THE UNIVERSITY OF ARIZONA
COLLEGE OF SCIENCE
Astronomy
& Steward Observatory

Director's Office
Steward Observatory
URL: www.as.arizona.edu

933 North Cherry Avenue
P.O. Box 210065
Tucson, AZ 85721-0065
Telephone: (520) 621-6524
buelljannuzi@email.arizona.edu

October 26, 2017

Dr. Denise Caldwell, Division Director, Division of Physics, NSF
Dr. Eric Saltzman, Section Head, Ant. Sciences, Div. of Polar Programs, NSF
Dr. James Siegrist, Associate Director of Science for High Energy Physics, DOE
Dr. Richard Green, Division Director, Division of Astronomy, NSF
Dr. Jim Ulvestad, Acting Assistant Director, Directorate for Mathematical and Physical Sciences, NSF

Dear Drs. Caldwell, Saltzman, Siegrist, Green, and Ulvestad:

In response to your letter of November 2016 requesting that the Astronomy and Astrophysics Advisory Committee (AAAC) establish a Cosmic Microwave Background Stage 4 Concept Definition Task force (CMB-S4 CDT) as a subcommittee of the AAAC to develop a concept for a CMB-S4 experiment, I appointed Charles Lawrence, Chief Scientist for Astronomy and Physics at the Jet Propulsion Laboratory, California Institute of Technology, as chair of the subcommittee.

Over the last nine months, he and his task force have expeditiously and quite impressively completed a tremendous amount of work to inform their response to your charge. Their report to the AAAC was provided to us on October 16, 2017. Members of the AAAC read the report and we sent questions to Charles Lawrence and his task force prior to our telecom with Charles on October 23rd. After a few minor revisions to clarify a couple of points in the report, the AAAC unanimously voted to accept the report of the Task Force. We hope their hard work meets your needs. A copy of the report accompanies this letter.

Sincerely,

Dr. Buell T. Jannuzi
Chair, Astronomy and Astrophysics Advisory Committee

Cc: Dr. Kathy Turner, DOE
Dr. Richard Barvainis, NSF
Dr. Paul Hertz, NASA

Jet Propulsion Laboratory
California Institute of Technology

4800 Oak Grove Drive
Pasadena, California 91109-8099

(818) 354-4321

23 October 2017

Dr. Buell Jannuzi
Chair, Astronomy and Astrophysics Advisory Committee

Dear Buell:

I am pleased to transmit to you the report of the CMB-S4 Concept Design Task Force.

The report presents a concept for implementing a ground-based "Stage-4" CMB experiment (CMB-S4), designed to tackle questions about fundamental physics from the earliest moments of the Universe through the epoch of dark energy domination.

The Report identifies three transformative science goals for CMB-S4:

- To measure the imprint of primordial gravitational waves on the CMB polarization anisotropy;
- To detect or strongly constrain departures from the thermal history of the Universe predicted by the standard model of particle physics; and
- To provide a legacy survey of nearly half the sky at centimeter to millimeter wavelengths.

We translate the first two goals into Science Requirements for CMB-S4, which we flow down to Measurement Requirements, Instrument Requirements, and Experiment Requirements. The third goal can be achieved with the same hardware required for the first two.

We develop a Strawperson Concept that meets the requirements, including operations and data handling and analysis, with schedule and cost including contingency appropriate for a concept design. This Strawperson is based on existing technology and computation and data-management models. We recommend investments to improve the reliability and production throughput of detector and readout components, and the continued development of a simulation framework to evaluate instrument designs and systematics.

On behalf of the CMB-S4 Concept Design Task Force,

Charles R. Lawrence
Chair

Cosmic Microwave Background Stage 4 Concept Definition Task Force

REPORT TO THE AAAC

23 October 2017

THE CMB-S4
CONCEPT DEFINITION TASK FORCE

Charles Lawrence, Jet Propulsion Laboratory, *Chair*

Julian Borrill, Lawrence Berkeley National Laboratory and University of California, Berkeley

John Carlstrom, University of Chicago and Argonne National Laboratory

Tom Crawford, University of Chicago

Mark Devlin, University of Pennsylvania

Jo Dunkley, Princeton University

Raphael Flauger, University of California, San Diego

Brenna Flaugher, Fermi National Accelerator Laboratory

Shaul Hanany, University of Minnesota

Kent Irwin, Stanford University and SLAC National Accelerator Laboratory

Bill Jones, Princeton University

Brian Keating, University of California, San Diego

John Kovac, Harvard University

Akito Kusaka, Lawrence Berkeley National Laboratory and University of California, Berkeley

