitivities required by CMB-S4. The ubiquity of TES detectors for CMB illustrates the direct connection between HEP-invented technology and CMB science.

8.2.1.1 Development of mass production capability at the national labs

Delivering a half-million background-limited bolometers necessitates a change in the execution of the U.S. ground-based CMB program. The current program consists of a number of independent (primarily university) efforts, each focused on the development and delivery of their own instrument. The involvement of HEP in the current process has been through small investments targeted at specific technical contributions. Current detector fabrication R&D for each of these independent efforts aims to field $\sim 10,000$ detectors over the course

of three years with throughput limited by fabrication and testing resources. This level of production capacity is sufficient for Stage 3 experiments, but is inadequate for the challenges of CMB-S4. Realizing CMB-S4 will require a radical change in this approach where HEP resources take the leading role. Of particular importance is an increased participation and support of national labs to provide resources which are unavailable to university groups. In particular, 1) leveraging micro-fabrication tools and expertise available at multi-purpose national labs is essential for the successful fabrication of the CMB-S4 detectors, 2) supporting the computing infrastructure to support the greatly increased data rate, dataset size and analysis complexity, and 3) leveraging experience and resources for partnering with industry and the commercial sector.

The CMB-S4 experimental program will build on the success of Stage 2 and 3 CMB experiments. It will be a coherent effort incorporating resources from both national labs and university groups.

The necessary steps required for developing array fabrication ability for CMB-S4 include:

- **Improved Production Reliability** The favored technology for CMB-S4 are TES bolometers coupled through superconducting microstrip. Critical for the TES technology is reliable and optimal superconducting microstrip performance at millimeter wavelengths. Recent work on microstrip-coupled CMB detectors have demonstrated that it is possible to make superconducting microstrip which is virtually loss-less at the required frequencies [579]; however, the fabrication yield needs to be improved for CMB-S4. Thus, one of the principle components of the CMB-S4 program is developing a reliable mass production process. Such work requires well maintained tooling, dedicated materials deposition, and understanding and control of all the materials dependent loss mechanisms.
- **Increased Production Volume and Throughput** Achieving $O(500,000)$ TES detectors demands new investment into TES array production resources. The required production throughput needs access to micro-fabrication resources with exclusive control of the thin film deposition systems. This exclusive access to microfabrication tooling falls squarely within the domain of national labs, or by technology transfer to a qualified industrial partner. Additionally, an extensive program of detector testing, characterization, and quality control is crucial for the mass production of 500,000 TES bolometers. This requirement will be met by establishing test facilities and organizing a quality assurance program among the universities and national labs.

8.2.2 Multiplexed Detector Readout

Multiplexed TES readouts are required for implementing focal planes with more than 1000 detector elements and will continue to be an active component of the CMB-S4 R&D program. Modest improvements over existing fielded Stage 2 multiplexer technology will be sufficient for the needs of CMB-S4. However, recent developments with microwave-based readout techniques for TES detectors may lead to new multiplexer technologies with broader applicability and lower cost, and could be synergistic with microwave kinetic inductance detector (MKID) readout development efforts.

8.2.3 Telescope Architecture and Reimaging Optics

There are several types of telescope architectures that are currently being used successfully in Stage 2 experiments. Most telescopes are reflector type (mirror-based), although the BICEP/KECK series of experiments are refractors (lens-based). Both types of telescope offer high performance, but the refractors

are limited in diameter and therefore angular resolution. In reflector telescopes, both standard Gregorian and crossed Dragone designs have been used successfully.

As part of the CMB-S4 technology roadmap, optics studies will be performed to develop a design with both large optical throughput to accommodate large detector arrays and high polarization purity.

The large size of the CMB-S4 focal planes together with the required sub-Kelvin operating temperature necessitates the development of new broad-band large aperture refractive reimaging optics which permit high throughput at millimeter wavelengths while blocking infrared thermal emission. Upcoming Stage 3 experiments will serve as a proving ground for some new cryogenic optics technologies. CMB-S4 will build on these Stage 3 accomplishments with the goal of developing manufacturing techniques to yield a large number of customized cryogenic optics with optimal performance.

8.2.4 Polarization Modulation

Polarization modulators, such as continuously rotated half-wave plates that are used in ABS, POLARBEAR, and AdvACT, and a variable-delay polarization modulator (VPM) used in CLASS can give improved rejection of atmospheric fluctuation noise and a reduction in intensity-to-polarization conversion when compared to bolometer differencing. In the CMB-S4 technology roadmap, we will use the experience from Stage 3 experiments to assess the tradeoffs of using polarization modulators, and we will develop large-throughput and high-bandwidth modulators.

References

- [1] **Virgo, LIGO Scientific** Collaboration, B. P. Abbott *et al.*, “Observation of Gravitational Waves from a Binary Black Hole Merger,” *Phys. Rev. Lett.* **116** no. 6, (2016) 061102, [arXiv:1602.03837 \[gr-qc\]](https://arxiv.org/abs/1602.03837).
- [2] J. Khoury, B. A. Ovrut, P. J. Steinhardt, and N. Turok, “The Ekpyrotic universe: Colliding branes and the origin of the hot big bang,” *Phys. Rev.* **D64** (2001) 123522, [arXiv:hep-th/0103239 \[hep-th\]](https://arxiv.org/abs/hep-th/0103239).
- [3] J. Khoury, B. A. Ovrut, N. Seiberg, P. J. Steinhardt, and N. Turok, “From big crunch to big bang,” *Phys. Rev.* **D65** (2002) 086007, [arXiv:hep-th/0108187 \[hep-th\]](https://arxiv.org/abs/hep-th/0108187).
- [4] P. J. Steinhardt and N. Turok, “Cosmic evolution in a cyclic universe,” *Phys. Rev.* **D65** (2002) 126003, [arXiv:hep-th/0111098 \[hep-th\]](https://arxiv.org/abs/hep-th/0111098).
- [5] A. Nayeri, R. H. Brandenberger, and C. Vafa, “Producing a scale-invariant spectrum of perturbations in a Hagedorn phase of string cosmology,” *Phys. Rev. Lett.* **97** (2006) 021302, [arXiv:hep-th/0511140 \[hep-th\]](https://arxiv.org/abs/hep-th/0511140).
- [6] R. H. Brandenberger, “The Matter Bounce Alternative to Inflationary Cosmology,” [arXiv:1206.4196 \[astro-ph.CO\]](https://arxiv.org/abs/1206.4196).
- [7] Y.-F. Cai and E. Wilson-Ewing, “A Λ CDM bounce scenario,” *JCAP* **1503** no. 03, (2015) 006, [arXiv:1412.2914 \[gr-qc\]](https://arxiv.org/abs/1412.2914).
- [8] J. de Haro and Y.-F. Cai, “An Extended Matter Bounce Scenario: current status and challenges,” *Gen. Rel. Grav.* **47** no. 8, (2015) 95, [arXiv:1502.03230 \[gr-qc\]](https://arxiv.org/abs/1502.03230).
- [9] A. Lewis, A. Challinor, and A. Lasenby, “Efficient computation of CMB anisotropies in closed FRW models,” *Astrophys. J.* **538** (2000) 473–476, [arXiv:astro-ph/9911177 \[astro-ph\]](https://arxiv.org/abs/astro-ph/9911177).
- [10] **BICEP2, Keck Array** Collaboration, P. A. R. Ade *et al.*, “BICEP2 / Keck Array V: Measurements of B-mode Polarization at Degree Angular Scales and 150 GHz by the Keck Array,” *Astrophys. J.* **811** (2015) 126, [arXiv:1502.00643 \[astro-ph.CO\]](https://arxiv.org/abs/1502.00643).
- [11] **SPT** Collaboration, R. Keisler *et al.*, “Measurements of Sub-degree B-mode Polarization in the Cosmic Microwave Background from 100 Square Degrees of SPTpol Data,” *Astrophys. J.* **807** no. 2, (2015) 151, [arXiv:1503.02315 \[astro-ph.CO\]](https://arxiv.org/abs/1503.02315).
- [12] T. Essinger-Hileman *et al.*, “CLASS: The Cosmology Large Angular Scale Surveyor,” *Proc. SPIE Int. Soc. Opt. Eng.* **9153** (2014) 91531I, [arXiv:1408.4788 \[astro-ph.IM\]](https://arxiv.org/abs/1408.4788).
- [13] D. H. Lyth, “What would we learn by detecting a gravitational wave signal in the cosmic microwave background anisotropy?,” *Phys. Rev. Lett.* **78** (1997) 1861–1863, [arXiv:hep-ph/9606387 \[hep-ph\]](https://arxiv.org/abs/hep-ph/9606387).
- [14] A. D. Linde, “Particle physics and inflationary cosmology,” *Contemp. Concepts Phys.* **5** (1990) 1–362, [arXiv:hep-th/0503203 \[hep-th\]](https://arxiv.org/abs/hep-th/0503203).
- [15] N. Kaloper, A. Lawrence, and L. Sorbo, “An Ignoble Approach to Large Field Inflation,” *JCAP* **1103** (2011) 023, [arXiv:1101.0026 \[hep-th\]](https://arxiv.org/abs/1101.0026).
- [16] C. Csaki, N. Kaloper, J. Serra, and J. Terning, “Inflation from Broken Scale Invariance,” *Phys. Rev. Lett.* **113** (2014) 161302, [arXiv:1406.5192 \[hep-th\]](https://arxiv.org/abs/1406.5192).

- [17] D. E. Kaplan and R. Rattazzi, “A Clockwork Axion,” [arXiv:1511.01827 \[hep-ph\]](https://arxiv.org/abs/1511.01827).
- [18] K. Choi and S. H. Im, “Realizing the relaxion from multiple axions and its UV completion with high scale supersymmetry,” *JHEP* **01** (2016) 149, [arXiv:1511.00132 \[hep-ph\]](https://arxiv.org/abs/1511.00132).
- [19] X. G. Wen and E. Witten, “World Sheet Instantons and the Peccei-Quinn Symmetry,” *Phys. Lett.* **B166** (1986) 397.
- [20] K. Freese, J. A. Frieman, and A. V. Olinto, “Natural inflation with pseudo - Nambu-Goldstone bosons,” *Phys.Rev.Lett.* **65** (1990) 3233–3236.
- [21] F. C. Adams, J. R. Bond, K. Freese, J. A. Frieman, and A. V. Olinto, “Natural inflation: Particle physics models, power law spectra for large scale structure, and constraints from COBE,” *Phys. Rev.* **D47** (1993) 426–455, [arXiv:hep-ph/9207245 \[hep-ph\]](https://arxiv.org/abs/hep-ph/9207245).
- [22] T. Banks, M. Dine, P. J. Fox, and E. Gorbatov, “On the possibility of large axion decay constants,” *JCAP* **0306** (2003) 001, [arXiv:hep-th/0303252 \[hep-th\]](https://arxiv.org/abs/hep-th/0303252).
- [23] N. Arkani-Hamed, L. Motl, A. Nicolis, and C. Vafa, “The String landscape, black holes and gravity as the weakest force,” *JHEP* **06** (2007) 060, [arXiv:hep-th/0601001 \[hep-th\]](https://arxiv.org/abs/hep-th/0601001).
- [24] J. E. Kim, H. P. Nilles, and M. Peloso, “Completing natural inflation,” *JCAP* **0501** (2005) 005, [arXiv:hep-ph/0409138 \[hep-ph\]](https://arxiv.org/abs/hep-ph/0409138).
- [25] T. Rudelius, “On the Possibility of Large Axion Moduli Spaces,” *JCAP* **1504** no. 04, (2015) 049, [arXiv:1409.5793 \[hep-th\]](https://arxiv.org/abs/1409.5793).
- [26] A. de la Fuente, P. Saraswat, and R. Sundrum, “Natural Inflation and Quantum Gravity,” *Phys. Rev. Lett.* **114** no. 15, (2015) 151303, [arXiv:1412.3457 \[hep-th\]](https://arxiv.org/abs/1412.3457).
- [27] T. Rudelius, “Constraints on Axion Inflation from the Weak Gravity Conjecture,” *JCAP* **1509** no. 09, (2015) 020, [arXiv:1503.00795 \[hep-th\]](https://arxiv.org/abs/1503.00795).
- [28] J. Brown, W. Cottrell, G. Shiu, and P. Soler, “Fencing in the Swampland: Quantum Gravity Constraints on Large Field Inflation,” *JHEP* **10** (2015) 023, [arXiv:1503.04783 \[hep-th\]](https://arxiv.org/abs/1503.04783).
- [29] T. C. Bachlechner, C. Long, and L. McAllister, “Planckian Axions and the Weak Gravity Conjecture,” *JHEP* **01** (2016) 091, [arXiv:1503.07853 \[hep-th\]](https://arxiv.org/abs/1503.07853).
- [30] J. Brown, W. Cottrell, G. Shiu, and P. Soler, “On Axionic Field Ranges, Loopholes and the Weak Gravity Conjecture,” [arXiv:1504.00659 \[hep-th\]](https://arxiv.org/abs/1504.00659).
- [31] B. Heidenreich, M. Reece, and T. Rudelius, “Weak Gravity Strongly Constrains Large-Field Axion Inflation,” *JHEP* **12** (2015) 108, [arXiv:1506.03447 \[hep-th\]](https://arxiv.org/abs/1506.03447).
- [32] B. Heidenreich, M. Reece, and T. Rudelius, “Sharpening the Weak Gravity Conjecture with Dimensional Reduction,” [arXiv:1509.06374 \[hep-th\]](https://arxiv.org/abs/1509.06374).
- [33] K. Kooner, S. Parameswaran, and I. Zavala, “Warping the Weak Gravity Conjecture,” [arXiv:1509.07049 \[hep-th\]](https://arxiv.org/abs/1509.07049).
- [34] E. Silverstein and A. Westphal, “Monodromy in the CMB: Gravity Waves and String Inflation,” *Phys.Rev.* **D78** (2008) 106003, [arXiv:0803.3085 \[hep-th\]](https://arxiv.org/abs/0803.3085).
- [35] L. McAllister, E. Silverstein, and A. Westphal, “Gravity Waves and Linear Inflation from Axion Monodromy,” *Phys.Rev.* **D82** (2010) 046003, [arXiv:0808.0706 \[hep-th\]](https://arxiv.org/abs/0808.0706).

- [36] N. Kaloper and L. Sorbo, “A Natural Framework for Chaotic Inflation,” *Phys. Rev. Lett.* **102** (2009) 121301, [arXiv:0811.1989 \[hep-th\]](#).
- [37] M. Berg, E. Pajer, and S. Sjors, “Dante’s Inferno,” *Phys. Rev.* **D81** (2010) 103535, [arXiv:0912.1341 \[hep-th\]](#).
- [38] E. Palti and T. Weigand, “Towards large r from $[p, q]$ -inflation,” *JHEP* **04** (2014) 155, [arXiv:1403.7507 \[hep-th\]](#).
- [39] L. McAllister, E. Silverstein, A. Westphal, and T. Wrane, “The Powers of Monodromy,” *JHEP* **09** (2014) 123, [arXiv:1405.3652 \[hep-th\]](#).
- [40] F. Marchesano, G. Shiu, and A. M. Uranga, “F-term Axion Monodromy Inflation,” *JHEP* **09** (2014) 184, [arXiv:1404.3040 \[hep-th\]](#).
- [41] R. Blumenhagen, C. Damian, A. Font, D. Herschmann, and R. Sun, “The Flux-Scaling Scenario: De Sitter Uplift and Axion Inflation,” [arXiv:1510.01522 \[hep-th\]](#).
- [42] A. Hebecker, J. Moritz, A. Westphal, and L. T. Witkowski, “Axion Monodromy Inflation with Warped KK-Modes,” *Phys. Lett.* **B754** (2016) 328–334, [arXiv:1512.04463 \[hep-th\]](#).
- [43] I. Ben-Dayan and R. Brustein, “Cosmic Microwave Background Observables of Small Field Models of Inflation,” *JCAP* **1009** (2010) 007, [arXiv:0907.2384 \[astro-ph.CO\]](#).
- [44] S. Hotchkiss, A. Mazumdar, and S. Nadathur, “Observable gravitational waves from inflation with small field excursions,” *JCAP* **1202** (2012) 008, [arXiv:1110.5389 \[astro-ph.CO\]](#).
- [45] A. Chatterjee and A. Mazumdar, “Bound on largest $r \lesssim 0.1$ from sub-Planckian excursions of inflaton,” *JCAP* **1501** no. 01, (2015) 031, [arXiv:1409.4442 \[astro-ph.CO\]](#).
- [46] A. Ashoorioon, K. Dimopoulos, M. M. Sheikh-Jabbari, and G. Shiu, “Non-Bunch?Davis initial state reconciles chaotic models with BICEP and Planck,” *Phys. Lett.* **B737** (2014) 98–102, [arXiv:1403.6099 \[hep-th\]](#).
- [47] H. Collins, R. Holman, and T. Vardanyan, “Do Mixed States save Effective Field Theory from BICEP?,” [arXiv:1403.4592 \[hep-th\]](#).
- [48] A. Aravind, D. Lorshbough, and S. Paban, “Bogoliubov Excited States and the Lyth Bound,” *JCAP* **1408** (2014) 058, [arXiv:1403.6216 \[astro-ph.CO\]](#).
- [49] R. Flauger, D. Green, and R. A. Porto, “On squeezed limits in single-field inflation. Part I,” *JCAP* **1308** (2013) 032, [arXiv:1303.1430 \[hep-th\]](#).
- [50] **BICEP2, Planck** Collaboration, P. Ade *et al.*, “Joint Analysis of BICEP2/*KeckArray* and *Planck* Data,” *Phys. Rev. Lett.* **114** (2015) 101301, [arXiv:1502.00612 \[astro-ph.CO\]](#).
- [51] V. Mukhanov, “Quantum Cosmological Perturbations: Predictions and Observations,” *Eur. Phys. J.* **C73** (2013) 2486, [arXiv:1303.3925 \[astro-ph.CO\]](#).
- [52] D. Roest, “Universality classes of inflation,” *JCAP* **1401** (2014) 007, [arXiv:1309.1285 \[hep-th\]](#).
- [53] P. Creminelli, S. Dubovsky, D. Lopez Nacir, M. Simonovic, G. Trevisan, G. Villadoro, and M. Zaldarriaga, “Implications of the scalar tilt for the tensor-to-scalar ratio,” *Phys. Rev.* **D92** no. 12, (2015) 123528, [arXiv:1412.0678 \[astro-ph.CO\]](#).
- [54] A. D. Linde, “Chaotic Inflation,” *Phys. Lett.* **B129** (1983) 177–181.

- [55] R. Flauger, L. McAllister, E. Pajer, A. Westphal, and G. Xu, “Oscillations in the CMB from Axion Monodromy Inflation,” *JCAP* **1006** (2010) 009, [arXiv:0907.2916 \[hep-th\]](#).
- [56] A. A. Starobinsky, “A New Type of Isotropic Cosmological Models Without Singularity,” *Phys.Lett.* **B91** (1980) 99–102.
- [57] D. Salopek, J. Bond, and J. M. Bardeen, “Designing Density Fluctuation Spectra in Inflation,” *Phys.Rev.* **D40** (1989) 1753.
- [58] F. L. Bezrukov and M. Shaposhnikov, “The Standard Model Higgs boson as the inflaton,” *Phys. Lett.* **B659** (2008) 703–706, [arXiv:0710.3755 \[hep-th\]](#).
- [59] A. B. Goncharov and A. D. Linde, “Chaotic Inflation in Supergravity,” *Phys. Lett.* **B139** (1984) 27.
- [60] R. Kallosh and A. Linde, “Universality Class in Conformal Inflation,” *JCAP* **1307** (2013) 002, [arXiv:1306.5220 \[hep-th\]](#).
- [61] J. L. Cook and L. Sorbo, “Particle production during inflation and gravitational waves detectable by ground-based interferometers,” *Phys. Rev.* **D85** (2012) 023534, [arXiv:1109.0022 \[astro-ph.CO\]](#). [Erratum: Phys. Rev.D86,069901(2012)].
- [62] L. Senatore, E. Silverstein, and M. Zaldarriaga, “New Sources of Gravitational Waves during Inflation,” *JCAP* **1408** (2014) 016, [arXiv:1109.0542 \[hep-th\]](#).
- [63] N. Barnaby, J. Moxon, R. Namba, M. Peloso, G. Shiu, and P. Zhou, “Gravity waves and non-Gaussian features from particle production in a sector gravitationally coupled to the inflaton,” *Phys. Rev.* **D86** (2012) 103508, [arXiv:1206.6117 \[astro-ph.CO\]](#).
- [64] P. D. Meerburg and E. Pajer, “Observational Constraints on Gauge Field Production in Axion Inflation,” *JCAP* **1302** (2013) 017, [arXiv:1203.6076 \[astro-ph.CO\]](#).
- [65] R. Z. Ferreira and M. S. Sloth, “Universal Constraints on Axions from Inflation,” *JHEP* **12** (2014) 139, [arXiv:1409.5799 \[hep-ph\]](#).
- [66] M. Mirbabayi, L. Senatore, E. Silverstein, and M. Zaldarriaga, “Gravitational Waves and the Scale of Inflation,” *Phys. Rev.* **D91** (2015) 063518, [arXiv:1412.0665 \[hep-th\]](#).
- [67] O. Özsoy, K. Sinha, and S. Watson, “How Well Can We Really Determine the Scale of Inflation?,” *Phys. Rev.* **D91** no. 10, (2015) 103509, [arXiv:1410.0016 \[hep-th\]](#).
- [68] M. Peloso, L. Sorbo, and C. Unal, “Rolling axions during inflation: perturbativity and signatures,” [arXiv:1606.00459 \[astro-ph.CO\]](#).
- [69] R. Namba, M. Peloso, M. Shiraishi, L. Sorbo, and C. Unal, “Scale-dependent gravitational waves from a rolling axion,” *JCAP* **1601** no. 01, (2016) 041, [arXiv:1509.07521 \[astro-ph.CO\]](#).
- [70] M. M. Anber and L. Sorbo, “N-flationary magnetic fields,” *JCAP* **0610** (2006) 018, [arXiv:astro-ph/0606534 \[astro-ph\]](#).
- [71] L. Sorbo, “Parity violation in the Cosmic Microwave Background from a pseudoscalar inflaton,” *JCAP* **1106** (2011) 003, [arXiv:1101.1525 \[astro-ph.CO\]](#).
- [72] A. Maleknejad and M. M. Sheikh-Jabbari, “Gauge-flation: Inflation From Non-Abelian Gauge Fields,” *Phys. Lett.* **B723** (2013) 224–228, [arXiv:1102.1513 \[hep-ph\]](#).

- [73] P. Adshead and M. Wyman, “Chromo-Natural Inflation: Natural inflation on a steep potential with classical non-Abelian gauge fields,” *Phys. Rev. Lett.* **108** (2012) 261302, [arXiv:1202.2366 \[hep-th\]](#).
- [74] P. Adshead and M. Wyman, “Gauge-flation trajectories in Chromo-Natural Inflation,” *Phys. Rev.* **D86** (2012) 043530, [arXiv:1203.2264 \[hep-th\]](#).
- [75] P. Adshead, E. Martinec, and M. Wyman, “Gauge fields and inflation: Chiral gravitational waves, fluctuations, and the Lyth bound,” *Phys. Rev.* **D88** no. 2, (2013) 021302, [arXiv:1301.2598 \[hep-th\]](#).
- [76] P. Adshead, E. Martinec, and M. Wyman, “Perturbations in Chromo-Natural Inflation,” *JHEP* **09** (2013) 087, [arXiv:1305.2930 \[hep-th\]](#).
- [77] E. Dimastrogiovanni, M. Fasiello, and A. J. Tolley, “Low-Energy Effective Field Theory for Chromo-Natural Inflation,” *JCAP* **1302** (2013) 046, [arXiv:1211.1396 \[hep-th\]](#).
- [78] E. Dimastrogiovanni and M. Peloso, “Stability analysis of chromo-natural inflation and possible evasion of Lyth’s bound,” *Phys. Rev.* **D87** no. 10, (2013) 103501, [arXiv:1212.5184 \[astro-ph.CO\]](#).
- [79] A. Lue, L.-M. Wang, and M. Kamionkowski, “Cosmological signature of new parity violating interactions,” *Phys. Rev. Lett.* **83** (1999) 1506–1509, [arXiv:astro-ph/9812088 \[astro-ph\]](#).
- [80] V. Gluscevic and M. Kamionkowski, “Testing Parity-Violating Mechanisms with Cosmic Microwave Background Experiments,” *Phys. Rev.* **D81** (2010) 123529, [arXiv:1002.1308 \[astro-ph.CO\]](#).
- [81] J. Bielefeld and R. R. Caldwell, “Chiral Imprint of a Cosmic Gauge Field on Primordial Gravitational Waves,” *Phys. Rev.* **D91** no. 12, (2015) 123501, [arXiv:1412.6104 \[astro-ph.CO\]](#).
- [82] J. Bielefeld and R. R. Caldwell, “Cosmological consequences of classical flavor-space locked gauge field radiation,” *Phys. Rev.* **D91** no. 12, (2015) 124004, [arXiv:1503.05222 \[gr-qc\]](#).
- [83] R. Namba, E. Dimastrogiovanni, and M. Peloso, “Gauge-flation confronted with Planck,” *JCAP* **1311** (2013) 045, [arXiv:1308.1366 \[astro-ph.CO\]](#).
- [84] M. B. Hindmarsh and T. W. B. Kibble, “Cosmic strings,” *Rept. Prog. Phys.* **58** (1995) 477–562, [arXiv:hep-ph/9411342 \[hep-ph\]](#).
- [85] A. Vilenkin, “Cosmological Density Fluctuations Produced by Vacuum Strings,” *Phys. Rev. Lett.* **46** (1981) 1169–1172. [Erratum: *Phys. Rev. Lett.* 46, 1496 (1981)].
- [86] L. Kofman, A. D. Linde, and A. A. Starobinsky, “Nonthermal phase transitions after inflation,” *Phys. Rev. Lett.* **76** (1996) 1011–1014, [arXiv:hep-th/9510119 \[hep-th\]](#).
- [87] I. Tkachev, S. Khlebnikov, L. Kofman, and A. D. Linde, “Cosmic strings from preheating,” *Phys. Lett.* **B440** (1998) 262–268, [arXiv:hep-ph/9805209 \[hep-ph\]](#).
- [88] R. Jeannerot, “A Supersymmetric SO(10) model with inflation and cosmic strings,” *Phys. Rev.* **D53** (1996) 5426–5436, [arXiv:hep-ph/9509365 \[hep-ph\]](#).
- [89] R. Jeannerot, J. Rocher, and M. Sakellariadou, “How generic is cosmic string formation in SUSY GUTs,” *Phys. Rev.* **D68** (2003) 103514, [arXiv:hep-ph/0308134 \[hep-ph\]](#).
- [90] J. Rocher and M. Sakellariadou, “D-term inflation, cosmic strings, and consistency with cosmic microwave background measurement,” *Phys. Rev. Lett.* **94** (2005) 011303, [arXiv:hep-ph/0412143 \[hep-ph\]](#).

- [91] S. Sarangi and S. H. H. Tye, “Cosmic string production towards the end of brane inflation,” *Phys. Lett.* **B536** (2002) 185–192, [arXiv:hep-th/0204074 \[hep-th\]](#).
- [92] N. T. Jones, H. Stoica, and S. H. H. Tye, “The Production, spectrum and evolution of cosmic strings in brane inflation,” *Phys. Lett.* **B563** (2003) 6–14, [arXiv:hep-th/0303269 \[hep-th\]](#).
- [93] E. J. Copeland, R. C. Myers, and J. Polchinski, “Cosmic F and D strings,” *JHEP* **06** (2004) 013, [arXiv:hep-th/0312067 \[hep-th\]](#).
- [94] J. Urrestilla, P. Mukherjee, A. R. Liddle, N. Bevis, M. Hindmarsh, and M. Kunz, “Degeneracy between primordial tensor modes and cosmic strings in future CMB data from the Planck satellite,” *Phys. Rev.* **D77** (2008) 123005, [arXiv:0803.2059 \[astro-ph\]](#).
- [95] U. Seljak and A. Slosar, “B polarization of cosmic microwave background as a tracer of strings,” *Phys. Rev.* **D74** (2006) 063523, [arXiv:astro-ph/0604143 \[astro-ph\]](#).
- [96] A. Avgoustidis, E. J. Copeland, A. Moss, L. Pogosian, A. Pourtsidou, and D. A. Steer, “Constraints on the fundamental string coupling from B-mode experiments,” *Phys. Rev. Lett.* **107** (2011) 121301, [arXiv:1105.6198 \[astro-ph.CO\]](#).
- [97] **Planck** Collaboration, P. A. R. Ade *et al.*, “Planck 2015 results. XVII. Constraints on primordial non-Gaussianity,” [arXiv:1502.01592 \[astro-ph.CO\]](#).
- [98] **Planck** Collaboration, P. A. R. Ade *et al.*, “Planck 2013 results. XXV. Searches for cosmic strings and other topological defects,” *Astron. Astrophys.* **571** (2014) A25, [arXiv:1303.5085 \[astro-ph.CO\]](#).
- [99] **NANOGrav** Collaboration, Z. Arzoumanian *et al.*, “The NANOGrav Nine-year Data Set: Limits on the Isotropic Stochastic Gravitational Wave Background,” [arXiv:1508.03024 \[astro-ph.GA\]](#).
- [100] J. Urrestilla, N. Bevis, M. Hindmarsh, M. Kunz, and A. R. Liddle, “Cosmic microwave anisotropies from BPS semilocal strings,” *JCAP* **0807** (2008) 010, [arXiv:0711.1842 \[astro-ph\]](#).
- [101] L. M. Krauss, “Gravitational waves from global phase transitions,” *Phys. Lett.* **B284** (1992) 229–233.
- [102] K. Jones-Smith, L. M. Krauss, and H. Mathur, “A Nearly Scale Invariant Spectrum of Gravitational Radiation from Global Phase Transitions,” *Phys. Rev. Lett.* **100** (2008) 131302, [arXiv:0712.0778 \[astro-ph\]](#).
- [103] D. Baumann and M. Zaldarriaga, “Causality and Primordial Tensor Modes,” *JCAP* **0906** (2009) 013, [arXiv:0901.0958 \[astro-ph.CO\]](#).
- [104] H. Lee, S. C. Su, and D. Baumann, “The Superhorizon Test of Future B-mode Experiments,” *JCAP* **1502** no. 02, (2015) 036, [arXiv:1408.6709 \[astro-ph.CO\]](#).
- [105] L. C. Price, H. V. Peiris, J. Frazer, and R. Easther, “Gravitational wave consistency relations for multifield inflation,” *Phys. Rev. Lett.* **114** no. 3, (2015) 031301, [arXiv:1409.2498 \[astro-ph.CO\]](#).
- [106] D. Baumann, H. Lee, and G. L. Pimentel, “High-Scale Inflation and the Tensor Tilt,” *JHEP* **01** (2016) 101, [arXiv:1507.07250 \[hep-th\]](#).
- [107] P. D. Meerburg, R. Hloæek, B. Hadzhiyska, and J. Meyers, “Multiwavelength constraints on the inflationary consistency relation,” *Phys. Rev.* **D91** no. 10, (2015) 103505, [arXiv:1502.00302 \[astro-ph.CO\]](#).
- [108] P. D. Lasky *et al.*, “Gravitational-wave cosmology across 29 decades in frequency,” *Phys. Rev.* **X6** no. 1, (2016) 011035, [arXiv:1511.05994 \[astro-ph.CO\]](#).

- [109] D. Baumann, D. Green, and R. A. Porto, “B-modes and the Nature of Inflation,” *JCAP* **1501** no. 01, (2015) 016, [arXiv:1407.2621 \[hep-th\]](#).
- [110] S. Renaux-Petel and K. Turzynski, “On reaching the adiabatic limit in multi-field inflation,” *JCAP* **1506** no. 06, (2015) 010, [arXiv:1405.6195 \[astro-ph.CO\]](#).
- [111] S. Alexander and J. Martin, “Birefringent gravitational waves and the consistency check of inflation,” *Phys. Rev.* **D71** (2005) 063526, [arXiv:hep-th/0410230 \[hep-th\]](#).
- [112] C. R. Contaldi, J. Magueijo, and L. Smolin, “Anomalous CMB polarization and gravitational chirality,” *Phys. Rev. Lett.* **101** (2008) 141101, [arXiv:0806.3082 \[astro-ph\]](#).
- [113] T. Takahashi and J. Soda, “Chiral Primordial Gravitational Waves from a Lifshitz Point,” *Phys. Rev. Lett.* **102** (2009) 231301, [arXiv:0904.0554 \[hep-th\]](#).
- [114] J. M. Maldacena and G. L. Pimentel, “On graviton non-Gaussianities during inflation,” *JHEP* **09** (2011) 045, [arXiv:1104.2846 \[hep-th\]](#).
- [115] J. M. Maldacena, “Non-Gaussian features of primordial fluctuations in single field inflationary models,” *JHEP* **05** (2003) 013, [arXiv:astro-ph/0210603 \[astro-ph\]](#).
- [116] S. Endlich, A. Nicolis, and J. Wang, “Solid Inflation,” *JCAP* **1310** (2013) 011, [arXiv:1210.0569 \[hep-th\]](#).
- [117] L. Bordin, P. Creminelli, M. Mirbabayi, and J. Noreña, “Tensor Squeezed Limits and the Higuchi Bound,” [arXiv:1605.08424 \[astro-ph.CO\]](#).
- [118] E. Dimastrogiovanni, M. Fasiello, and M. Kamionkowski, “Imprints of Massive Primordial Fields on Large-Scale Structure,” *JCAP* **1602** (2016) 017, [arXiv:1504.05993 \[astro-ph.CO\]](#).
- [119] P. D. Meerburg, J. Meyers, A. van Engelen, and Y. Ali-Haïmoud, “CMB B-mode non-Gaussianity,” *Phys. Rev.* **D93** no. 12, (2016) 123511, [arXiv:1603.02243 \[astro-ph.CO\]](#).
- [120] R. H. Brandenberger and C. Vafa, “Superstrings in the Early Universe,” *Nucl. Phys.* **B316** (1989) 391–410.
- [121] A. A. Tseytlin and C. Vafa, “Elements of string cosmology,” *Nucl. Phys.* **B372** (1992) 443–466, [arXiv:hep-th/9109048 \[hep-th\]](#).
- [122] T. Battefeld and S. Watson, “String gas cosmology,” *Rev. Mod. Phys.* **78** (2006) 435–454, [arXiv:hep-th/0510022 \[hep-th\]](#).
- [123] D. Baumann *et al.*, “CMBPol Mission Concept Study: Probing Inflation with CMB Polarization,” *AIP Conf. Proc.* **1141** (2009) 10–120, [arXiv:0811.3919 \[astro-ph\]](#).
- [124] P. Creminelli, M. A. Luty, A. Nicolis, and L. Senatore, “Starting the Universe: Stable Violation of the Null Energy Condition and Non-standard Cosmologies,” *JHEP* **0612** (2006) 080, [arXiv:hep-th/0606090 \[hep-th\]](#).
- [125] E. I. Buchbinder, J. Khoury, and B. A. Ovrut, “New Ekpyrotic cosmology,” *Phys. Rev.* **D76** (2007) 123503, [arXiv:hep-th/0702154 \[hep-th\]](#).
- [126] P. Creminelli and L. Senatore, “A Smooth bouncing cosmology with scale invariant spectrum,” *JCAP* **0711** (2007) 010, [arXiv:hep-th/0702165 \[hep-th\]](#).

- [127] R. Kallosh, J. U. Kang, A. D. Linde, and V. Mukhanov, “The New ekpyrotic ghost,” *JCAP* **0804** (2008) 018, [arXiv:0712.2040 \[hep-th\]](#).
- [128] R. Brustein and G. Veneziano, “The Graceful exit problem in string cosmology,” *Phys. Lett.* **B329** (1994) 429–434, [arXiv:hep-th/9403060 \[hep-th\]](#).
- [129] N. Kaloper, R. Madden, and K. A. Olive, “Axions and the graceful exit problem in string cosmology,” *Phys. Lett.* **B371** (1996) 34–40, [arXiv:hep-th/9510117 \[hep-th\]](#).
- [130] N. Kaloper, R. Madden, and K. A. Olive, “Towards a singularity - free inflationary universe?,” *Nucl. Phys.* **B452** (1995) 677–704, [arXiv:hep-th/9506027 \[hep-th\]](#).
- [131] N. Kaloper and S. Watson, “Geometric precipices in string cosmology,” *Phys. Rev.* **D77** (2008) 066002, [arXiv:0712.1820 \[hep-th\]](#).
- [132] J. Quintin, Z. Sherkatghanad, Y.-F. Cai, and R. H. Brandenberger, “Evolution of cosmological perturbations and the production of non-Gaussianities through a nonsingular bounce: Indications for a no-go theorem in single field matter bounce cosmologies,” *Phys. Rev.* **D92** no. 6, (2015) 063532, [arXiv:1508.04141 \[hep-th\]](#).
- [133] I. Ben-Dayan, “Gravitational Waves in Bouncing Cosmologies from Gauge Field Production,” [arXiv:1604.07899 \[astro-ph.CO\]](#).
- [134] S. L. Dubovsky, “Phases of massive gravity,” *JHEP* **10** (2004) 076, [arXiv:hep-th/0409124 \[hep-th\]](#).
- [135] K. Hinterbichler, “Theoretical Aspects of Massive Gravity,” *Rev. Mod. Phys.* **84** (2012) 671–710, [arXiv:1105.3735 \[hep-th\]](#).
- [136] S. Dubovsky, R. Flauger, A. Starobinsky, and I. Tkachev, “Signatures of a Graviton Mass in the Cosmic Microwave Background,” *Phys. Rev.* **D81** (2010) 023523, [arXiv:0907.1658 \[astro-ph.CO\]](#).
- [137] L. S. Finn and P. J. Sutton, “Bounding the mass of the graviton using binary pulsar observations,” *Phys. Rev.* **D65** (2002) 044022, [arXiv:gr-qc/0109049 \[gr-qc\]](#).
- [138] P. D. Meerburg, “Alleviating the tension at low ℓ through axion monodromy,” *Phys. Rev.* **D90** no. 6, (2014) 063529, [arXiv:1406.3243 \[astro-ph.CO\]](#).
- [139] R. Flauger, L. McAllister, E. Silverstein, and A. Westphal, “Drifting Oscillations in Axion Monodromy,” [arXiv:1412.1814 \[hep-th\]](#).
- [140] S. R. Behbahani, A. Dymarsky, M. Mirbabayi, and L. Senatore, “(Small) Resonant non-Gaussianities: Signatures of a Discrete Shift Symmetry in the Effective Field Theory of Inflation,” *JCAP* **1212** (2012) 036, [arXiv:1111.3373 \[hep-th\]](#).
- [141] S. R. Behbahani and D. Green, “Collective Symmetry Breaking and Resonant Non-Gaussianity,” *JCAP* **1211** (2012) 056, [arXiv:1207.2779 \[hep-th\]](#).
- [142] A. Achucarro, J.-O. Gong, S. Hardeman, G. A. Palma, and S. P. Patil, “Features of heavy physics in the CMB power spectrum,” *JCAP* **1101** (2011) 030, [arXiv:1010.3693 \[hep-ph\]](#).
- [143] P. D. Meerburg, J. P. van der Schaar, and P. S. Corasaniti, “Signatures of Initial State Modifications on Bispectrum Statistics,” *JCAP* **0905** (2009) 018, [arXiv:0901.4044 \[hep-th\]](#).
- [144] P. D. Meerburg and M. Münchmeyer, “Optimal CMB estimators for bispectra from excited states,” *Phys. Rev.* **D92** no. 6, (2015) 063527, [arXiv:1505.05882 \[astro-ph.CO\]](#).