Adrian Lee, University of California, Berkeley and Lawrence Berkeley National Laboratory

Jeff McMahon, University of Michigan

Michael Niemack, Cornell University

Steve Padin, University of Chicago

Clem Pryke, University of Minnesota

Suzanne Staggs, Princeton University

Edward Wollack, NASA/Goddard Space Flight Center

TABLE OF CONTENTS

1 EXECUTIVE SUMMARY **1**

2 SCIENCE JUSTIFICATION AND GOALS **5**
- 2.1 Overview . . . 5
- 2.2 Gravitational Waves and Inflation . . . 5
- 2.3 Light Relics . . . 6
- 2.4 Neutrino Masses . . . 8
- 2.5 Evolution of Cosmic Structure . . . 8

3 SCIENCE AND MEASUREMENT REQUIREMENTS **10**
- 3.1 Measurement Challenges . . . 10
- 3.2 Science Traceability Matrix . . . 11

4 INSTRUMENT DESIGN **13**
- 4.1 Sites . . . 13
- 4.2 Heritage and Design Considerations . . . 14
- 4.3 CMB-S4 Strawperson Concept . . . 17

5 OPERATIONS **18**

6 DATA MANAGEMENT **19**
- 6.1 Overview . . . 19
- 6.2 Computational Resources . . . 20
- 6.3 Instrument Data . . . 20
- 6.4 Time Domain . . . 22
- 6.5 Science Analyses . . . 23
- 6.6 Simulations . . . 23
- 6.7 Publication . . . 24

7 COLLABORATION STRUCTURE **24**

8 COMPLEMENTARITY WITH MEASUREMENTS FROM SPACE AND BALLOONS **25**

9 TECHNOLOGY DEVELOPMENT **27**

10 STAGING AND SCHEDULE **28**

11 COST **29**
- 11.1 Approach . . . 29
- 11.2 Basis of Estimate . . . 31
- 11.3 Cost for Strawperson Concept . . . 32
- 11.4 Cost Range . . . 33

REFERENCES **34**

A SIMULATIONS **36**
A.1 Small-Area Survey Targeting the Degree-Scale Signature of Gravitational Waves . . 37
A.1.1 Map noise simulations . . . 37
A.1.2 Foreground models . . . 38
A.1.3 Instrumental systematics . . . 39
A.1.4 Delensing . . . 39
A.1.5 Analysis methods . . . 40
A.1.6 Results . . . 40
A.2 Large Area Survey . . . 43

B TECHNOLOGY DEVELOPMENT **44**
B.1 Focal-Plane Sensors and Readouts . . . 44
B.2 Focal-Plane Optical Coupling . . . 44
B.3 Telescopes . . . 45
B.4 Camera Cryogenics and Optics . . . 45
B.5 Data Management . . . 46

C THE CONCEPT DEFINITION TASKFORCE **47**
C.1 Formation and Purpose . . . 47
C.2 Process . . . 47

D CHARGE TO THE CDT **49**

1 EXECUTIVE SUMMARY

This report presents a concept for implementing a ground-based Cosmic Microwave Background Stage-4 experiment (CMB-S4) designed to tackle questions about fundamental physics from the earliest moments in the Universe through to the epoch of dark energy domination. CMB-S4 will search for inflation at one end of the energy scale, and for sub-eV particles at the other, including neutrinos and as-yet-undiscovered light relics. CMB-S4 data will trace out the behavior of gravity across cosmological distances and anchor our understanding of how structures form under the gravitational collapse of dark matter moderated by the influence of baryons. These aims require CMB measurements with sensitivity, precision, and accuracy far beyond those obtained to date, roughly two orders of magnitude more sensitive than the summation of all measurements acquired so far, and an order of magnitude more sensitive than Stage-3 experiments. To address this challenge, the CMB community has come together to advocate a single comprehensive CMB-S4 experiment, which has been endorsed by the 2014 report of the Particle Physics Project Prioritization Panel (P5) *Building for Discovery* and the 2015 NAS/NRC report *A Strategic Vision for NSF Investments in Antarctic and Southern Ocean Research*.