- [145] R. Flauger and E. Pajer, “Resonant Non-Gaussianity,” *JCAP* **1101** (2011) 017, [arXiv:1002.0833 \[hep-th\]](#).
- [146] P. D. Meerburg, M. Münchmeyer, and B. Wandelt, “Joint resonant CMB power spectrum and bispectrum estimation,” *Phys. Rev.* **D93** no. 4, (2016) 043536, [arXiv:1510.01756 \[astro-ph.CO\]](#).
- [147] J. R. Fergusson, H. F. Gruetjen, E. P. S. Shellard, and M. Liguori, “Combining power spectrum and bispectrum measurements to detect oscillatory features,” *Phys. Rev.* **D91** no. 2, (2015) 023502, [arXiv:1410.5114 \[astro-ph.CO\]](#).
- [148] N. Arkani-Hamed and J. Maldacena, “Cosmological Collider Physics,” [arXiv:1503.08043 \[hep-th\]](#).
- [149] X. Chen, M. H. Namjoo, and Y. Wang, “Quantum Primordial Standard Clocks,” *JCAP* **1602** no. 02, (2016) 013, [arXiv:1509.03930 \[astro-ph.CO\]](#).
- [150] P. Creminelli and M. Zaldarriaga, “Single-Field Consistency Relation for the 3-Point Function,” *JCAP* **0410** (2004) 006, [arXiv:astro-ph/0407059 \[astro-ph\]](#).
- [151] E. Nelson and S. Shandera, “Statistical Naturalness and non-Gaussianity in a Finite Universe,” *Phys. Rev. Lett.* **110** no. 13, (2013) 131301, [arXiv:1212.4550 \[astro-ph.CO\]](#).
- [152] M. LoVerde, E. Nelson, and S. Shandera, “Non-Gaussian Mode Coupling and the Statistical Cosmological Principle,” *JCAP* **1306** (2013) 024, [arXiv:1303.3549 \[astro-ph.CO\]](#).
- [153] S. Nurmi, C. T. Byrnes, and G. Tasinato, “A non-Gaussian landscape,” *JCAP* **1306** (2013) 004, [arXiv:1301.3128 \[astro-ph.CO\]](#).
- [154] B. Bonga, S. Brahma, A.-S. Deutsch, and S. Shandera, “Cosmic variance in inflation with two light scalars,” [arXiv:1512.05365 \[astro-ph.CO\]](#).
- [155] M. Zaldarriaga, “Non-Gaussianities in models with a varying inflaton decay rate,” *Phys. Rev.* **D69** (2004) 043508, [arXiv:astro-ph/0306006 \[astro-ph\]](#).
- [156] J.-L. Lehners and S. Renaux-Petel, “Multifield Cosmological Perturbations at Third Order and the Ekpyrotic Trispectrum,” *Phys. Rev.* **D80** (2009) 063503, [arXiv:0906.0530 \[hep-th\]](#).
- [157] E. Komatsu and D. N. Spergel, “Acoustic signatures in the primary microwave background bispectrum,” *Phys. Rev.* **D63** (2001) 063002, [arXiv:astro-ph/0005036 \[astro-ph\]](#).
- [158] M. Alvarez *et al.*, “Testing Inflation with Large Scale Structure: Connecting Hopes with Reality,” [arXiv:1412.4671 \[astro-ph.CO\]](#).
- [159] D. H. Lyth and D. Wands, “Generating the curvature perturbation without an inflaton,” *Phys. Lett.* **B524** (2002) 5–14, [arXiv:hep-ph/0110002 \[hep-ph\]](#).
- [160] O. Dorøsø *et al.*, “Cosmology with the SPHEREX All-Sky Spectral Survey,” [arXiv:1412.4872 \[astro-ph.CO\]](#).
- [161] D. Babich, P. Creminelli, and M. Zaldarriaga, “The Shape of Non-Gaussianities,” *JCAP* **0408** (2004) 009, [arXiv:astro-ph/0405356 \[astro-ph\]](#).
- [162] L. Senatore, K. M. Smith, and M. Zaldarriaga, “Non-Gaussianities in Single Field Inflation and their Optimal Limits from the WMAP 5-year Data,” *JCAP* **1001** (2010) 028, [arXiv:0905.3746 \[astro-ph.CO\]](#).

- [163] E. Silverstein and D. Tong, “Scalar speed limits and cosmology: Acceleration from D-cceleration,” *Phys. Rev.* **D70** (2004) 103505, [arXiv:hep-th/0310221 \[hep-th\]](#).
- [164] N. Arkani-Hamed, P. Creminelli, S. Mukohyama, and M. Zaldarriaga, “Ghost inflation,” *JCAP* **0404** (2004) 001, [arXiv:hep-th/0312100 \[hep-th\]](#).
- [165] M. Alishahiha, E. Silverstein, and D. Tong, “DBI in the sky,” *Phys. Rev.* **D70** (2004) 123505, [arXiv:hep-th/0404084 \[hep-th\]](#).
- [166] X. Chen, M.-x. Huang, S. Kachru, and G. Shiu, “Observational signatures and non-Gaussianities of general single field inflation,” *JCAP* **0701** (2007) 002, [arXiv:hep-th/0605045 \[hep-th\]](#).
- [167] C. Cheung, P. Creminelli, A. L. Fitzpatrick, J. Kaplan, and L. Senatore, “The Effective Field Theory of Inflation,” *JHEP* **03** (2008) 014, [arXiv:0709.0293 \[hep-th\]](#).
- [168] X. Chen and Y. Wang, “Quasi-Single Field Inflation and Non-Gaussianities,” *JCAP* **1004** (2010) 027, [arXiv:0911.3380 \[hep-th\]](#).
- [169] A. J. Tolley and M. Wyman, “The Gelaton Scenario: Equilateral non-Gaussianity from multi-field dynamics,” *Phys. Rev.* **D81** (2010) 043502, [arXiv:0910.1853 \[hep-th\]](#).
- [170] S. Cremonini, Z. Lalak, and K. Turzynski, “Strongly Coupled Perturbations in Two-Field Inflationary Models,” *JCAP* **1103** (2011) 016, [arXiv:1010.3021 \[hep-th\]](#).
- [171] D. Baumann and D. Green, “Signatures of Supersymmetry from the Early Universe,” *Phys. Rev.* **D85** (2012) 103520, [arXiv:1109.0292 \[hep-th\]](#).
- [172] N. Barnaby and S. Shandera, “Feeding your Inflaton: Non-Gaussian Signatures of Interaction Structure,” *JCAP* **1201** (2012) 034, [arXiv:1109.2985 \[astro-ph.CO\]](#).
- [173] P. Creminelli, “On non-Gaussianities in single-field inflation,” *JCAP* **0310** (2003) 003, [arXiv:astro-ph/0306122 \[astro-ph\]](#).
- [174] D. Baumann, D. Green, H. Lee, and R. A. Porto, “Signs of Analyticity in Single-Field Inflation,” *Phys. Rev.* **D93** no. 2, (2016) 023523, [arXiv:1502.07304 \[hep-th\]](#).
- [175] D. Baumann and D. Green, “Equilateral Non-Gaussianity and New Physics on the Horizon,” *JCAP* **1109** (2011) 014, [arXiv:1102.5343 \[hep-th\]](#).
- [176] D. Green, M. Lewandowski, L. Senatore, E. Silverstein, and M. Zaldarriaga, “Anomalous Dimensions and Non-Gaussianity,” *JHEP* **10** (2013) 171, [arXiv:1301.2630 \[hep-th\]](#).
- [177] V. Assassi, D. Baumann, D. Green, and L. McAllister, “Planck-Suppressed Operators,” *JCAP* **1401** (2014) 033, [arXiv:1304.5226 \[hep-th\]](#).
- [178] **Planck** Collaboration, P. Ade *et al.*, “Planck 2013 Results. XXIV. Constraints on primordial non-Gaussianity,” [arXiv:1303.5084 \[astro-ph.CO\]](#).
- [179] J. R. Bond and G. Efstathiou, “Cosmic background radiation anisotropies in universes dominated by nonbaryonic dark matter,” *Astrophys. J.* **285** (1984) L45–L48.
- [180] H. Kodama and M. Sasaki, “Evolution of Isocurvature Perturbations. 1. Photon - Baryon Universe,” *Int. J. Mod. Phys.* **A1** (1986) 265.
- [181] H. Kodama and M. Sasaki, “Evolution of Isocurvature Perturbations. 2. Radiation Dust Universe,” *Int. J. Mod. Phys.* **A2** (1987) 491.

- [182] W. Hu and N. Sugiyama, “Toward understanding CMB anisotropies and their implications,” *Phys. Rev.* **D51** (1995) 2599–2630, [astro-ph/9411008](#).
- [183] K. Moodley, M. Bucher, J. Dunkley, P. G. Ferreira, and C. Skordis, “Constraints on isocurvature models from the WMAP first-year data,” *Phys. Rev.* **D70** (2004) 103520, [arXiv:astro-ph/0407304](#) [astro-ph].
- [184] R. Bean, J. Dunkley, and E. Pierpaoli, “Constraining Isocurvature Initial Conditions with WMAP 3-year data,” *Phys. Rev.* **D74** (2006) 063503, [astro-ph/0606685](#).
- [185] M. Bucher, K. Moodley, and N. Turok, “The General primordial cosmic perturbation,” *Phys. Rev.* **D62** (2000) 083508, [arXiv:astro-ph/9904231](#) [astro-ph].
- [186] M. Bucher, J. Dunkley, P. G. Ferreira, K. Moodley, and C. Skordis, “The Initial conditions of the universe: How much isocurvature is allowed?,” *Phys. Rev. Lett.* **93** (2004) 081301, [arXiv:astro-ph/0401417](#) [astro-ph].
- [187] WMAP Collaboration, J. Dunkley *et al.*, “Five-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Likelihoods and Parameters from the WMAP data,” *Astrophys. J. Suppl.* **180** (2009) 306–329, [arXiv:0803.0586](#) [astro-ph].
- [188] Planck Collaboration, P. A. R. Ade *et al.*, “Planck 2015 results. XX. Constraints on inflation,” [arXiv:1502.02114](#) [astro-ph.CO].
- [189] K. Enqvist, H. Kurki-Suonio, and J. Valiviita, “Limits on isocurvature fluctuations from boomerang and MAXIMA,” *Phys. Rev.* **D62** (2000) 103003, [arXiv:astro-ph/0006429](#) [astro-ph].
- [190] C. J. MacTavish *et al.*, “Cosmological parameters from the 2003 flight of BOOMERANG,” *Astrophys. J.* **647** (2006) 799–812, [arXiv:astro-ph/0507503](#) [astro-ph].
- [191] C. Gordon and A. Lewis, “Observational constraints on the curvaton model of inflation,” *Phys. Rev.* **D67** (2003) 123513, [astro-ph/0212248](#).
- [192] A. Lewis and A. Challinor, “Evolution of cosmological dark matter perturbations,” *Phys. Rev.* **D66** (2002) 023531, [arXiv:astro-ph/0203507](#) [astro-ph].
- [193] A. Lewis and A. Challinor, “The 21cm angular-power spectrum from the dark ages,” *Phys. Rev.* **D76** (2007) 083005, [arXiv:astro-ph/0702600](#) [ASTRO-PH].
- [194] C. Gordon and J. R. Pritchard, “Forecasted 21 cm constraints on compensated isocurvature perturbations,” *Phys. Rev.* **D80** (2009) 063535, [arXiv:0907.5400](#) [astro-ph.CO].
- [195] S. Mollerach, “Isocurvature Baryon Perturbations and Inflation,” *Phys. Rev.* **D42** (1990) 313–325.
- [196] V. F. Mukhanov, H. Feldman, and R. H. Brandenberger, “Theory of cosmological perturbations. Part 1. Classical perturbations. Part 2. Quantum theory of perturbations. Part 3. Extensions,” *Phys. Rept.* **215** (1992) 203–333.
- [197] T. Moroi and T. Takahashi, “Effects of cosmological moduli fields on cosmic microwave background,” *Phys. Lett.* **B522** (2001) 215–221, [hep-ph/0110096](#).
- [198] D. H. Lyth, C. Ungarelli, and D. Wands, “The Primordial density perturbation in the curvaton scenario,” *Phys. Rev.* **D67** (2003) 023503, [astro-ph/0208055](#).
- [199] M. Postma, “The Curvaton scenario in supersymmetric theories,” *Phys. Rev.* **D67** (2003) 063518, [hep-ph/0212005](#).

- [200] S. Kasuya, M. Kawasaki, and F. Takahashi, “MSSM curvaton in the gauge mediated SUSY breaking,” *Phys. Lett.* **B578** (2004) 259–268, [hep-ph/0305134](#).
- [201] M. Ikegami and T. Moroi, “Curvaton scenario with Affleck-Dine baryogenesis,” *Phys. Rev.* **D70** (2004) 083515, [hep-ph/0404253](#).
- [202] A. Mazumdar and A. Perez-Lorenzana, “Sneutrino condensate as a candidate for the hot big bang cosmology,” *Phys. Rev.* **D70** (2004) 083526, [hep-ph/0406154](#).
- [203] R. Allahverdi, K. Enqvist, A. Jokinen, and A. Mazumdar, “Identifying the curvaton within MSSM,” *JCAP* **0610** (2006) 007, [hep-ph/0603255](#).
- [204] E. Papantonopoulos and V. Zamarias, “Curvaton Dynamics in Brane-worlds,” *JCAP* **0611** (2006) 005, [gr-qc/0608026](#).
- [205] A. Mazumdar and J. Rocher, “Particle physics models of inflation and curvaton scenarios,” *Phys. Rept.* **497** (2011) 85–215, [arXiv:1001.0993](#).
- [206] A. Mazumdar and S. Nadathur, “The curvaton scenario in the MSSM and predictions for non-Gaussianity,” *Phys. Rev. Lett.* **108** (2012) 111302, [arXiv:1107.4078](#).
- [207] C. Gordon and K. A. Malik, “WMAP, neutrino degeneracy and non-Gaussianity constraints on isocurvature perturbations in the curvaton model of inflation,” *Phys. Rev.* **D69** (2004) 063508, [arXiv:astro-ph/0311102 \[astro-ph\]](#).
- [208] E. Di Valentino, M. Lattanzi, G. Mangano, A. Melchiorri, and P. Serpico, “Future constraints on neutrino isocurvature perturbations in the curvaton scenario,” *Phys. Rev.* **D85** (2012) 043511, [arXiv:1111.3810](#).
- [209] T. L. Smith and D. Grin, “Probing A Panoply of Curvaton-Decay Scenarios Using CMB Data,” [arXiv:1511.07431 \[astro-ph.CO\]](#).
- [210] J. B. Muñoz, D. Grin, L. Dai, M. Kamionkowski, and E. D. Kovetz, “Search for Compensated Isocurvature Perturbations with Planck Power Spectra,” *Phys. Rev.* **D93** no. 4, (2016) 043008, [arXiv:1511.04441 \[astro-ph.CO\]](#).
- [211] D. Grin, O. Doré, and M. Kamionkowski, “Do baryons trace dark matter in the early universe?,” *Phys. Rev. Lett.* **107** (2011) 261301, [arXiv:1107.1716 \[astro-ph.CO\]](#).
- [212] D. Grin, O. Doré, and M. Kamionkowski, “Compensated Isocurvature Perturbations and the Cosmic Microwave Background,” *Phys. Rev.* **D84** (2011) 123003, [arXiv:1107.5047 \[astro-ph.CO\]](#).
- [213] D. Grin, D. Hanson, G. P. Holder, O. Doré, and M. Kamionkowski, “Baryons do trace dark matter 380,000 years after the big bang: Search for compensated isocurvature perturbations with WMAP 9-year data,” *Phys. Rev.* **D89** no. 2, (2014) 023006, [arXiv:1306.4319 \[astro-ph.CO\]](#).
- [214] C. He, D. Grin, and W. Hu, “Compensated isocurvature perturbations in the curvaton model,” *Phys. Rev.* **D92** no. 6, (2015) 063018, [arXiv:1505.00639 \[astro-ph.CO\]](#).
- [215] W. Hu and T. Okamoto, “Mass reconstruction with cmb polarization,” *Astrophys.J.* **574** (2002) 566–574, [arXiv:astro-ph/0111606 \[astro-ph\]](#).
- [216] A. H. Guth and Y. Nomura, “What can the observation of nonzero curvature tell us?,” *Phys. Rev.* **D86** (2012) 023534, [arXiv:1203.6876 \[hep-th\]](#).

- [217] M. Kleban and M. Schillo, “Spatial Curvature Falsifies Eternal Inflation,” *JCAP* **1206** (2012) 029, [arXiv:1202.5037 \[astro-ph.CO\]](#).
- [218] A. Font-Ribera, P. McDonald, N. Mostek, B. A. Reid, H.-J. Seo, *et al.*, “DESI and other dark energy experiments in the era of neutrino mass measurements,” [arXiv:1308.4164 \[astro-ph.CO\]](#).
- [219] C. D. Leonard, P. Bull, and R. Allison, “The Spatial Curvature Endgame,” [arXiv:1604.01410 \[astro-ph.CO\]](#).
- [220] M. Zaldarriaga, D. N. Spergel, and U. Seljak, “Microwave background constraints on cosmological parameters,” *Astrophys. J.* **488** (1997) 1–13, [arXiv:astro-ph/9702157 \[astro-ph\]](#).
- [221] J. R. Bond, G. Efstathiou, and M. Tegmark, “Forecasting cosmic parameter errors from microwave background anisotropy experiments,” *Mon. Not. Roy. Astron. Soc.* **291** (1997) L33–L41, [arXiv:astro-ph/9702100 \[astro-ph\]](#).
- [222] J. Errard, S. M. Feeney, H. V. Peiris, and A. H. Jaffe, “Robust forecasts on fundamental physics from the foreground-obscured, gravitationally-lensed CMB polarization,” *JCAP* **1603** no. 03, (2016) 052, [arXiv:1509.06770 \[astro-ph.CO\]](#).
- [223] C. M. Hirata and U. Seljak, “Analyzing weak lensing of the cosmic microwave background using the likelihood function,” *Phys. Rev.* **D67** (2003) 043001, [arXiv:astro-ph/0209489 \[astro-ph\]](#).
- [224] K. M. Smith, D. Hanson, M. LoVerde, C. M. Hirata, and O. Zahn, “Delensing CMB Polarization with External Datasets,” *JCAP* **1206** (2012) 014, [arXiv:1010.0048 \[astro-ph.CO\]](#).
- [225] W. L. K. Wu, J. Errard, C. Dvorkin, C. L. Kuo, A. T. Lee, P. McDonald, A. Slosar, and O. Zahn, “A Guide to Designing Future Ground-based Cosmic Microwave Background Experiments,” *Astrophys. J.* **788** (2014) 138, [arXiv:1402.4108 \[astro-ph.CO\]](#).
- [226] **Topical Conveners:** K.N. Abazajian, J.E. Carlstrom, A.T. Lee Collaboration, K. N. Abazajian *et al.*, “Neutrino Physics from the Cosmic Microwave Background and Large Scale Structure,” *Astropart. Phys.* **63** (2015) 66–80, [arXiv:1309.5383 \[astro-ph.CO\]](#).
- [227] W. Hu, D. J. Eisenstein, and M. Tegmark, “Weighing neutrinos with galaxy surveys,” *Phys. Rev. Lett.* **80** (1998) 5255–5258, [arXiv:astro-ph/9712057 \[astro-ph\]](#).
- [228] J. Lesgourgues and S. Pastor, “Massive neutrinos and cosmology,” *Phys. Rept.* **429** (2006) 307–379, [arXiv:astro-ph/0603494 \[astro-ph\]](#).
- [229] K. Abazajian, E. Calabrese, A. Cooray, F. De Bernardis, S. Dodelson, *et al.*, “Cosmological and Astrophysical Neutrino Mass Measurements,” *Astropart. Phys.* **35** (2011) 177–184, [arXiv:1103.5083 \[astro-ph.CO\]](#).
- [230] **Planck** Collaboration, P. A. R. Ade *et al.*, “Planck 2015 results. XIII. Cosmological parameters,” [arXiv:1502.01589 \[astro-ph.CO\]](#).
- [231] **Particle Data Group** Collaboration, K. A. Olive *et al.*, “Review of Particle Physics,” *Chin. Phys. C38* (2014) 090001.
- [232] J. R. Bond and A. S. Szalay, “The Collisionless Damping of Density Fluctuations in an Expanding Universe,” *Astrophys. J.* **274** (1983) 443–468.
- [233] C.-P. Ma, “Linear power spectra in cold + hot dark matter models: Analytical approximations and applications,” *Astrophys. J.* **471** (1996) 13–23, [arXiv:astro-ph/9605198 \[astro-ph\]](#).

- [234] W. Hu and D. J. Eisenstein, “Small scale perturbations in a general MDM cosmology,” *Astrophys.J.* **498** (1998) 497, [arXiv:astro-ph/9710216 \[astro-ph\]](#).
- [235] M. Kaplinghat, L. Knox, and Y.-S. Song, “Determining neutrino mass from the CMB alone,” *Phys.Rev.Lett.* **91** (2003) 241301, [arXiv:astro-ph/0303344 \[astro-ph\]](#).
- [236] J. L. Tinker, A. V. Kravtsov, A. Klypin, K. Abazajian, M. S. Warren, G. Yepes, S. Gottlöber, and D. E. Holz, “Toward a halo mass function for precision cosmology: The Limits of universality,” *Astrophys. J.* **688** (2008) 709–728, [arXiv:0803.2706 \[astro-ph\]](#).
- [237] S. Bhattacharya, K. Heitmann, M. White, Z. Lukic, C. Wagner, and S. Habib, “Mass Function Predictions Beyond LCDM,” *Astrophys. J.* **732** (2011) 122, [arXiv:1005.2239 \[astro-ph.CO\]](#).
- [238] M. Costanzi, F. Villaescusa-Navarro, M. Viel, J.-Q. Xia, S. Borgani, E. Castorina, and E. Sefusatti, “Cosmology with massive neutrinos III: the halo mass function and an application to galaxy clusters,” *JCAP* **1312** (2013) 012, [arXiv:1311.1514 \[astro-ph.CO\]](#).
- [239] M. LoVerde, “Spherical collapse in $\nu\Lambda$ CDM,” *Phys. Rev.* **D90** no. 8, (2014) 083518, [arXiv:1405.4858 \[astro-ph.CO\]](#).
- [240] M. Hasselfield *et al.*, “The Atacama Cosmology Telescope: Sunyaev-Zel’dovich selected galaxyclusters at 148 GHz from three seasons of data,” *JCAP* **1307** (2013) 008, [arXiv:1301.0816 \[astro-ph.CO\]](#).
- [241] A. B. Mantz *et al.*, “Weighing the giants – IV. Cosmology and neutrino mass,” *Mon. Not. Roy. Astron. Soc.* **446** (2015) 2205–2225, [arXiv:1407.4516 \[astro-ph.CO\]](#).
- [242] **Planck** Collaboration, P. A. R. Ade *et al.*, “Planck 2015 results. XXIV. Cosmology from Sunyaev-Zeldovich cluster counts,” [arXiv:1502.01597 \[astro-ph.CO\]](#).
- [243] **SPT** Collaboration, T. de Haan *et al.*, “Cosmological Constraints from Galaxy Clusters in the 2500 square-degree SPT-SZ Survey,” *Submitted to: Astrophys. J.* (2016) , [arXiv:1603.06522 \[astro-ph.CO\]](#).
- [244] S. Wang, Z. Haiman, W. Hu, J. Khouri, and M. May, “Weighing neutrinos with galaxy cluster surveys,” *Phys.Rev.Lett.* **95** (2005) 011302, [arXiv:astro-ph/0505390 \[astro-ph\]](#).
- [245] K. Andersson *et al.*, “X-ray Properties of the First SZE-selected Galaxy Cluster Sample from the South Pole Telescope,” *Astrophys. J.* **738** (2011) 48, [arXiv:1006.3068 \[astro-ph.CO\]](#).
- [246] V. L. Corless and L. J. King, “Cosmology with the cluster mass function: mass estimators and shape systematics in large weak lensing surveys,” *Mon. Not. Roy. Astron. Soc.* **396** (2009) 315, [arXiv:0901.3434 \[astro-ph.CO\]](#).
- [247] M. R. Becker and A. V. Kravtsov, “On the Accuracy of Weak Lensing Cluster Mass Reconstructions,” *Astrophys. J.* **740** (2011) 25, [arXiv:1011.1681 \[astro-ph.CO\]](#).
- [248] D. E. Applegate, A. von der Linden, P. L. Kelly, M. T. Allen, S. W. Allen, P. R. Burchat, D. L. Burke, H. Ebeling, A. Mantz, and R. G. Morris, “Weighing the Giants ? III. Methods and measurements of accurate galaxy cluster weak-lensing masses,” *Mon. Not. Roy. Astron. Soc.* **439** no. 1, (2014) 48–72, [arXiv:1208.0605 \[astro-ph.CO\]](#).
- [249] **LSST Dark Energy Science** Collaboration, A. Abate *et al.*, “Large Synoptic Survey Telescope: Dark Energy Science Collaboration,” [arXiv:1211.0310 \[astro-ph.CO\]](#).

- [250] **ACT** Collaboration, M. Madhavacheril *et al.*, “Evidence of Lensing of the Cosmic Microwave Background by Dark Matter Halos,” *Phys. Rev. Lett.* **114** no. 15, (2015) 151302, [arXiv:1411.7999 \[astro-ph.CO\]](#). [Addendum: Phys. Rev. Lett.114,no.18,189901(2015)].
- [251] E. J. Baxter *et al.*, “A Measurement of Gravitational Lensing of the Cosmic Microwave Background by Galaxy Clusters Using Data from the South Pole Telescope,” *Astrophys. J.* **806** no. 2, (2015) 247, [arXiv:1412.7521 \[astro-ph.CO\]](#).
- [252] W. Hu, S. DeDeo, and C. Vale, “Cluster Mass Estimators from CMB Temperature and Polarization Lensing,” *New J. Phys.* **9** (2007) 441, [arXiv:astro-ph/0701276 \[astro-ph\]](#).
- [253] I. Tereno, C. Schimd, J.-P. Uzan, M. Kilbinger, F. H. Vincent, and L. Fu, “CFHTLS weak-lensing constraints on the neutrino masses,” *Astron. Astrophys.* **500** (2009) 657–665, [arXiv:0810.0555 \[astro-ph\]](#).
- [254] J.-Q. Xia, B. R. Granett, M. Viel, S. Bird, L. Guzzo, M. G. Haehnelt, J. Coupon, H. J. McCracken, and Y. Mellier, “Constraints on Massive Neutrinos from the CFHTLS Angular Power Spectrum,” *JCAP* **1206** (2012) 010, [arXiv:1203.5105 \[astro-ph.CO\]](#).
- [255] A. J. Cuesta, V. Niro, and L. Verde, “Neutrino mass limits: robust information from the power spectrum of galaxy surveys,” *Phys. Dark Univ.* **13** (2016) 77–86, [arXiv:1511.05983 \[astro-ph.CO\]](#).
- [256] N. Palanque-Delabrouille *et al.*, “Constraint on neutrino masses from SDSS-III/BOSS Ly α forest and other cosmological probes,” *JCAP* **1502** no. 02, (2015) 045, [arXiv:1410.7244 \[astro-ph.CO\]](#).
- [257] **BOSS** Collaboration, F. Beutler *et al.*, “The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: Signs of neutrino mass in current cosmological datasets,” *Mon. Not. Roy. Astron. Soc.* **444** (2014) 3501, [arXiv:1403.4599 \[astro-ph.CO\]](#).
- [258] M. LoVerde, “Halo bias in mixed dark matter cosmologies,” *Phys. Rev.* **D90** no. 8, (2014) 083530, [arXiv:1405.4855 \[astro-ph.CO\]](#).
- [259] Y. Takeuchi and K. Kadota, “Probing Neutrinos from Planck and Forthcoming Galaxy Redshift Surveys,” *JCAP* **1401** (2014) 046, [arXiv:1310.0037 \[astro-ph.CO\]](#).
- [260] R. Pearson and O. Zahn, “Cosmology from cross correlation of CMB lensing and galaxy surveys,” *Phys. Rev.* **D89** no. 4, (2014) 043516, [arXiv:1311.0905 \[astro-ph.CO\]](#).
- [261] E.-M. Mueller, F. de Bernardis, R. Bean, and M. D. Niemack, “Constraints on massive neutrinos from the pairwise kinematic Sunyaev-Zel’dovich effect,” *Phys. Rev.* **D92** no. 6, (2015) 063501, [arXiv:1412.0592 \[astro-ph.CO\]](#).
- [262] S. Das, J. Errard, and D. Spergel, “Can CMB Lensing Help Cosmic Shear Surveys?”, [arXiv:1311.2338 \[astro-ph.CO\]](#).
- [263] C. Kraus *et al.*, “Final results from phase II of the Mainz neutrino mass search in tritium beta decay,” *Eur. Phys. J.* **C40** (2005) 447–468, [arXiv:hep-ex/0412056 \[hep-ex\]](#).
- [264] **Troitsk** Collaboration, V. N. Aseev *et al.*, “An upper limit on electron antineutrino mass from Troitsk experiment,” *Phys. Rev.* **D84** (2011) 112003, [arXiv:1108.5034 \[hep-ex\]](#).
- [265] **KATRIN** Collaboration, J. Angrik *et al.*, “KATRIN design report 2004.”
- [266] **ECHo** Collaboration, S. Eliseev *et al.*, “Direct Measurement of the Mass Difference of ^{163}Ho and ^{163}Dy Solves the Q -Value Puzzle for the Neutrino Mass Determination,” *Phys. Rev. Lett.* **115** no. 6, (2015) 062501, [arXiv:1604.04210 \[physics.ins-det\]](#).

- [267] **HOLMES** Collaboration, V. Ceriale, “HOLMES: The Electron Capture Decay of ^{163}Ho to Measure the Electron Neutrino Mass with sub-eV sensitivity: TES detector and array production.,” *PoS NEUTEL2015* (2015) 066.
- [268] M. P. Croce *et al.*, “Development of holmium-163 electron-capture spectroscopy with transition-edge sensors,” [arXiv:1510.03874 \[physics.ins-det\]](#).
- [269] **Project 8** Collaboration, D. Asner *et al.*, “Single electron detection and spectroscopy via relativistic cyclotron radiation,” *Phys. Rev. Lett.* **114** no. 16, (2015) 162501, [arXiv:1408.5362 \[physics.ins-det\]](#).
- [270] **GERDA** Collaboration, M. Agostini *et al.*, “Results on Neutrinoless Double- β Decay of ^{76}Ge from Phase I of the GERDA Experiment,” *Phys. Rev. Lett.* **111** no. 12, (2013) 122503, [arXiv:1307.4720 \[nucl-ex\]](#).
- [271] **EXO-200** Collaboration, M. Auger *et al.*, “Search for Neutrinoless Double-Beta Decay in ^{136}Xe with EXO-200,” *Phys. Rev. Lett.* **109** (2012) 032505, [arXiv:1205.5608 \[hep-ex\]](#).
- [272] **CUORE** Collaboration, D. R. Artusa *et al.*, “Searching for neutrinoless double-beta decay of ^{130}Te with CUORE,” *Adv. High Energy Phys.* **2015** (2015) 879871, [arXiv:1402.6072 \[physics.ins-det\]](#).
- [273] A. Aprahamian *et al.*, “Reaching for the horizon: The 2015 long range plan for nuclear science.”.
- [274] R. B. Patterson, “Prospects for Measurement of the Neutrino Mass Hierarchy,” *Ann. Rev. Nucl. Part. Sci.* **65** (2015) 177–192, [arXiv:1506.07917 \[hep-ex\]](#).
- [275] **JUNO** Collaboration, F. An *et al.*, “Neutrino Physics with JUNO,” *J. Phys.* **G43** no. 3, (2016) 030401, [arXiv:1507.05613 \[physics.ins-det\]](#).
- [276] **Hyper-Kamiokande Proto-Collaboration** Collaboration, K. Abe *et al.*, “Physics potential of a long-baseline neutrino oscillation experiment using a J-PARC neutrino beam and Hyper-Kamiokande,” *PTEP* **2015** (2015) 053C02, [arXiv:1502.05199 \[hep-ex\]](#).
- [277] M. Goodman, “The Deep Underground Neutrino Experiment,” *Adv. High Energy Phys.* **2015** (2015) 256351.
- [278] **KM3Net** Collaboration, S. Adrian-Martinez *et al.*, “Letter of Intent for KM3NeT2.0,” [arXiv:1601.07459 \[astro-ph.IM\]](#).
- [279] **NOvA** Collaboration, P. Adamson *et al.*, “First measurement of electron neutrino appearance in NOvA,” *Phys. Rev. Lett.* **116** no. 15, (2016) 151806, [arXiv:1601.05022 \[hep-ex\]](#).
- [280] S. Weinberg, *Cosmology*. OUP Oxford, 2008.
- [281] A. D. Dolgov, “Neutrinos in cosmology,” *Phys. Rept.* **370** (2002) 333–535, [arXiv:hep-ph/0202122 \[hep-ph\]](#).
- [282] G. Mangano, G. Miele, S. Pastor, T. Pinto, O. Pisanti, *et al.*, “Relic neutrino decoupling including flavor oscillations,” *Nucl.Phys.* **B729** (2005) 221–234, [arXiv:hep-ph/0506164 \[hep-ph\]](#).
- [283] E. Grohs, G. M. Fuller, C. T. Kishimoto, M. W. Paris, and A. Vlasenko, “Neutrino energy transport in weak decoupling and big bang nucleosynthesis,” [arXiv:1512.02205 \[astro-ph.CO\]](#).
- [284] **SLD Electroweak Group, DELPHI, ALEPH, SLD, SLD Heavy Flavour Group, OPAL, LEP Electroweak Working Group, L3** Collaboration, S. Schael *et al.*, “Precision electroweak measurements on the Z resonance,” *Phys. Rept.* **427** (2006) 257–454, [arXiv:hep-ex/0509008 \[hep-ex\]](#).