The charge to the Concept Definition Task Force (CDT) included the following tasks: delineating transformative science goals unique to CMB-S4, along with the instrumental requirements needed to achieve them; developing a strawperson design concept, including plans for its implementation and a cost estimate range; evaluating the readiness of the technology and prioritizing development areas for investment; and developing models for data management and collaboration. This report presents the response of the CDT to the charge. The concept presented here is supported by work from a broad community of CMB scientists. The report builds on community efforts as represented in the first editions of the CMB-S4 Science Book and the CMB-S4 Technology Book. In defining the concept, the CDT often consulted with experts from the community.

The two key elements of the CDT concept for CMB-S4 are; 1) it requires multiple cameras and telescopes distributed across two sites; and 2) the experiment will be undertaken by a single collaboration and run as one project. The magnitude of the increase in science reach and complexity over existing CMB projects mandates the latter.

The CDT has identified the following transformative science goals for CMB-S4. The first two translate to requirements that cross clear, critical science thresholds, which cannot be achieved with the upgrades planned for the current (Stage-3) CMB experiments.

- The first goal and requirement for CMB-S4 is to measure the imprint of primordial gravitational waves on the CMB polarization anisotropy, quantified by the tensor-to-scalar ratio r. Specifically, CMB-S4 will be designed to provide a detection of $r \geq 0.003$. In the absence of a signal, CMB-S4 will be designed to constrain $r < 0.001$ at the 95% confidence level, nearly two orders of magnitude more stringent than current constraints. This will test many of the simplest models of inflation, including those based on symmetry principles, that occur at high energy and large inflaton field range. The r requirements have been translated into measurement requirements consistent with projecting out foregrounds and other contamination as detailed in Appendix A.
- The second goal and requirement for CMB-S4 is to detect or strongly constrain departures from the thermal history of the Universe predicted by the standard model of particle physics. Many well-motivated extensions of the standard model to higher energies predict low-mass relic particles. Departures from the standard history are conveniently quantified by the contribution of light relic particles to the effective number of relativistic species in the early Universe, $N_{\rm eff}$. CMB measurements are sensitive to the contribution of relic particles to the

energy density in the early Universe and therefore only depend on their interaction cross sections with standard model particles through the temperature at which the light relics decouple from the standard model. CMB-S4 will constrain $\Delta N_{\rm eff} \leq 0.06$ at the 95% confidence level allowing detection of, or constraints on, a wide range of light relic particles even if they are too weakly interacting to be detected by lab-based experiments. CMB-S4 will be the most robust and precise probe of the thermal history of our Universe, and will improve bounds on the decoupling temperature compared to Stage-3 CMB experiments or planned large scale structure surveys by a factor of a few for scalars, pushing it above the QCD phase transition, and two orders of magnitude or more for particles with spin.

- The third goal for CMB-S4 is to provide a unique legacy survey of nearly half the sky at centimeter to millimeter wavelengths. This powerful survey will complement and enhance the Large Synoptic Survey Telescope (LSST) optical survey of the same region, as well as other planned and yet-to-be-imagined surveys and data from both ground- and space-based instruments. Gravitational interactions with dark matter and electromagnetic interactions with baryonic matter leave signatures on the CMB as it traverses the expanse of the Universe. CMB-S4 maps will thus provide highly complementary data for investigations of dark energy, modifications to general relativity, and neutrino properties. For example, the CMB-S4 legacy survey will enable two independent determinations of the sum of the neutrino masses, using weak gravitational lensing and the evolution of the number density of galaxy clusters. The maps will also provide a unique and powerful probe of the influence of baryonic feedback on the formation of galaxies and clusters of galaxies. This third CMB-S4 science goal will have the broadest benefit to both the cosmological and astronomical communities. The design of CMB-S4 should preserve the ability to provide this legacy science as much as possible without substantially impacting the project cost or compromising the performance for r and $N_{\rm eff}$.

The first two science requirements drive the measurement requirements, which also satisfy those for the last science goal. The science and measurement requirements are given in a Science Traceability Matrix in the report. The main findings for the latter are:

- Deep measurements (noise levels $< 1\mu$K-arcmin) over exceptionally low foreground regions covering 3% to 8% of the sky are required to meet the goals for r. These measurements must have high fidelity and low contamination over a wide range of angular scales and frequencies. Large-angular-scale measurements with resolution of order $30'$ at 90 GHz and well determined beam properties and excellent control of systematic contamination are needed to image the B-mode polarization signature of the primordial gravitational waves. Small-angular-scale measurements with resolution of order $1\farcm5$ at 150 GHz are needed for removing the contamination of the degree-scale B modes caused by gravitational lensing of the much stronger CMB E-mode polarization ("delensing").
- Multifrequency coverage is required for foreground mitigation. As current measurements have shown, the rms fluctuation in Galactic foreground emission is an order of magnitude larger than the predicted B-mode fluctuations for $r = 0.001$, even in the cleanest regions of sky. Simulations based on the current best knowledge of the dust and synchrotron foreground emission indicate CMB-S4 can meet its r goals using nine frequency bands spanning 20 GHz to 270 GHz for the low-angular-resolution measurements. Fewer bands are needed for the high-angular-resolution delensing, $N_{\rm eff}$, and sky survey data.
- High-resolution ($\leq 1'.5$) measurements at a noise level of $\sim 1\mu$K-arcmin over a sky coverage of 40% are required to meet the $N_{\rm eff}$ goals.

The CDT has defined and costed a strawperson design concept that will meet these measurement requirements and therefore achieve the CMB-S4 science goals. In the development of the design concept, the CDT made extensive use of simulations based on our current understanding of the expected level and complexity of the foreground emission. Noise levels in the simulations are based on assumptions for the instrumental performance that are directly scaled from existing CMB instruments. These simulations are described in more detail in Appendix A. The strawperson design concept is based around two surveys with instrumentation configured as follows:

- An r survey covering 3% to 8% of the sky to be conducted using 1) fourteen 0.5-m aperture cameras, each measuring in two of the eight targeted frequency bands between 30 and 270 GHz; and 2) a 6-m class telescope equipped with detectors distributed over seven bands from 20 GHz to 270 GHz. Measurements at degree angular scales and larger by cameras with ~ 0.5 m apertures have been demonstrated to deliver high-fidelity, low-contamination polarization measurements at these scales. The combination of small-aperture cameras with the 6-m telescope therefore provides low-resolution B-mode measurements with excellent control of systematic contamination, as well as the high-resolution measurements required for delensing. Simulations indicate that a four-year survey[1] with the above configuration should be sufficient to achieve the $r < 0.001$ upper limit requirement or a 5σ detection for $r \geq 0.004$. Up to an additional four years are estimated to be required to achieve a 5σ detection for $r \geq 0.003$. The total detector count is of order 170,000 for the 0.5-m cameras and 70,000 for the 6-m telescope, with the bulk of the detectors allocated to the 85 GHz to 270 GHz bands.
- An $N_{\rm eff}$ survey covering 40% of the sky to be conducted over seven years using two 6-m class telescopes, each equipped with order 70,000 background-limited detectors distributed over seven frequency bands spanning 20 GHz to 270 GHz, with the bulk of the detectors allocated to the 95 GHz and 145 GHz bands.

The surveys will be conducted from both the Chilean and South Pole sites. The South Pole site provides superior atmospheric stability, and its unique geographical location enables continuous observation of a small low-foreground patch of the sky but restricts the total sky coverage available. The Chilean site is also exceptional for mm-wave observations and provides access to the 40% sky coverage required to meet the $N_{\rm eff}$ and legacy science goals. The optimal distribution of the 6-m and small-aperture cameras between the two sites has not been determined and has little effect on cost.

The CDT developed a comprehensive cost-estimating tool to capture all aspects of the experiment, with fidelity appropriate for a concept design, taking care not to miss any scope that might drive the cost. The CDT costing exercise included an optimization of cost versus experiment configuration, identified the major cost drivers, and is informative for the planning of pre-construction technology development. A preliminary estimate of the operations cost was also developed. A panel of DOE experts reviewed the CDT costing approach on 12 September 2017 and recommendations from that review have been included in the cost estimate.

The total project cost for the CDT design concept is \$412 M in 2017 dollars, including 45% contingency. The construction project effort is 21%, 32%, and 10% of the total project cost for scientists (non-DOE lab), engineers, and technicians, respectively. The top-level management effort, which includes the project manager, project controls, agency reviews, safety, quality assurance, and outreach, is estimated at 10% of the total project cost, with 10% contingency. The cost assumptions

[1] Here and throughout this report, observing times are given in "wall-clock" years. All overheads, inefficiencies, site, and instrumental factors are included in the calculations of data quality that would be achieved. See the right-hand column of Table 1 and Appendix A for details.

do not include potential savings from the use of legacy instrumentation that may be available at each site.

The CDT finds that CMB-S4 can be developed based on existing technology and computation and data-management models used in Stage-3 experiments. Pre-project investments in technology development will reduce risks in cost, schedule, and performance. The CDT recommends investments to improve the reliability and production throughput of detector and readout components, and the continued development of a simulation framework to evaluate instrument designs and systematics.