- [285] K. Abazajian, G. M. Fuller, and M. Patel, “Sterile neutrino hot, warm, and cold dark matter,” *Phys. Rev.* **D64** (2001) 023501, [arXiv:astro-ph/0101524 \[astro-ph\]](#).
- [286] A. Strumia and F. Vissani, “Neutrino masses and mixings and...,” [arXiv:hep-ph/0606054 \[hep-ph\]](#).
- [287] A. Boyarsky, O. Ruchayskiy, and M. Shaposhnikov, “The Role of sterile neutrinos in cosmology and astrophysics,” *Ann. Rev. Nucl. Part. Sci.* **59** (2009) 191–214, [arXiv:0901.0011 \[hep-ph\]](#).
- [288] L. A. Boyle and A. Buonanno, “Relating gravitational wave constraints from primordial nucleosynthesis, pulsar timing, laser interferometers, and the CMB: Implications for the early Universe,” *Phys. Rev.* **D78** (2008) 043531, [arXiv:0708.2279 \[astro-ph\]](#).
- [289] A. Stewart and R. Brandenberger, “Observational Constraints on Theories with a Blue Spectrum of Tensor Modes,” *JCAP* **0808** (2008) 012, [arXiv:0711.4602 \[astro-ph\]](#).
- [290] L. Ackerman, M. R. Buckley, S. M. Carroll, and M. Kamionkowski, “Dark Matter and Dark Radiation,” *Phys. Rev. D* **79** (2009) 023519, [arXiv:0810.5126 \[hep-ph\]](#).
- [291] D. E. Kaplan, G. Z. Krnjaic, K. R. Rehermann, and C. M. Wells, “Dark Atoms: Asymmetry and Direct Detection,” *JCAP* **1110** (2011) 011, [arXiv:1105.2073 \[hep-ph\]](#).
- [292] F.-Y. Cyr-Racine and K. Sigurdson, “Cosmology of atomic dark matter,” *Phys. Rev.* **D87** no. 10, (2013) 103515, [arXiv:1209.5752 \[astro-ph.CO\]](#).
- [293] D. Cadamuro, S. Hannestad, G. Raffelt, and J. Redondo, “Cosmological bounds on sub-MeV mass axions,” *JCAP* **1102** (2011) 003, [arXiv:1011.3694 \[hep-ph\]](#).
- [294] S. Weinberg, “Goldstone Bosons as Fractional Cosmic Neutrinos,” *Phys. Rev. Lett.* **110** (2013) 241301, [arXiv:1305.1971 \[astro-ph.CO\]](#).
- [295] G. Steigman, “Equivalent Neutrinos, Light WIMPs, and the Chimera of Dark Radiation,” *Phys. Rev.* **D87** no. 10, (2013) 103517, [arXiv:1303.0049 \[astro-ph.CO\]](#).
- [296] C. Boehm, M. J. Dolan, and C. McCabe, “A Lower Bound on the Mass of Cold Thermal Dark Matter from Planck,” *JCAP* **1308** (2013) 041, [arXiv:1303.6270 \[hep-ph\]](#).
- [297] Z. Chacko, Y. Cui, S. Hong, and T. Okui, “Hidden dark matter sector, dark radiation, and the CMB,” *Phys. Rev.* **D92** (2015) 055033, [arXiv:1505.04192 \[hep-ph\]](#).
- [298] D. Baumann, D. Green, J. Meyers, and B. Wallisch, “Phases of New Physics in the CMB,” [arXiv:1508.06342 \[astro-ph.CO\]](#).
- [299] M. Zaldarriaga and D. D. Harari, “Analytic approach to the polarization of the cosmic microwave background in flat and open universes,” *Phys. Rev.* **D52** (1995) 3276–3287, [arXiv:astro-ph/9504085 \[astro-ph\]](#).
- [300] Z. Hou, R. Keisler, L. Knox, M. Millea, and C. Reichardt, “How Massless Neutrinos Affect the Cosmic Microwave Background Damping Tail,” *Phys. Rev.* **D87** (2013) 083008, [arXiv:1104.2333 \[astro-ph.CO\]](#).
- [301] S. Bashinsky and U. Seljak, “Neutrino perturbations in CMB anisotropy and matter clustering,” *Phys. Rev.* **D69** (2004) 083002, [arXiv:astro-ph/0310198 \[astro-ph\]](#).
- [302] B. Follin, L. Knox, M. Millea, and Z. Pan, “First Detection of the Acoustic Oscillation Phase Shift Expected from the Cosmic Neutrino Background,” *Phys. Rev. Lett.* **115** no. 9, (2015) 091301, [arXiv:1503.07863 \[astro-ph.CO\]](#).

- [303] R. D. Peccei and H. R. Quinn, “CP Conservation in the Presence of Instantons,” *Phys. Rev. Lett.* **38** (1977) 1440–1443.
- [304] P. W. Graham, D. E. Kaplan, and S. Rajendran, “Cosmological Relaxation of the Electroweak Scale,” *Phys. Rev. Lett.* **115** no. 22, (2015) 221801, [arXiv:1504.07551 \[hep-ph\]](#).
- [305] A. Arvanitaki, S. Dimopoulos, S. Dubovsky, N. Kaloper, and J. March-Russell, “String Axiverse,” *Phys. Rev.* **D81** (2010) 123530, [arXiv:0905.4720 \[hep-th\]](#).
- [306] P. Langacker, “Neutrino Masses from the Top Down,” *Ann. Rev. Nucl. Part. Sci.* **62** (2012) 215–235, [arXiv:1112.5992 \[hep-ph\]](#).
- [307] LSND Collaboration, C. Athanassopoulos *et al.*, “Evidence for nu/mu → nu/e neutrino oscillations from lsnd,” *Phys. Rev. Lett.* **81** (1998) 1774–1777, [nucl-ex/9709006](#).
- [308] MiniBooNE Collaboration, A. A. Aguilar-Arevalo *et al.*, “Improved Search for $\bar{\nu}_\mu \rightarrow \bar{\nu}_e$ Oscillations in the MiniBooNE Experiment,” *Phys. Rev. Lett.* **110** (2013) 161801, [arXiv:1207.4809 \[hep-ex\]](#).
- [309] G. Mention, M. Fechner, T. Lasserre, T. A. Mueller, D. Lhuillier, M. Cribier, and A. Letourneau, “The Reactor Antineutrino Anomaly,” *Phys. Rev.* **D83** (2011) 073006, [arXiv:1101.2755 \[hep-ex\]](#).
- [310] P. Huber, “On the determination of anti-neutrino spectra from nuclear reactors,” *Phys. Rev.* **C84** (2011) 024617, [arXiv:1106.0687 \[hep-ph\]](#). [Erratum: Phys. Rev.C85,029901(2012)].
- [311] J. N. Bahcall, P. I. Krastev, and E. Lisi, “Limits on electron-neutrino oscillations from the GALLEX Cr-51 source experiment,” *Phys. Lett.* **B348** (1995) 121–123, [arXiv:hep-ph/9411414 \[hep-ph\]](#).
- [312] C. Giunti and M. Laveder, “Statistical Significance of the Gallium Anomaly,” *Phys. Rev.* **C83** (2011) 065504, [arXiv:1006.3244 \[hep-ph\]](#).
- [313] J. M. Conrad, C. M. Ignarra, G. Karagiorgi, M. H. Shaevitz, and J. Spitz, “Sterile Neutrino Fits to Short Baseline Neutrino Oscillation Measurements,” *Adv. High Energy Phys.* **2013** (2013) 163897, [arXiv:1207.4765 \[hep-ex\]](#).
- [314] J. Kopp, P. A. N. Machado, M. Maltoni, and T. Schwetz, “Sterile Neutrino Oscillations: The Global Picture,” *JHEP* **05** (2013) 050, [arXiv:1303.3011 \[hep-ph\]](#).
- [315] C. Giunti and E. M. Zavanin, “Appearance–disappearance relation in $3 + N_s$ short-baseline neutrino oscillations,” *Mod. Phys. Lett.* **A31** no. 01, (2015) 1650003, [arXiv:1508.03172 \[hep-ph\]](#).
- [316] K. N. Abazajian, “Telling three from four neutrinos with cosmology,” *Astropart. Phys.* **19** (2003) 303–312, [arXiv:astro-ph/0205238 \[astro-ph\]](#).
- [317] A. de Gouvea and W.-C. Huang, “Constraining the (Low-Energy) Type-I Seesaw,” *Phys. Rev.* **D85** (2012) 053006, [arXiv:1110.6122 \[hep-ph\]](#).
- [318] A. Donini, P. Hernandez, J. Lopez-Pavon, M. Maltoni, and T. Schwetz, “The minimal 3+2 neutrino model versus oscillation anomalies,” *JHEP* **07** (2012) 161, [arXiv:1205.5230 \[hep-ph\]](#).
- [319] T. D. Jacques, L. M. Krauss, and C. Lunardini, “Additional Light Sterile Neutrinos and Cosmology,” *Phys. Rev.* **D87** no. 8, (2013) 083515, [arXiv:1301.3119 \[astro-ph.CO\]](#). [Erratum: Phys. Rev.D88,no.10,109901(2013)].
- [320] M. Wyman, D. H. Rudd, R. A. Vanderveld, and W. Hu, “Neutrinos Help Reconcile Planck Measurements with the Local Universe,” *Phys. Rev. Lett.* **112** no. 5, (2014) 051302, [arXiv:1307.7715 \[astro-ph.CO\]](#).

- [321] R. A. Battye and A. Moss, “Evidence for Massive Neutrinos from Cosmic Microwave Background and Lensing Observations,” *Phys. Rev. Lett.* **112** no. 5, (2014) 051303, [arXiv:1308.5870 \[astro-ph.CO\]](#).
- [322] E. Giusarma, E. Di Valentino, M. Lattanzi, A. Melchiorri, and O. Mena, “Relic Neutrinos, thermal axions and cosmology in early 2014,” *Phys. Rev.* **D90** no. 4, (2014) 043507, [arXiv:1403.4852 \[astro-ph.CO\]](#).
- [323] S. Joudaki, K. N. Abazajian, and M. Kaplinghat, “Are Light Sterile Neutrinos Preferred or Disfavored by Cosmology?,” *Phys. Rev.* **D87** (2013) 065003, [arXiv:1208.4354 \[astro-ph.CO\]](#).
- [324] M. Archidiacono, N. Fornengo, C. Giunti, S. Hannestad, and A. Melchiorri, “Sterile neutrinos: Cosmology versus short-baseline experiments,” *Phys. Rev.* **D87** no. 12, (2013) 125034, [arXiv:1302.6720 \[astro-ph.CO\]](#).
- [325] K. N. Abazajian *et al.*, “Light Sterile Neutrinos: A White Paper,” [arXiv:1204.5379 \[hep-ph\]](#).
- [326] C. Brust, D. E. Kaplan, and M. T. Walters, “New Light Species and the CMB,” *JHEP* **12** (2013) 058, [arXiv:1303.5379 \[hep-ph\]](#).
- [327] K. Osato, T. Sekiguchi, M. Shirasaki, A. Kamada, and N. Yoshida, “Cosmological Constraint on the Light Gravitino Mass from CMB Lensing and Cosmic Shear,” [arXiv:1601.07386 \[astro-ph.CO\]](#).
- [328] L. A. Boyle, P. J. Steinhardt, and N. Turok, “The Cosmic gravitational wave background in a cyclic universe,” *Phys. Rev.* **D69** (2004) 127302, [arXiv:hep-th/0307170 \[hep-th\]](#).
- [329] R. A. Isaacson, “Gravitational Radiation in the Limit of High Frequency. II. Nonlinear Terms and the Effective Stress Tensor,” *Phys. Rev.* **166** (1968) 1272–1279.
- [330] C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation*. W. H. Freeman, San Francisco, 1973.
- [331] Y. Watanabe and E. Komatsu, “Improved Calculation of the Primordial Gravitational Wave Spectrum in the Standard Model,” *Phys. Rev.* **D73** (2006) 123515, [arXiv:astro-ph/0604176 \[astro-ph\]](#).
- [332] M. Maggiore, *Gravitational Waves. Vol. 1: Theory and Experiments*. Oxford Master Series in Physics. Oxford University Press, 2007. <http://www.oup.com/uk/catalogue/?ci=9780198570745>.
- [333] R. A. Alpher, H. Bethe, and G. Gamow, “The origin of chemical elements,” *Phys. Rev.* **73** (1948) 803–804.
- [334] R. H. Cyburt, B. D. Fields, K. A. Olive, and T.-H. Yeh, “Big Bang Nucleosynthesis: 2015,” [arXiv:1505.01076 \[astro-ph.CO\]](#).
- [335] R. V. Wagoner, W. A. Fowler, and F. Hoyle, “On the Synthesis of elements at very high temperatures,” *Astrophys. J.* **148** (1967) 3–49.
- [336] R. H. Cyburt, B. D. Fields, and K. A. Olive, “The NACRE thermonuclear reaction compilation and big bang nucleosynthesis,” *New Astron.* **6** (2001) 215–238, [arXiv:astro-ph/0102179 \[astro-ph\]](#).
- [337] O. Pisanti, A. Cirillo, S. Esposito, F. Iocco, G. Mangano, G. Miele, and P. D. Serpico, “PArthENoPE: Public Algorithm Evaluating the Nucleosynthesis of Primordial Elements,” *Comput. Phys. Commun.* **178** (2008) 956–971, [arXiv:0705.0290 \[astro-ph\]](#).
- [338] B. D. Fields, “The primordial lithium problem,” *Ann. Rev. Nucl. Part. Sci.* **61** (2011) 47–68, [arXiv:1203.3551 \[astro-ph.CO\]](#).

- [339] P. D. Serpico, S. Esposito, F. Iocco, G. Mangano, G. Miele, and O. Pisanti, “Nuclear reaction network for primordial nucleosynthesis: A Detailed analysis of rates, uncertainties and light nuclei yields,” *JCAP* **0412** (2004) 010, [arXiv:astro-ph/0408076 \[astro-ph\]](#).
- [340] R. Cooke, M. Pettini, R. A. Jorgenson, M. T. Murphy, and C. C. Steidel, “Precision measures of the primordial abundance of deuterium,” *Astrophys. J.* **781** no. 1, (2014) 31, [arXiv:1308.3240 \[astro-ph.CO\]](#).
- [341] W. Fischler and J. Meyers, “Dark Radiation Emerging After Big Bang Nucleosynthesis?,” *Phys. Rev.* **D83** (2011) 063520, [arXiv:1011.3501 \[astro-ph.CO\]](#).
- [342] J. L. Menestrina and R. J. Scherrer, “Dark Radiation from Particle Decays during Big Bang Nucleosynthesis,” *Phys. Rev.* **D85** (2012) 047301, [arXiv:1111.0605 \[astro-ph.CO\]](#).
- [343] D. Hooper, F. S. Queiroz, and N. Y. Gnedin, “Non-Thermal Dark Matter Mimicking An Additional Neutrino Species In The Early Universe,” *Phys. Rev.* **D85** (2012) 063513, [arXiv:1111.6599 \[astro-ph.CO\]](#).
- [344] M. Millea, L. Knox, and B. Fields, “New Bounds for Axions and Axion-Like Particles with keV-GeV Masses,” *Phys. Rev.* **D92** no. 2, (2015) 023010, [arXiv:1501.04097 \[astro-ph.CO\]](#).
- [345] G. Gubitosi, F. Piazza, and F. Vernizzi, “The Effective Field Theory of Dark Energy,” *JCAP* **1302** (2013) 032, [arXiv:1210.0201 \[hep-th\]](#). [*JCAP*1302,032(2013)].
- [346] J. K. Bloomfield, . . Flanagan, M. Park, and S. Watson, “Dark energy or modified gravity? An effective field theory approach,” *JCAP* **1308** (2013) 010, [arXiv:1211.7054 \[astro-ph.CO\]](#).
- [347] S. Weinberg, “Effective Field Theory for Inflation,” *Phys. Rev.* **D77** (2008) 123541, [arXiv:0804.4291 \[hep-th\]](#).
- [348] P. Creminelli, G. D’Amico, J. Norena, and F. Vernizzi, “The Effective Theory of Quintessence: the $w < -1$ Side Unveiled,” *JCAP* **0902** (2009) 018, [arXiv:0811.0827 \[astro-ph\]](#).
- [349] M. Park, K. M. Zurek, and S. Watson, “A Unified Approach to Cosmic Acceleration,” *Phys. Rev.* **D81** (2010) 124008, [arXiv:1003.1722 \[hep-th\]](#).
- [350] R. Jimenez, P. Talavera, and L. Verde, “An effective theory of accelerated expansion,” *Int. J. Mod. Phys.* **A27** (2012) 1250174, [arXiv:1107.2542 \[astro-ph.CO\]](#).
- [351] M. Raveri, B. Hu, N. Frusciante, and A. Silvestri, “Effective Field Theory of Cosmic Acceleration: constraining dark energy with CMB data,” *Phys. Rev.* **D90** no. 4, (2014) 043513, [arXiv:1405.1022 \[astro-ph.CO\]](#).
- [352] J. Gleyzes, D. Langlois, F. Piazza, and F. Vernizzi, “Essential Building Blocks of Dark Energy,” *JCAP* **1308** (2013) 025, [arXiv:1304.4840 \[hep-th\]](#).
- [353] J. Bloomfield, “A Simplified Approach to General Scalar-Tensor Theories,” *JCAP* **1312** (2013) 044, [arXiv:1304.6712 \[astro-ph.CO\]](#).
- [354] F. Piazza and F. Vernizzi, “Effective Field Theory of Cosmological Perturbations,” *Class. Quant. Grav.* **30** (2013) 214007, [arXiv:1307.4350 \[hep-th\]](#).
- [355] J. Gleyzes, D. Langlois, and F. Vernizzi, “A unifying description of dark energy,” *Int. J. Mod. Phys.* **D23** no. 13, (2015) 1443010, [arXiv:1411.3712 \[hep-th\]](#).

- [356] J. Gleyzes, D. Langlois, F. Piazza, and F. Vernizzi, “Healthy theories beyond Horndeski,” *Phys. Rev. Lett.* **114** no. 21, (2015) 211101, [arXiv:1404.6495 \[hep-th\]](#).
- [357] N. Frusciante, M. Raveri, D. Vernieri, B. Hu, and A. Silvestri, “Ho?ava Gravity in the Effective Field Theory formalism: From cosmology to observational constraints,” *Phys. Dark Univ.* **13** (2016) 7–24, [arXiv:1508.01787 \[astro-ph.CO\]](#).
- [358] N. Frusciante, G. Papadomanolakis, and A. Silvestri, “An Extended action for the effective field theory of dark energy: a stability analysis and a complete guide to the mapping at the basis of EFTCAMB,” [arXiv:1601.04064 \[gr-qc\]](#).
- [359] F. Zwicky, “Die Rotverschiebung von extragalaktischen Nebeln,” *Helv. Phys. Acta* **6** (1933) 110–127.
- [360] S. D. M. White, J. F. Navarro, A. E. Evrard, and C. S. Frenk, “The Baryon content of galaxy clusters: A Challenge to cosmological orthodoxy,” *Nature* **366** (1993) 429–433.
- [361] M. Donahue, G. M. Voit, I. M. Gioia, G. Luppino, J. P. Hughes, and J. T. Stocke, “A very hot, high redshift cluster of galaxies: more trouble for $\omega = 1$,” *Astrophys. J.* **502** (1998) 550, [arXiv:astro-ph/9707010 \[astro-ph\]](#).
- [362] N. A. Bahcall and X.-h. Fan, “The Most massive distant clusters: Determining omega and sigma8,” *Astrophys. J.* **504** (1998) 1, [arXiv:astro-ph/9803277 \[astro-ph\]](#).
- [363] D. Clowe, M. Bradac, A. H. Gonzalez, M. Markevitch, S. W. Randall, C. Jones, and D. Zaritsky, “A direct empirical proof of the existence of dark matter,” *Astrophys. J.* **648** (2006) L109–L113, [arXiv:astro-ph/0608407 \[astro-ph\]](#).
- [364] A. Vikhlinin *et al.*, “Chandra Cluster Cosmology Project III: Cosmological Parameter Constraints,” *Astrophys. J.* **692** (2009) 1060–1074, [arXiv:0812.2720 \[astro-ph\]](#).
- [365] A. Mantz, S. W. Allen, D. Rapetti, and H. Ebeling, “The Observed Growth of Massive Galaxy Clusters I: Statistical Methods and Cosmological Constraints,” *Mon. Not. Roy. Astron. Soc.* **406** (2010) 1759–1772, [arXiv:0909.3098 \[astro-ph.CO\]](#).
- [366] D. Rapetti, C. Blake, S. W. Allen, A. Mantz, D. Parkinson, and F. Beutler, “A combined measurement of cosmic growth and expansion from clusters of galaxies, the CMB and galaxy clustering,” *Mon. Not. Roy. Astron. Soc.* **432** (2013) 973, [arXiv:1205.4679 \[astro-ph.CO\]](#).
- [367] B. A. Benson *et al.*, “Cosmological Constraints from Sunyaev-Zel’dovich-Selected Clusters with X-ray Observations in the First 178 Square Degrees of the South Pole Telescope Survey,” *Astrophys. J.* **763** (2013) 147, [arXiv:1112.5435 \[astro-ph.CO\]](#).
- [368] A. B. Mantz, S. W. Allen, R. G. Morris, D. A. Rapetti, D. E. Applegate, P. L. Kelly, A. von der Linden, and R. W. Schmidt, “Cosmology and astrophysics from relaxed galaxy clusters – II. Cosmological constraints,” *Mon. Not. Roy. Astron. Soc.* **440** no. 3, (2014) 2077–2098, [arXiv:1402.6212 \[astro-ph.CO\]](#).
- [369] D. H. Weinberg, M. J. Mortonson, D. J. Eisenstein, C. Hirata, A. G. Riess, and E. Rozo, “Observational Probes of Cosmic Acceleration,” *Phys. Rept.* **530** (2013) 87–255, [arXiv:1201.2434 \[astro-ph.CO\]](#).
- [370] R. A. Sunyaev and Ya. B. Zeldovich, “The Observations of relic radiation as a test of the nature of X-Ray radiation from the clusters of galaxies,” *Comments Astrophys. Space Phys.* **4** (1972) 173–178.
- [371] D. Nagai, “The impact of galaxy formation on the sunyaev-zeldovich effect of galaxy clusters,” *Astrophys. J.* **650** (2006) 538–549, [arXiv:astro-ph/0512208 \[astro-ph\]](#).

- [372] D. Nagai, A. V. Kravtsov, and A. Vikhlinin, “Effects of Galaxy Formation on Thermodynamics of the Intrachuster Medium,” *Astrophys. J.* **668** (2007) 1–14, [arXiv:astro-ph/0703661 \[astro-ph\]](#).
- [373] A. Kravtsov and S. Borgani, “Formation of Galaxy Clusters,” *Ann. Rev. Astron. Astrophys.* **50** (2012) 353–409, [arXiv:1205.5556 \[astro-ph.CO\]](#).
- [374] Z. Staniszewski *et al.*, “Galaxy clusters discovered with a Sunyaev-Zel’dovich effect survey,” *Astrophys. J.* **701** (2009) 32–41, [arXiv:0810.1578 \[astro-ph\]](#).
- [375] K. Vanderlinde *et al.*, “Galaxy Clusters Selected with the Sunyaev-Zel’dovich Effect from 2008 South Pole Telescope Observations,” *Astrophys. J.* **722** (2010) 1180–1196, [arXiv:1003.0003 \[astro-ph.CO\]](#).
- [376] C. L. Reichardt *et al.*, “Galaxy clusters discovered via the Sunyaev-Zel’dovich effect in the first 720 square degrees of the South Pole Telescope survey,” *Astrophys. J.* **763** (2013) 127, [arXiv:1203.5775 \[astro-ph.CO\]](#).
- [377] **Planck** Collaboration, P. A. R. Ade *et al.*, “Planck 2013 results. XXIX. The Planck catalogue of Sunyaev-Zeldovich sources,” *Astron. Astrophys.* **571** (2014) A29, [arXiv:1303.5089 \[astro-ph.CO\]](#).
- [378] **SPT** Collaboration, L. E. Bleem *et al.*, “Galaxy Clusters Discovered via the Sunyaev-Zel’dovich Effect in the 2500-square-degree SPT-SZ survey,” *Astrophys. J. Suppl.* **216** no. 2, (2015) 27, [arXiv:1409.0850 \[astro-ph.CO\]](#).
- [379] **Planck** Collaboration, P. A. R. Ade *et al.*, “Planck 2013 results. XXXII. The updated Planck catalogue of Sunyaev-Zeldovich sources,” *Astron. Astrophys.* **581** (2015) A14, [arXiv:1502.00543 \[astro-ph.CO\]](#).
- [380] A. Pillepich, C. Porciani, and T. H. Reiprich, “The X-ray cluster survey with eROSITA: forecasts for cosmology, cluster physics, and primordial non-Gaussianity,” *Mon. Not. Roy. Astron. Soc.* **422** (2012) 44–69, [arXiv:1111.6587 \[astro-ph.CO\]](#).
- [381] **SPT-3G** Collaboration, B. A. Benson *et al.*, “SPT-3G: A Next-Generation Cosmic Microwave Background Polarization Experiment on the South Pole Telescope,” *Proc. SPIE Int. Soc. Opt. Eng.* **9153** (2014) 91531P, [arXiv:1407.2973 \[astro-ph.IM\]](#).
- [382] E. Calabrese *et al.*, “Precision Epoch of Reionization studies with next-generation CMB experiments,” *JCAP* **1408** (2014) 010, [arXiv:1406.4794 \[astro-ph.CO\]](#).
- [383] S. Bhattacharya and A. Kosowsky, “Dark Energy Constraints from Galaxy Cluster Peculiar Velocities,” *Phys. Rev.* **D77** (2008) 083004, [arXiv:0712.0034 \[astro-ph\]](#).
- [384] A. Kosowsky and S. Bhattacharya, “A Future Test of Gravitation Using Galaxy Cluster Velocities,” *Phys. Rev.* **D80** (2009) 062003, [arXiv:0907.4202 \[astro-ph.CO\]](#).
- [385] E.-M. Mueller, F. de Bernardis, R. Bean, and M. D. Niemack, “Constraints on gravity and dark energy from the pairwise kinematic Sunyaev-Zeldovich effect,” *Astrophys. J.* **808** no. 1, (2015) 47, [arXiv:1408.6248 \[astro-ph.CO\]](#).
- [386] **Planck** Collaboration, R. Adam *et al.*, “Planck 2015 results. I. Overview of products and scientific results,” [arXiv:1502.01582 \[astro-ph.CO\]](#).
- [387] S. Joudaki *et al.*, “CFHTLenS revisited: assessing concordance with Planck including astrophysical systematics,” [arXiv:1601.05786 \[astro-ph.CO\]](#).

- [388] H. Hildebrandt *et al.*, “KiDS-450: Cosmological parameter constraints from tomographic weak gravitational lensing,” [arXiv:1606.05338 \[astro-ph.CO\]](#).
- [389] **DES** Collaboration, T. Abbott *et al.*, “Cosmology from Cosmic Shear with DES Science Verification Data,” [arXiv:1507.05552 \[astro-ph.CO\]](#).
- [390] B. Hu, M. Raveri, N. Frusciante, and A. Silvestri, “EFTCAMB/EFTCosmoMC: Numerical Notes v2.0,” [arXiv:1405.3590 \[astro-ph.IM\]](#).
- [391] E. Bellini and I. Sawicki, “Maximal freedom at minimum cost: linear large-scale structure in general modifications of gravity,” *JCAP* **1407** (2014) 050, [arXiv:1404.3713 \[astro-ph.CO\]](#).
- [392] E. V. Linder, G. Sengr, and S. Watson, “Is the Effective Field Theory of Dark Energy Effective?,” *JCAP* **1605** no. 05, (2016) 053, [arXiv:1512.06180 \[astro-ph.CO\]](#).
- [393] X.-L. Chen and M. Kamionkowski, “Particle decays during the cosmic dark ages,” *Phys. Rev.* **D70** (2004) 043502, [arXiv:astro-ph/0310473 \[astro-ph\]](#).
- [394] N. Padmanabhan and D. P. Finkbeiner, “Detecting dark matter annihilation with CMB polarization: Signatures and experimental prospects,” *Phys. Rev.* **D72** (2005) 023508, [arXiv:astro-ph/0503486 \[astro-ph\]](#).
- [395] M. S. Madhavacheril, N. Sehgal, and T. R. Slatyer, “Current Dark Matter Annihilation Constraints from CMB and Low-Redshift Data,” *Phys. Rev.* **D89** (2014) 103508, [arXiv:1310.3815 \[astro-ph.CO\]](#).
- [396] C. Dvorkin, K. Blum, and M. Zaldarriaga, “Perturbed Recombination from Dark Matter Annihilation,” *Phys. Rev.* **D87** no. 10, (2013) 103522, [arXiv:1302.4753 \[astro-ph.CO\]](#).
- [397] H. Goldberg and L. J. Hall, “A NEW CANDIDATE FOR DARK MATTER,” *Phys. Lett.* **B174** (1986) 151.
- [398] E. D. Carlson, M. E. Machacek, and L. J. Hall, “Self-interacting dark matter,” *Astrophys. J.* **398** (1992) 43–52.
- [399] B.-A. Gradwohl and J. A. Frieman, “Dark matter, long range forces, and large scale structure,” *Astrophys. J.* **398** (1992) 407–424.
- [400] M. E. Machacek, “Growth of adiabatic perturbations in selfinteracting dark matter,” *Astrophys. J.* **431** (1994) 41–51.
- [401] A. A. de Laix, R. J. Scherrer, and R. K. Schaefer, “Constraints of selfinteracting dark matter,” *Astrophys. J.* **452** (1995) 495, [arXiv:astro-ph/9502087 \[astro-ph\]](#).
- [402] F. Atrio-Barandela and S. Davidson, “Interacting hot dark matter,” *Phys. Rev. D* **55** (1997) 5886–5894, [arXiv:astro-ph/9702236 \[astro-ph\]](#).
- [403] C. Boehm, A. Riazuelo, S. H. Hansen, and R. Schaeffer, “Interacting dark matter disguised as warm dark matter,” *Phys. Rev. D* **66** (2002) 083505, [arXiv:astro-ph/0112522 \[astro-ph\]](#).
- [404] R. Foot, “Mirror matter-type dark matter,” *Int. J. Mod. Phys. D* **13** (2004) 2161–2192, [arXiv:astro-ph/0407623 \[astro-ph\]](#).
- [405] A. M. Green, S. Hofmann, and D. J. Schwarz, “The First wimpy halos,” *JCAP* **0508** (2005) 003, [arXiv:astro-ph/0503387 \[astro-ph\]](#).

- [406] S. Profumo, K. Sigurdson, and M. Kamionkowski, “What mass are the smallest protohalos?,” *Phys. Rev. Lett.* **97** (2006) 031301, [arXiv:astro-ph/0603373 \[astro-ph\]](#).
- [407] G. Mangano, A. Melchiorri, P. Serra, A. Cooray, and M. Kamionkowski, “Cosmological bounds on dark matter-neutrino interactions,” *Phys. Rev. D* **74** (2006) 043517, [arXiv:astro-ph/0606190 \[astro-ph\]](#).
- [408] N. Arkani-Hamed, D. P. Finkbeiner, T. R. Slatyer, and N. Weiner, “A Theory of Dark Matter,” *Phys. Rev. D* **79** (2009) 015014, [arXiv:0810.0713 \[hep-ph\]](#).
- [409] J. L. Feng, M. Kaplinghat, H. Tu, and H.-B. Yu, “Hidden Charged Dark Matter,” *jcap* **0907** (2009) 004, [arXiv:0905.3039 \[hep-ph\]](#).
- [410] P. Serra, F. Zalamea, A. Cooray, G. Mangano, and A. Melchiorri, “Constraints on neutrino – dark matter interactions from cosmic microwave background and large scale structure data,” *Phys. Rev. D* **81** (2010) 043507, [arXiv:0911.4411 \[astro-ph.CO\]](#).
- [411] T. Bringmann, “Particle Models and the Small-Scale Structure of Dark Matter,” *New J. Phys.* **11** (2009) 105027, [arXiv:0903.0189 \[astro-ph.CO\]](#).
- [412] D. E. Kaplan, G. Z. Krnjaic, K. R. Rehmann, and C. M. Wells, “Atomic Dark Matter,” *jcap* **1005** (2010) 021, [arXiv:0909.0753 \[hep-ph\]](#).
- [413] S. D. McDermott, H.-B. Yu, and K. M. Zurek, “Turning off the Lights: How Dark is Dark Matter?,” *Phys. Rev. D* **83** (2011) 063509, [arXiv:1011.2907 \[hep-ph\]](#).
- [414] L. G. van den Aarssen, T. Bringmann, and C. Pfrommer, “Is dark matter with long-range interactions a solution to all small-scale problems of Λ CDM cosmology?,” *Phys. Rev. Lett.* **109** (2012) 231301, [arXiv:1205.5809 \[astro-ph.CO\]](#).
- [415] R. Diamanti, E. Giusarma, O. Mena, M. Archidiacono, and A. Melchiorri, “Dark Radiation and interacting scenarios,” *Phys. Rev. D* **87** (2013) 063509, [arXiv:1212.6007 \[astro-ph.CO\]](#).
- [416] M. Baldi, “Structure formation in Multiple Dark Matter cosmologies with long-range scalar interactions,” *Mon. Not. Roy. Astron. Soc.* **428** (2013) 2074, [arXiv:1206.2348 \[astro-ph.CO\]](#).
- [417] J. M. Cline, Z. Liu, and W. Xue, “Millicharged Atomic Dark Matter,” *Phys. Rev. D* **85** (2012) 101302, [arXiv:1201.4858 \[hep-ph\]](#).
- [418] J. Fan, A. Katz, L. Randall, and M. Reece, “Double-Disk Dark Matter,” *Phys. Dark Univ.* **2** (2013) 139–156, [arXiv:1303.1521 \[astro-ph.CO\]](#).
- [419] J. J. Fan, A. Katz, L. Randall, and M. Reece, “Dark-Disk Universe,” *Phys. Rev. Lett.* **110** (2013) 211302, [arXiv:1303.3271 \[hep-ph\]](#).
- [420] F.-Y. Cyr-Racine, R. de Putter, A. Raccanelli, and K. Sigurdson, “Constraints on Large-Scale Dark Acoustic Oscillations from Cosmology,” *Phys. Rev. D* **89** no. 6, (2014) 063517, [arXiv:1310.3278 \[astro-ph.CO\]](#).
- [421] T. Bringmann, J. Hasenkamp, and J. Kersten, “Tight bonds between sterile neutrinos and dark matter,” *JCAP* **1407** (2014) 042, [arXiv:1312.4947 \[hep-ph\]](#).
- [422] R. J. Wilkinson, J. Lesgourgues, and C. Boehm, “Using the CMB angular power spectrum to study Dark Matter-photon interactions,” *jcap* **1404** (2014) 026, [arXiv:1309.7588 \[astro-ph.CO\]](#).

- [423] C. Dvorkin, K. Blum, and M. Kamionkowski, “Constraining Dark Matter-Baryon Scattering with Linear Cosmology,” *Phys. Rev. D* **89** no. 2, (2014) 023519, [arXiv:1311.2937 \[astro-ph.CO\]](#).
- [424] C. Boehm, J. A. Schewtschenko, R. J. Wilkinson, C. M. Baugh, and S. Pascoli, “Using the Milky Way satellites to study interactions between cold dark matter and radiation,” *Mon. Not. Roy. Astron. Soc.* **445** (2014) L31–L35, [arXiv:1404.7012 \[astro-ph.CO\]](#).
- [425] R. J. Wilkinson, C. Boehm, and J. Lesgourges, “Constraining Dark Matter-Neutrino Interactions using the CMB and Large-Scale Structure,” *JCAP* **1405** (2014) 011, [arXiv:1401.7597 \[astro-ph.CO\]](#).
- [426] M. Escudero, O. Mena, A. C. Vincent, R. J. Wilkinson, and C. Boehm, “Exploring dark matter microphysics with galaxy surveys,” *jcap* **1509** no. 09, (2015) 034, [arXiv:1505.06735 \[astro-ph.CO\]](#).
- [427] X. Chu and B. Dasgupta, “Dark Radiation Alleviates Problems with Dark Matter Halos,” *Phys. Rev. Lett.* **113** no. 16, (2014) 161301, [arXiv:1404.6127 \[hep-ph\]](#).
- [428] M. Archidiacono, S. Hannestad, R. S. Hansen, and T. Tram, “Cosmology with self-interacting sterile neutrinos and dark matter - A pseudoscalar model,” *Phys. Rev. D* **91** no. 6, (2015) 065021, [arXiv:1404.5915 \[astro-ph.CO\]](#).
- [429] M. A. Buen-Abad, G. Marques-Tavares, and M. Schmaltz, “Non-Abelian dark matter and dark radiation,” *Phys. Rev. D* **92** no. 2, (2015) 023531, [arXiv:1505.03542 \[hep-ph\]](#).
- [430] J. Lesgourges, G. Marques-Tavares, and M. Schmaltz, “Evidence for dark matter interactions in cosmological precision data?,” *JCAP* **1602** no. 02, (2016) 037, [arXiv:1507.04351 \[astro-ph.CO\]](#).
- [431] W. J. G. de Blok and S. S. McGaugh, “The Dark and visible matter content of low surface brightness disk galaxies,” *Mon. Not. Roy. Astron. Soc.* **290** (1997) 533–552, [arXiv:astro-ph/9704274 \[astro-ph\]](#).
- [432] A. A. Klypin, A. V. Kravtsov, O. Valenzuela, and F. Prada, “Where are the missing Galactic satellites?,” *Astrophys. J.* **522** (1999) 82–92, [arXiv:astro-ph/9901240 \[astro-ph\]](#).
- [433] B. Moore, S. Ghigna, F. Governato, G. Lake, T. R. Quinn, J. Stadel, and P. Tozzi, “Dark matter substructure within galactic halos,” *Astrophys. J.* **524** (1999) L19–L22, [arXiv:astro-ph/9907411 \[astro-ph\]](#).
- [434] J. Zavala, Y. P. Jing, A. Faltenbacher, G. Yepes, Y. Hoffman, S. Gottlober, and B. Catinella, “The velocity function in the local environment from LCDM and LWDM constrained simulations,” *Astrophys. J.* **700** (2009) 1779–1793, [arXiv:0906.0585 \[astro-ph.CO\]](#).
- [435] S.-H. Oh, W. J. G. de Blok, E. Brinks, F. Walter, and R. C. Kennicutt, Jr, “Dark and luminous matter in THINGS dwarf galaxies,” *Astron. J.* **141** (2011) 193, [arXiv:1011.0899 \[astro-ph.CO\]](#).
- [436] M. Boylan-Kolchin, J. S. Bullock, and M. Kaplinghat, “Too big to fail? The puzzling darkness of massive Milky Way subhaloes,” *Mon. Not. Roy. Astron. Soc.* **415** (2011) L40, [arXiv:1103.0007 \[astro-ph.CO\]](#).
- [437] E. Papastergis, A. M. Martin, R. Giovanelli, and M. P. Haynes, “The velocity width function of galaxies from the 40% ALFALFA survey: shedding light on the cold dark matter overabundance problem,” *Astrophys. J.* **739** (2011) 38, [arXiv:1106.0710 \[astro-ph.CO\]](#).
- [438] M. G. Walker and J. Penarrubia, “A Method for Measuring (Slopes of) the Mass Profiles of Dwarf Spheroidal Galaxies,” *Astrophys. J.* **742** (2011) 20, [arXiv:1108.2404 \[astro-ph.CO\]](#).