The CDT finds that the project timeline also benefits from the prior experience with designing and deploying CMB experiments. The project schedule to deliver the strawperson concept allows two years for design and development, four years for construction, and two years for commissioning. Allowing for some overlap of these phases, the CDT baselines seven years to deliver the project for the start of operations.

The design concept is highly modular and thus allows flexibility in staging the project implementation. Funding considerations and the specific prioritization requirements for the agencies will impact the timeline, however, with DOE and NSF having different constraints. It is therefore advantageous to work with DOE and NSF to develop a staged implementation plan that leads to the full CMB-S4 program, and the report gives one such example.

Lastly the report offers suggestions for establishing the CMB-S4 collaboration and its relationship to the CMB-S4 project. The CDT notes that the informal grass-roots CMB-S4 collaboration has been highly effective, and is pleased to report that the collaboration is taking steps to establish the formal collaboration over the next several months.

2 SCIENCE JUSTIFICATION AND GOALS

2.1 Overview

Since the first detection of the cosmic microwave background (CMB) over 50 years ago, CMB measurements have continuously transformed our understanding of the early Universe. Measurements of the CMB using ground-based, balloon, and satellite experiments have provided conclusive evidence that our Universe evolved into its present state from an early hot, dense state; have determined the age and the composition of our Universe at the percent level; and have provided the strongest evidence that dark matter cannot consist of non-luminous baryonic matter. Furthermore, observations of the polarization of the CMB have demonstrated in a model-independent way that the observed fluctuations in CMB intensity were already present around the time of recombination and, according to Einstein's theory of general relativity, must have been generated by some process long before the moment when our Universe became filled with a hot and dense plasma (i.e., the "hot big bang"). As a consequence, measurements of the CMB provide us with a remarkable opportunity to study the earliest moments of our Universe.

Observations have nearly exhausted the information accessible in primary temperature anisotropies, and Stage-3 experiments have just begun to make precise measurements of the polarization of the CMB as well as measurements of secondary effects such as the weak gravitational lensing of the CMB by large-scale structure and the thermal and kinetic Sunyaev-Zel'dovich effects (tSZ and kSZ). CMB-S4 is the natural next step for the CMB program. It will exploit the enormous potential of CMB measurements to once again transform our understanding of the early Universe and particle physics, fulfilling the goals set out in the 2010 Astronomy and Astrophysics Decadal Survey and the 2014 report of the Particle Physics Project Prioritization Panel to search for primordial gravitational waves as well as light relic particles. In doing so, CMB-S4 will also constrain neutrino properties and provide critical measurements of the evolution of cosmic structure from the early universe to the present day, advancing a key goal of the 2010 Decadal Survey.

2.2 Gravitational Waves and Inflation

CMB-S4 will be exquisitely sensitive to the imprint of primordial gravitational waves left on the CMB. According to Einstein's general theory of relativity, any gravitational waves present at recombination with wavelengths observable by CMB-S4 must have been generated before the hot big bang, and are pristine relics from the primordial Universe. Because these gravitational waves are fluctuations in the metric of spacetime itself rather than fluctuations in the density, a detection would open a new window into the early Universe and would transform our understanding of several aspects of fundamental physics. For the foreseeable future, precision measurements of the CMB polarization are our only way to detect primordial gravitational waves.

Inflation, a period of accelerated expansion of the early Universe, is the leading paradigm for explaining the origin of the primordial density perturbations that grew into the CMB anisotropies and eventually into the stars and galaxies we see around us. The majority of inflationary models predict that the primordial density perturbations are adiabatic, Gaussian, and nearly scale-invariant, in agreement with existing observations.

In addition to creating primordial density perturbations, the rapid expansion of spacetime creates primordial gravitational waves that imprint a characteristic polarization pattern onto the CMB. Many models of inflation predict a signal large enough to be detected with CMB-S4. Because the polarization pattern is due to quantum fluctuations in the metric of spacetime generated during inflation, it would provide insights into the quantum nature of gravity. The strength of the signal is encoded in the tensor-to-scalar ratio r, the ratio of the power stored in gravitational waves relative

to that in the already measured density perturbations. Its value provides a measurement of the expansion rate of the Universe during inflation. The Friedmann equation would then provide us with the inflationary energy scale. CMB-S4 will be able to detect primordial gravitational waves for $r > 0.003$, and a detection would point to inflationary physics near the energy scale associated with grand unified theories. As a consequence, a detection would provide additional evidence in favor of the idea of the unification of forces, and would probe energy scales far beyond the reach of the LHC or any conceivable collider experiment. The knowledge of the energy scale of inflation would have broad implications for many other aspects of fundamental physics, including ubiquitous ingredients of string theory like axions and moduli, fields that control the shapes and sizes of the compact dimensions.