- [439] M. S. Pawłowski, P. Kroupa, and H. Jerjen, “Dwarf Galaxy Planes: the discovery of symmetric structures in the Local Group,” *Mon. Not. Roy. Astron. Soc.* **435** (2013) 1928, [arXiv:1307.6210 \[astro-ph.CO\]](#).
- [440] A. Klypin, I. Karachentsev, D. Makarov, and O. Nasonova, “Abundance of Field Galaxies,” *Mon. Not. Roy. Astron. Soc.* **454** no. 2, (2015) 1798–1810, [arXiv:1405.4523 \[astro-ph.CO\]](#).
- [441] K. A. Oman *et al.*, “The unexpected diversity of dwarf galaxy rotation curves,” *Mon. Not. Roy. Astron. Soc.* **452** no. 4, (2015) 3650–3665, [arXiv:1504.01437 \[astro-ph.GA\]](#).
- [442] E. Papastergis, R. Giovanelli, M. P. Haynes, and F. Shankar, “Is there a “too big to fail” problem in the field?,” *Astron. Astrophys.* **574** (2015) A113, [arXiv:1407.4665 \[astro-ph.GA\]](#).
- [443] F.-Y. Cyr-Racine, K. Sigurdson, J. Zavala, T. Bringmann, M. Vogelsberger, and C. Pfrommer, “ETHOS - An Effective Theory of Structure Formation: From dark particle physics to the matter distribution of the Universe,” [arXiv:1512.05344 \[astro-ph.CO\]](#).
- [444] F.-Y. Cyr-Racine and K. Sigurdson, “Limits on Neutrino-Neutrino Scattering in the Early Universe,” *Phys. Rev.* **D90** no. 12, (2014) 123533, [arXiv:1306.1536 \[astro-ph.CO\]](#).
- [445] M. Vogelsberger, J. Zavala, F.-Y. Cyr-Racine, C. Pfrommer, T. Bringmann, and K. Sigurdson, “ETHOS - An Effective Theory of Structure Formation: Dark matter physics as a possible explanation of the small-scale CDM problems,” *Mon. Not. Roy. Astron. Soc.* **460** (2016) 1399, [arXiv:1512.05349 \[astro-ph.CO\]](#).
- [446] D. J. E. Marsh, “Axion Cosmology,” [arXiv:1510.07633 \[astro-ph.CO\]](#).
- [447] R. Hlozek, D. Grin, D. J. Marsh, and P. G. Ferreira, “A search for ultralight axions using precision cosmological data,” *Phys. Rev.* **D91** no. 10, (2015) 103512, [arXiv:1410.2896 \[astro-ph.CO\]](#).
- [448] D. J. E. Marsh, “WarmAndFuzzy: the halo model beyond CDM,” [arXiv:1605.05973 \[astro-ph.CO\]](#).
- [449] M. Axenides, R. H. Brandenberger, and M. S. Turner, “Development of Axion Perturbations in an Axion Dominated Universe,” *Phys. Lett.* **B126** (1983) 178.
- [450] P. Fox, A. Pierce, and S. D. Thomas, “Probing a QCD string axion with precision cosmological measurements,” [arXiv:hep-th/0409059 \[hep-th\]](#).
- [451] M. P. Hertzberg, M. Tegmark, and F. Wilczek, “Axion Cosmology and the Energy Scale of Inflation,” *Phys. Rev.* **D78** (2008) 083507, [arXiv:0807.1726 \[astro-ph\]](#).
- [452] P. Svrcek and E. Witten, “Axions In String Theory,” *JHEP* **06** (2006) 051, [arXiv:hep-th/0605206 \[hep-th\]](#).
- [453] J. Preskill, M. B. Wise, and F. Wilczek, “Cosmology of the Invisible Axion,” *Phys. Lett.* **B120** (1983) 127–132.
- [454] L. F. Abbott and P. Sikivie, “A Cosmological Bound on the Invisible Axion,” *Phys. Lett.* **B120** (1983) 133–136.
- [455] M. Dine and W. Fischler, “The Not So Harmless Axion,” *Phys. Lett.* **B120** (1983) 137–141.
- [456] M. S. Turner, “Coherent Scalar Field Oscillations in an Expanding Universe,” *Phys. Rev.* **D28** (1983) 1243.
- [457] P. J. Steinhardt and M. S. Turner, “Saving the Invisible Axion,” *Phys. Lett.* **B129** (1983) 51.

- [458] D. J. E. Marsh and P. G. Ferreira, “Ultra-Light Scalar Fields and the Growth of Structure in the Universe,” *Phys. Rev.* **D82** (2010) 103528, [arXiv:1009.3501 \[hep-ph\]](#).
- [459] **ADMX** Collaboration, S. J. Asztalos *et al.*, “A SQUID-based microwave cavity search for dark-matter axions,” *Phys. Rev. Lett.* **104** (2010) 041301, [arXiv:0910.5914 \[astro-ph.CO\]](#).
- [460] D. Budker, P. W. Graham, M. Ledbetter, S. Rajendran, and A. Sushkov, “Proposal for a Cosmic Axion Spin Precession Experiment (CASPER),” *Phys. Rev.* **X4** no. 2, (2014) 021030, [arXiv:1306.6089 \[hep-ph\]](#).
- [461] L. Visinelli and P. Gondolo, “Axion cold dark matter in view of BICEP2 results,” *Phys. Rev. Lett.* **113** (2014) 011802, [arXiv:1403.4594 \[hep-ph\]](#).
- [462] D. J. E. Marsh, D. Grin, R. Hlozek, and P. G. Ferreira, “Tensor Interpretation of BICEP2 Results Severely Constrains Axion Dark Matter,” *Phys. Rev. Lett.* **113** no. 1, (2014) 011801, [arXiv:1403.4216 \[astro-ph.CO\]](#).
- [463] T. Higaki, K. S. Jeong, and F. Takahashi, “Solving the Tension between High-Scale Inflation and Axion Isocurvature Perturbations,” *Phys. Lett.* **B734** (2014) 21–26, [arXiv:1403.4186 \[hep-ph\]](#).
- [464] M. Fairbairn, R. Hogan, and D. J. E. Marsh, “Unifying inflation and dark matter with the Peccei-Quinn field: observable axions and observable tensors,” *Phys. Rev.* **D91** no. 2, (2015) 023509, [arXiv:1410.1752 \[hep-ph\]](#).
- [465] Y. Nomura, S. Rajendran, and F. Sanches, “Axion Isocurvature and Magnetic Monopoles,” *Phys. Rev. Lett.* **116** no. 14, (2016) 141803, [arXiv:1511.06347 \[hep-ph\]](#).
- [466] D. J. E. Marsh, D. Grin, R. Hlozek, and P. G. Ferreira, “Axiverse cosmology and the energy scale of inflation,” *Phys. Rev.* **D87** (2013) 121701, [arXiv:1303.3008 \[astro-ph.CO\]](#).
- [467] D. J. E. Marsh and J. Silk, “A Model For Halo Formation With Axion Mixed Dark Matter,” *Mon. Not. Roy. Astron. Soc.* **437** no. 3, (2014) 2652–2663, [arXiv:1307.1705 \[astro-ph.CO\]](#).
- [468] B. Bozek, D. J. E. Marsh, J. Silk, and R. F. G. Wyse, “Galaxy UV-luminosity function and reionization constraints on axion dark matter,” *Mon. Not. Roy. Astron. Soc.* **450** no. 1, (2015) 209–222, [arXiv:1409.3544 \[astro-ph.CO\]](#).
- [469] H.-Y. Schive, T. Chiueh, T. Broadhurst, and K.-W. Huang, “Contrasting Galaxy Formation from Quantum Wave Dark Matter, ψ DM, with Λ CDM, using Planck and Hubble Data,” *Astrophys. J.* **818** no. 1, (2016) 89, [arXiv:1508.04621 \[astro-ph.GA\]](#).
- [470] P. W. Graham, I. G. Irastorza, S. K. Lamoreaux, A. Lindner, and K. A. van Bibber, “Experimental Searches for the Axion and Axion-Like Particles,” *Ann. Rev. Nucl. Part. Sci.* **65** (2015) 485–514, [arXiv:1602.00039 \[hep-ex\]](#).
- [471] G. G. Raffelt, “Neutrinos and the stars,” *Proc. Int. Sch. Phys. Fermi* **182** (2012) 61–143, [arXiv:1201.1637 \[astro-ph.SR\]](#).
- [472] E. Di Valentino, E. Giusarma, M. Lattanzi, O. Mena, A. Melchiorri, and J. Silk, “Cosmological Axion and neutrino mass constraints from Planck 2015 temperature and polarization data,” *Phys. Lett.* **B752** (2016) 182–185, [arXiv:1507.08665 \[astro-ph.CO\]](#).
- [473] D. Cadamuro and J. Redondo, “Cosmological bounds on pseudo Nambu-Goldstone bosons,” *JCAP* **1202** (2012) 032, [arXiv:1110.2895 \[hep-ph\]](#).

- [474] D. Baumann, D. Green, and B. Wallisch, “A New Target for Cosmic Axion Searches,” [arXiv:1604.08614 \[astro-ph.CO\]](https://arxiv.org/abs/1604.08614).
- [475] S. M. Carroll, “Quintessence and the rest of the world,” *Phys. Rev. Lett.* **81** (1998) 3067–3070, [arXiv:astro-ph/9806099 \[astro-ph\]](https://arxiv.org/abs/astro-ph/9806099).
- [476] M. Kamionkowski, “How to De-Rotate the Cosmic Microwave Background Polarization,” *Phys. Rev. Lett.* **102** (2009) 111302, [arXiv:0810.1286 \[astro-ph\]](https://arxiv.org/abs/0810.1286).
- [477] V. Gluscevic, M. Kamionkowski, and A. Cooray, “De-Rotation of the Cosmic Microwave Background Polarization: Full-Sky Formalism,” *Phys. Rev.* **D80** (2009) 023510, [arXiv:0905.1687 \[astro-ph.CO\]](https://arxiv.org/abs/0905.1687).
- [478] V. Gluscevic, D. Hanson, M. Kamionkowski, and C. M. Hirata, “First CMB Constraints on Direction-Dependent Cosmological Birefringence from WMAP-7,” *Phys. Rev.* **D86** (2012) 103529, [arXiv:1206.5546 \[astro-ph.CO\]](https://arxiv.org/abs/1206.5546).
- [479] **POLARBEAR** Collaboration, P. A. R. Ade *et al.*, “POLARBEAR Constraints on Cosmic Birefringence and Primordial Magnetic Fields,” *Phys. Rev.* **D92** (2015) 123509, [arXiv:1509.02461 \[astro-ph.CO\]](https://arxiv.org/abs/1509.02461).
- [480] R. Durrer and A. Neronov, “Cosmological Magnetic Fields: Their Generation, Evolution and Observation,” *Astron. Astrophys. Rev.* **21** (2013) 62, [arXiv:1303.7121 \[astro-ph.CO\]](https://arxiv.org/abs/1303.7121).
- [481] D. Grasso and H. R. Rubinstein, “Magnetic fields in the early universe,” *Phys. Rept.* **348** (2001) 163–266, [arXiv:astro-ph/0009061 \[astro-ph\]](https://arxiv.org/abs/astro-ph/0009061).
- [482] T. Vachaspati, “Magnetic fields from cosmological phase transitions,” *Phys. Lett.* **B265** (1991) 258–261.
- [483] M. S. Turner and L. M. Widrow, “Inflation Produced, Large Scale Magnetic Fields,” *Phys. Rev.* **D37** (1988) 2743.
- [484] B. Ratra, “Cosmological ‘seed’ magnetic field from inflation,” *Astrophys. J.* **391** (1992) L1–L4.
- [485] C. Bonvin, C. Caprini, and R. Durrer, “Magnetic fields from inflation: the transition to the radiation era,” *Phys. Rev.* **D86** (2012) 023519, [arXiv:1112.3901 \[astro-ph.CO\]](https://arxiv.org/abs/1112.3901).
- [486] C. Caprini, R. Durrer, and G. Servant, “Gravitational wave generation from bubble collisions in first-order phase transitions: An analytic approach,” *Phys. Rev.* **D77** (2008) 124015, [arXiv:0711.2593 \[astro-ph\]](https://arxiv.org/abs/0711.2593).
- [487] T. Vachaspati, “Estimate of the primordial magnetic field helicity,” *Phys. Rev. Lett.* **87** (2001) 251302, [arXiv:astro-ph/0101261 \[astro-ph\]](https://arxiv.org/abs/astro-ph/0101261).
- [488] K. Subramanian and J. D. Barrow, “Microwave background signals from tangled magnetic fields,” *Phys. Rev. Lett.* **81** (1998) 3575–3578, [arXiv:astro-ph/9803261 \[astro-ph\]](https://arxiv.org/abs/astro-ph/9803261).
- [489] A. Mack, T. Kahniashvili, and A. Kosowsky, “Microwave background signatures of a primordial stochastic magnetic field,” *Phys. Rev.* **D65** (2002) 123004, [arXiv:astro-ph/0105504 \[astro-ph\]](https://arxiv.org/abs/astro-ph/0105504).
- [490] A. Lewis, “CMB anisotropies from primordial inhomogeneous magnetic fields,” *Phys. Rev.* **D70** (2004) 043011, [arXiv:astro-ph/0406096 \[astro-ph\]](https://arxiv.org/abs/astro-ph/0406096).
- [491] J. R. Shaw and A. Lewis, “Massive Neutrinos and Magnetic Fields in the Early Universe,” *Phys. Rev.* **D81** (2010) 043517, [arXiv:0911.2714 \[astro-ph.CO\]](https://arxiv.org/abs/0911.2714).

- [492] D. Paoletti and F. Finelli, “CMB Constraints on a Stochastic Background of Primordial Magnetic Fields,” *Phys. Rev.* **D83** (2011) 123533, [arXiv:1005.0148 \[astro-ph.CO\]](#).
- [493] K. E. Kunze and E. Komatsu, “Constraints on primordial magnetic fields from the optical depth of the cosmic microwave background,” *JCAP* **1506** no. 06, (2015) 027, [arXiv:1501.00142 \[astro-ph.CO\]](#).
- [494] A. Kosowsky, T. Kahnashvili, G. Lavrelashvili, and B. Ratra, “Faraday rotation of the Cosmic Microwave Background polarization by a stochastic magnetic field,” *Phys. Rev.* **D71** (2005) 043006, [arXiv:astro-ph/0409767 \[astro-ph\]](#).
- [495] L. Pogosian, A. P. S. Yadav, Y.-F. Ng, and T. Vachaspati, “Primordial Magnetism in the CMB: Exact Treatment of Faraday Rotation and WMAP7 Bounds,” *Phys. Rev.* **D84** (2011) 043530, [arXiv:1106.1438 \[astro-ph.CO\]](#). [Erratum: Phys. Rev.D84,089903(2011)].
- [496] **Planck** Collaboration, P. A. R. Ade *et al.*, “Planck 2015 results. XIX. Constraints on primordial magnetic fields,” [arXiv:1502.01594 \[astro-ph.CO\]](#).
- [497] J. Lizarraga, J. Urrestilla, D. Dáverio, M. Hindmarsh, M. Kunz, and A. R. Liddle, “Constraining topological defects with temperature and polarization anisotropies,” *Phys. Rev.* **D90** no. 10, (2014) 103504, [arXiv:1408.4126 \[astro-ph.CO\]](#).
- [498] A. Lazanu and P. Shellard, “Constraints on the Nambu-Goto cosmic string contribution to the CMB power spectrum in light of new temperature and polarisation data,” *JCAP* **1502** no. 02, (2015) 024, [arXiv:1410.5046 \[astro-ph.CO\]](#).
- [499] A. Moss and L. Pogosian, “Did BICEP2 see vector modes? First B-mode constraints on cosmic defects,” *Phys. Rev. Lett.* **112** (2014) 171302, [arXiv:1403.6105 \[astro-ph.CO\]](#).
- [500] **POLARBEAR** Collaboration, P. A. R. Ade *et al.*, “A Measurement of the Cosmic Microwave Background B-Mode Polarization Power Spectrum at Sub-Degree Scales with POLARBEAR,” *Astrophys. J.* **794** no. 2, (2014) 171, [arXiv:1403.2369 \[astro-ph.CO\]](#).
- [501] **BICEP2** Collaboration, P. A. R. Ade *et al.*, “Detection of B-Mode Polarization at Degree Angular Scales by BICEP2,” *Phys. Rev. Lett.* **112** no. 24, (2014) 241101, [arXiv:1403.3985 \[astro-ph.CO\]](#).
- [502] S. Foreman, A. Moss, and D. Scott, “Predicted Constraints on Cosmic String Tension from Planck and Future CMB Polarization Measurements,” *Phys. Rev.* **D84** (2011) 043522, [arXiv:1106.4018 \[astro-ph.CO\]](#).
- [503] L. Pogosian and T. Vachaspati, “Cosmic microwave background anisotropy from wiggly strings,” *Phys. Rev.* **D60** (1999) 083504, [arXiv:astro-ph/9903361 \[astro-ph\]](#).
- [504] M. Tegmark, “How to make maps from CMB data without losing information,” *Astrophys. J.* **480** (1997) L87–L90, [arXiv:astro-ph/9611130 \[astro-ph\]](#).
- [505] T. J. Poutanen, D. Maino, H. Kurki-Suonio, E. Keihanen, and E. Hivon, “Cosmic microwave background power spectrum estimation with the destriping technique,” *Mon. Not. Roy. Astron. Soc.* **353** (2004) 43, [arXiv:astro-ph/0404134 \[astro-ph\]](#).
- [506] E. Keihanen, H. Kurki-Suonio, T. Poutanen, D. Maino, and C. Burigana, “A maximum likelihood approach to the destriping technique,” *Astron. Astrophys.* **428** (2004) 287–298, [arXiv:astro-ph/0304411 \[astro-ph\]](#).
- [507] E. Keihanen, R. Keskitalo, H. Kurki-Suonio, T. Poutanen, and A. S. Sirvio, “Making CMB temperature and polarization maps with Madam,” *Astron. Astrophys.* **510** (2010) A57, [arXiv:0907.0367 \[astro-ph.CO\]](#).

- [508] M. Tristram, C. Filliard, O. Perdereau, S. Plaszczynski, R. Stompor, and F. Touze, “Iterative destriping and photometric calibration for Planck-HFI, polarized, multi-detector map-making,” *Astron. Astrophys.* **534** (2011) A88, [arXiv:1103.2281 \[astro-ph.CO\]](#).
- [509] **Planck** Collaboration, P. A. R. Ade *et al.*, “Planck 2015 results. VI. LFI mapmaking,” [arXiv:1502.01585 \[astro-ph.CO\]](#).
- [510] **Planck** Collaboration, R. Adam *et al.*, “Planck 2015 results. VIII. High Frequency Instrument data processing: Calibration and maps,” [arXiv:1502.01587 \[astro-ph.CO\]](#).
- [511] R. Stompor *et al.*, “Making maps of the cosmic microwave background: The MAXIMA example,” *Phys. Rev.* **D65** (2002) 022003, [arXiv:astro-ph/0106451 \[astro-ph\]](#).
- [512] C. M. Cantalupo, J. D. Borrill, A. H. Jaffe, T. S. Kisner, and R. Stompor, “MADmap: A Massively Parallel Maximum-Likelihood Cosmic Microwave Background Map-Maker,” *Astrophys. J. Suppl.* **187** (2010) 212–227, [arXiv:0906.1775 \[astro-ph.CO\]](#).
- [513] R. Dunner *et al.*, “The Atacama Cosmology Telescope: Data Characterization and Map Making,” *Astrophys. J.* **762** (2013) 10, [arXiv:1208.0050 \[astro-ph.IM\]](#).
- [514] C. Armitage and B. D. Wandelt, “Deconvolution map-making for cosmic microwave background observations,” *Phys. Rev.* **D70** (2004) 123007, [arXiv:astro-ph/0410092 \[astro-ph\]](#).
- [515] D. L. Harrison, F. van Leeuwen, and M. A. J. Ashdown, “A deconvolution map-making method for experiments with circular scanning strategies,” *Astron. Astrophys.* **532** (2011) A55, [arXiv:1106.4237 \[astro-ph.IM\]](#).
- [516] E. Keihanen and M. Reinecke, “ArtDeco: A beam deconvolution code for absolute CMB measurements,” *Astron. Astrophys.* (2012) 20183,, [arXiv:1208.1399 \[astro-ph.IM\]](#). [*Astron. Astrophys.* 548,A110(2012)].
- [517] E. Hivon, K. M. Gorski, C. B. Netterfield, B. P. Crill, S. Prunet, and F. Hansen, “Master of the cosmic microwave background anisotropy power spectrum: a fast method for statistical analysis of large and complex cosmic microwave background data sets,” *Astrophys. J.* **567** (2002) 2, [arXiv:astro-ph/0105302 \[astro-ph\]](#).
- [518] T. Culverhouse *et al.*, “The QUaD Galactic Plane Survey 1: Maps And Analysis of Diffuse Emission,” *Astrophys. J.* **722** (2010) 1057–1077, [arXiv:1001.1333 \[astro-ph.GA\]](#).
- [519] K. K. Schaffer *et al.*, “The First Public Release of South Pole Telescope Data: Maps of a 95-square-degree Field from 2008 Observations,” *Astrophys. J.* **743** (2011) 90, [arXiv:1111.7245 \[astro-ph.CO\]](#).
- [520] **Planck** Collaboration, R. Adam *et al.*, “Planck 2015 results. IX. Diffuse component separation: CMB maps,” [arXiv:1502.05956 \[astro-ph.CO\]](#).
- [521] **WMAP** Collaboration, L. Page *et al.*, “Three year Wilkinson Microwave Anisotropy Probe (WMAP) observations: polarization analysis,” *Astrophys. J. Suppl.* **170** (2007) 335, [arXiv:astro-ph/0603450 \[astro-ph\]](#).
- [522] Y. Fantaye, C. Baccigalupi, S. Leach, and A. P. S. Yadav, “CMB lensing reconstruction in the presence of diffuse polarized foregrounds,” *JCAP* **1212** (2012) 017, [arXiv:1207.0508 \[astro-ph.CO\]](#).
- [523] J. Delabrouille and J.-F. Cardoso, “Diffuse Source Separation in CMB Observations,” in *Data Analysis in Cosmology*, V. J. Martínez, E. Saar, E. Martínez-González, and M.-J. Pons-Bordería, eds., vol. 665 of *Lecture Notes in Physics*, Berlin Springer Verlag, pp. 159–205. 2009.

- [524] F. Stivoli, J. Grain, S. M. Leach, M. Tristram, C. Baccigalupi, and R. Stompor, “Maximum likelihood, parametric component separation and CMB B-mode detection in suborbital experiments,” *Mon. Not. Roy. Astron. Soc.* **408** (2010) 2319–2335, [arXiv:1004.4756 \[astro-ph.CO\]](#).
- [525] Y. Fantaye, F. Stivoli, J. Grain, S. M. Leach, M. Tristram, C. Baccigalupi, and R. Stompor, “Estimating the tensor-to-scalar ratio and the effect of residual foreground contamination,” *JCAP* **1108** (2011) 001, [arXiv:1104.1367 \[astro-ph.CO\]](#).
- [526] W. N. Brandt, C. R. Lawrence, A. C. S. Readhead, J. N. Pakianathan, and T. M. Fiola, “Separation of foreground radiation from cosmic microwave background anisotropy using multifrequency measurements,” *Ap. J.* **424** (Mar., 1994) 1–21.
- [527] H. K. Eriksen *et al.*, “CMB component separation by parameter estimation,” *Astrophys. J.* **641** (2006) 665–682, [arXiv:astro-ph/0508268 \[astro-ph\]](#).
- [528] **SDSS** Collaboration, M. Tegmark *et al.*, “Cosmological parameters from SDSS and WMAP,” *Phys. Rev.* **D69** (2004) 103501, [arXiv:astro-ph/0310723 \[astro-ph\]](#).
- [529] **WMAP** Collaboration, G. Hinshaw *et al.*, “Three-year Wilkinson Microwave Anisotropy Probe (WMAP) observations: temperature analysis,” *Astrophys. J. Suppl.* **170** (2007) 288, [arXiv:astro-ph/0603451 \[astro-ph\]](#).
- [530] E. Martinez-Gonzalez, J. M. Diego, P. Vielva, and J. Silk, “CMB power spectrum estimation and map reconstruction with the expectation - Maximization algorithm,” *Mon. Not. Roy. Astron. Soc.* **345** (2003) 1101, [arXiv:astro-ph/0302094 \[astro-ph\]](#).
- [531] J.-F. Cardoso, M. Le Jeune, J. Delabrouille, M. Betoule, and G. Patanchon, “Component Separation With Flexible Models - Application to Multichannel Astrophysical Observations,” *IEEE Journal of Selected Topics in Signal Processing* **2** (Nov., 2008) 735–746.
- [532] N. Katayama and E. Komatsu, “Simple foreground cleaning algorithm for detecting primordial B-mode polarization of the cosmic microwave background,” *Astrophys. J.* **737** (2011) 78, [arXiv:1101.5210 \[astro-ph.CO\]](#).
- [533] B. Gold *et al.*, “Seven-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Galactic Foreground Emission,” *Astrophys. J. Suppl.* **192** (2011) 15, [arXiv:1001.4555 \[astro-ph.GA\]](#).
- [534] **Planck** Collaboration, N. Aghanim *et al.*, “Planck 2015 results. XI. CMB power spectra, likelihoods, and robustness of parameters,” *Submitted to: Astron. Astrophys.* (2015) , [arXiv:1507.02704 \[astro-ph.CO\]](#).
- [535] **POLARBEAR** Collaboration, J. Errard *et al.*, “Modeling atmospheric emission for CMB ground-based observations,” *Astrophys. J.* **809** (2015) 63, [arXiv:1501.07911 \[astro-ph.IM\]](#).
- [536] E. S. Battistelli *et al.*, “Intensity and polarization of the atmospheric emission at millimetric wavelengths at Dome Concordia,” *Mon. Not. Roy. Astron. Soc.* **423** (2012) 1293–1299, [arXiv:1203.5615 \[astro-ph.IM\]](#).
- [537] E. L. Wright, “Scanning and mapping strategies for CMB experiments,” [arXiv:astro-ph/9612006 \[astro-ph\]](#).
- [538] J. R. Bond, A. H. Jaffe, and L. Knox, “Estimating the power spectrum of the cosmic microwave background,” *Phys. Rev.* **D57** (1998) 2117–2137, [arXiv:astro-ph/9708203 \[astro-ph\]](#).

- [539] J. Borrill, “The challenge of data analysis for future CMB observations,” *AIP Conf. Proc.* **476** (1999) 277–282, [arXiv:astro-ph/9903204 \[astro-ph\]](#).
- [540] **ACTPol** Collaboration, S. Naess *et al.*, “The Atacama Cosmology Telescope: CMB Polarization at $200 < \ell < 9000$,” *JCAP* **1410** no. 10, (2014) 007, [arXiv:1405.5524 \[astro-ph.CO\]](#).
- [541] **SPT** Collaboration, A. T. Crites *et al.*, “Measurements of E-Mode Polarization and Temperature-E-Mode Correlation in the Cosmic Microwave Background from 100 Square Degrees of SPTpol Data,” *Astrophys. J.* **805** no. 1, (2015) 36, [arXiv:1411.1042 \[astro-ph.CO\]](#).
- [542] A. Challinor and A. Lewis, “Lensed CMB power spectra from all-sky correlation functions,” *Phys. Rev.* **D71** (2005) 103010, [arXiv:astro-ph/0502425 \[astro-ph\]](#).
- [543] K. M. Smith, “Pseudo-c(l) estimators which do not mix e and b modes,” *Phys. Rev.* **D74** (2006) 083002, [arXiv:astro-ph/0511629 \[astro-ph\]](#).
- [544] T. Okamoto and W. Hu, “CMB lensing reconstruction on the full sky,” *Phys. Rev.* **D67** (2003) 083002, [arXiv:astro-ph/0301031 \[astro-ph\]](#).
- [545] **Planck** Collaboration, P. Ade *et al.*, “Planck 2013 results. XVI. Cosmological parameters,” [arXiv:1303.5076 \[astro-ph.CO\]](#).
- [546] A. Lewis and S. Bridle, “Cosmological parameters from CMB and other data: a Monte- Carlo approach,” *Phys. Rev.* **D66** (2002) 103511, [astro-ph/0205436](#).
- [547] M. M. Schmittfull, A. Challinor, D. Hanson, and A. Lewis, “Joint analysis of CMB temperature and lensing-reconstruction power spectra,” *Phys. Rev.* **D88** no. 6, (2013) 063012, [arXiv:1308.0286 \[astro-ph.CO\]](#).
- [548] C. Dvorkin and K. M. Smith, “Reconstructing Patchy Reionization from the Cosmic Microwave Background,” *Phys. Rev.* **D79** (2009) 043003, [arXiv:0812.1566 \[astro-ph\]](#).
- [549] D. Hanson, A. Challinor, G. Efstathiou, and P. Bielewicz, “CMB temperature lensing power reconstruction,” *Phys. Rev.* **D83** (2011) 043005, [arXiv:1008.4403 \[astro-ph.CO\]](#).
- [550] T. Namikawa, D. Hanson, and R. Takahashi, “Bias-Hardened CMB Lensing,” *Mon. Not. Roy. Astron. Soc.* **431** (2013) 609–620, [arXiv:1209.0091 \[astro-ph.CO\]](#).
- [551] A. Lewis, A. Challinor, and D. Hanson, “The shape of the CMB lensing bispectrum,” *JCAP* **1103** (2011) 018, [arXiv:1101.2234 \[astro-ph.CO\]](#).
- [552] **Planck Collaboration** Collaboration, P. Ade *et al.*, “Planck 2013 results. XVII. Gravitational lensing by large-scale structure,” [arXiv:1303.5077 \[astro-ph.CO\]](#).
- [553] K. M. Smith, W. Hu, and M. Kaplinghat, “Cosmological Information from Lensed CMB Power Spectra,” *Phys. Rev.* **D74** (2006) 123002, [arXiv:astro-ph/0607315 \[astro-ph\]](#).
- [554] C. Li, T. L. Smith, and A. Cooray, “Non-Gaussian Covariance of CMB B-modes of Polarization and Parameter Degradation,” *Phys. Rev.* **D75** (2007) 083501, [arXiv:astro-ph/0607494 \[astro-ph\]](#).
- [555] A. Benoit-Levy, K. M. Smith, and W. Hu, “Non-Gaussian structure of the lensed CMB power spectra covariance matrix,” *Phys. Rev.* **D86** (2012) 123008, [arXiv:1205.0474 \[astro-ph.CO\]](#).
- [556] B. D. Sherwin and M. Schmittfull, “Delensing the CMB with the Cosmic Infrared Background,” *Phys. Rev.* **D92** no. 4, (2015) 043005, [arXiv:1502.05356 \[astro-ph.CO\]](#).

- [557] G. Simard, D. Hanson, and G. Holder, “Prospects for Delensing the Cosmic Microwave Background for Studying Inflation,” *Astrophys. J.* **807** no. 2, (2015) 166, [arXiv:1410.0691 \[astro-ph.CO\]](https://arxiv.org/abs/1410.0691).
- [558] M. Remazeilles, C. Dickinson, H. K. K. Eriksen, and I. K. Wehus, “Sensitivity and foreground modelling for large-scale cosmic microwave background B-mode polarization satellite missions,” *Mon. Not. Roy. Astron. Soc.* **458** no. 2, (2016) 2032–2050, [arXiv:1509.04714 \[astro-ph.CO\]](https://arxiv.org/abs/1509.04714).
- [559] N. Jarosik *et al.*, “Seven-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Sky Maps, Systematic Errors, and Basic Results,” *Astrophys. J. Suppl.* **192** (2011) 14, [arXiv:1001.4744 \[astro-ph.CO\]](https://arxiv.org/abs/1001.4744).
- [560] J. Delabrouille *et al.*, “The pre-launch Planck Sky Model: a model of sky emission at submillimetre to centimetre wavelengths,” *Astron. Astrophys.* **553** (2013) A96, [arXiv:1207.3675 \[astro-ph.CO\]](https://arxiv.org/abs/1207.3675).
- [561] C. G. T. Haslam, U. Klein, C. J. Salter, H. Stoffel, W. E. Wilson, M. N. Cleary, D. J. Cooke, and P. Thomasson, “A 408 MHz all-sky continuum survey. I - observations at southern declinations and for the north polar region,” *A & A* **100** (July, 1981) 209–219. http://adsabs.harvard.edu/cgi-bin/nph-bib_query?bibcode=1981A%26A...100..209H&db_key=AST.
- [562] M. A. Miville-Deschenes, N. Ysard, A. Lavabre, N. Ponthieu, J. F. Macias-Perez, J. Aumont, and J. P. Bernard, “Separation of anomalous and synchrotron emissions using WMAP polarization data,” *Astron. Astrophys.* **490** (2008) 1093, [arXiv:0802.3345 \[astro-ph\]](https://arxiv.org/abs/0802.3345).
- [563] **Planck** Collaboration, R. Adam *et al.*, “Planck 2015 results. X. Diffuse component separation: Foreground maps,” [arXiv:1502.01588 \[astro-ph.CO\]](https://arxiv.org/abs/1502.01588).
- [564] B. D. Wandelt and F. K. Hansen, “Fast, exact CMB power spectrum estimation for a certain class of observational strategies,” *Phys. Rev.* **D67** (2003) 023001, [arXiv:astro-ph/0106515 \[astro-ph\]](https://arxiv.org/abs/astro-ph/0106515).
- [565] T. M. Crawford, “Power spectrum sensitivity of raster-scanned CMB experiments in the presence of 1/f noise,” *Phys. Rev.* **D76** (2007) 063008, [arXiv:astro-ph/0702608 \[astro-ph\]](https://arxiv.org/abs/astro-ph/0702608).
- [566] M. Bucher, “An exact result concerning the 1/f noise contribution to the large-angle error in CMB temperature and polarization maps,” [arXiv:1602.08274 \[astro-ph.IM\]](https://arxiv.org/abs/1602.08274).
- [567] C. Armitage-Caplan, J. Dunkley, H. K. Eriksen, and C. Dickinson, “Large-Scale Polarized Foreground Component Separation for Planck,” *Mon. Not. Roy. Astron. Soc.* **418** (2011) 1498, [arXiv:1103.2554 \[astro-ph.CO\]](https://arxiv.org/abs/1103.2554).
- [568] R. Allison, P. Caucal, E. Calabrese, J. Dunkley, and T. Louis, “Towards a cosmological neutrino mass detection,” *Phys. Rev.* **D92** no. 12, (2015) 123535, [arXiv:1509.07471 \[astro-ph.CO\]](https://arxiv.org/abs/1509.07471).
- [569] N. Hand *et al.*, “Evidence of Galaxy Cluster Motions with the Kinematic Sunyaev-Zel’dovich Effect,” *Phys. Rev. Lett.* **109** (2012) 041101, [arXiv:1203.4219 \[astro-ph.CO\]](https://arxiv.org/abs/1203.4219).
- [570] R. Keisler and F. Schmidt, “Prospects for measuring the relative velocities of galaxy clusters in photometric surveys using the kinetic Sunyaev-Zel’dovich Effect,” *Astrophys. J.* **765** (2013) L32, [arXiv:1211.0668 \[astro-ph.CO\]](https://arxiv.org/abs/1211.0668).
- [571] **Planck** Collaboration, P. A. R. Ade *et al.*, “Planck intermediate results. XXXVII. Evidence of unbound gas from the kinetic Sunyaev-Zeldovich effect,” *Astron. Astrophys.* **586** (2016) A140, [arXiv:1504.03339 \[astro-ph.CO\]](https://arxiv.org/abs/1504.03339).
- [572] **ACTPol** Collaboration, E. Schaan *et al.*, “Evidence for the kinematic Sunyaev-Zel’dovich effect with the Atacama Cosmology Telescope and velocity reconstruction from the Baryon Oscillation Spectroscopic Survey,” *Phys. Rev.* **D93** no. 8, (2016) 082002, [arXiv:1510.06442 \[astro-ph.CO\]](https://arxiv.org/abs/1510.06442).

- [573] J. C. Hill, S. Ferraro, N. Battaglia, J. Liu, and D. N. Spergel, “The Kinematic Sunyaev-Zel’dovich Effect with Projected Fields: A Novel Probe of the Baryon Distribution with Planck, WMAP, and WISE Data,” [arXiv:1603.01608 \[astro-ph.CO\]](https://arxiv.org/abs/1603.01608).
- [574] **DES, SPT** Collaboration, B. Soergel *et al.*, “Detection of the kinematic Sunyaev-Zel’dovich effect with DES Year 1 and SPT,” *Mon. Not. Roy. Astron. Soc.* (2016) , [arXiv:1603.03904 \[astro-ph.CO\]](https://arxiv.org/abs/1603.03904).
- [575] K. D. Irwin, “An application of electrothermal feedback for high resolution cryogenic particle detection,” *Applied Physics Letters* **66** no. 15, (1995) 1998–2000.
- [576] T. M. Lanting, K. Arnold, H.-M. Cho, J. Clarke, M. Dobbs, W. Holzapfel, A. T. Lee, M. Lueker, P. L. Richards, A. D. Smith, and H. G. Spieler, “Frequency-domain readout multiplexing of transition-edge sensor arrays,” *Nuclear Instruments and Methods in Physics Research A* **559** (Apr., 2006) 793–795.
- [577] K. D. Irwin, L. R. Vale, N. E. Bergren, S. Deiker, E. N. Grossman, G. C. Hilton, S. W. Nam, C. D. Reintsema, D. A. Rudman, and M. E. Huber, “Time-division SQUID multiplexers,” *Low Temperature Detectors* **605** (Feb., 2002) 301–304.
- [578] K. D. Irwin and K. W. Lehnert, “Microwave SQUID multiplexer,” *Applied Physics Letters* **85** (Sept., 2004) 2107.
- [579] J. W. Henning *et al.*, “Feedhorn-coupled TES polarimeter camera modules at 150 GHz for CMB polarization measurements with SPTpol,” *Proc. SPIE Int. Soc. Opt. Eng.* **8452** (2012) 84523A, [arXiv:1210.4969 \[astro-ph.IM\]](https://arxiv.org/abs/1210.4969).