In the absence of a detection, the upper limit of $r < 0.001$ at 95% CL achievable by CMB-S4 is nearly two orders of magnitude stronger than current limits, and an order of magnitude stronger than projected limits for Stage-3 experiments. The CMB-S4 measurement would significantly advance our understanding of inflation. It would rule out large classes of inflationary models and dramatically impact how we think about the theory. To some, the remaining class of models would be contrived enough to give up on inflation altogether.

Furthermore, CMB-S4 is in a unique position to probe the statistical properties of primordial density perturbations through measurements of primary anisotropies in the temperature and polarization of the CMB on small angular scales with unprecedented precision. The Planck satellite has recently detected the departure from scale-invariance of the power spectrum of primordial density perturbations, assuming that a power law characterized by a spectral index n_s. CMB-S4 would decrease the uncertainty on the spectral index of density perturbations by more than a factor of two, would similarly improve existing constraints on the so-called running, the scale-dependence of the spectral index, by a factor of two to three, and would tighten departures from adiabaticity by as much as a factor of five.

The degree of Gaussianity of the perturbations provides a direct probe of the couplings and field content of inflation. Planck has recently placed tight constraints on departures from Gaussianity. CMB-S4 alone will tighten constraints on all types of non-Gaussianity by factors of two to three. By itself it is not able to improve the current constraints on local non-Gaussianity of $f_{\rm NL}^{\rm local} = 0.8 \pm 5$ to the levels predicted by many models. However, the combination of the mass map made by CMB-S4 using gravitational lensing of the CMB with data from wide area, photometric galaxy surveys has the potential to improve the constraints on local non-Gaussianity by an order of magnitude or more. A discovery of $f_{\rm NL}^{\rm local} > 1$ would rule out all inflation models with only a single "clock", and shed light on models where light degrees of freedom other than the inflaton contribute to the observed fluctuations. Many alternative models predict $f_{\rm NL}^{\rm local} \sim O(1)$, while the most popular models predict $f_{\rm NL}^{\rm local} \ll 1$.

2.3 Light Relics

Light particles that are stable on cosmological time scales appear in many attempts to understand both the observed laws of physics and their extensions to higher energies. These light particles are often deeply tied to the underlying symmetries of nature and can play crucial roles in understanding some of the major outstanding problems in physics. In most cases, these particles interact too weakly to be produced at an appreciable level in Earth-based experiments, making them experimentally elusive. At the very high temperatures believed to be present in the early Universe, however, even extremely weakly coupled particles can be produced prolifically and can reach thermal equilibrium with the Standard Model particles. Light particles (masses less than 0.1 eV) produced at early times survive until the time when the CMB decoupled from the baryons

and observations become possible. Neutrinos are one example of such a relic found in the Standard Model. Extensions of the Standard Model also include a wide variety of possible light relics including axions, sterile neutrinos, hidden photons, and gravitinos. As a result, the search for light relics from the early Universe with CMB-S4 can shed light on some of the most important questions in fundamental physics, complementing existing collider searches and efforts to detect these light particles in the lab.

Light relics contribute to the total energy density in radiation in the Universe and significantly alter the power spectrum of the CMB at small angular scales (high multipole number ℓ). The energy density in radiation controls both the expansion rate of the Universe at that time and the fluctuations in the gravitational potential in which the baryons and photons evolve. Through these effects, CMB-S4 can provide an exquisite measurement of the total energy density in light, weakly coupled particles.

The contribution to the energy density from relativistic particles is conventionally parametrized by the quantity $N_{\rm eff}$, defined so that a species of light relics that was in equilibrium with the standard model particles and decoupled just before electrons and positrons annihilate contributes $\Delta N_{\rm eff} = 1$. In this convention, the contribution from each family of active neutrinos is $\Delta N_{\rm eff} \approx 1.015$, just above unity both because of QED corrections and because neutrinos have not quite decoupled when electron-positron freeze-out begins.

After a species of light relics decouples from the standard model, its temperature redshifts like the inverse of the scale factor. The weaker their interactions with the standard model, the earlier light relics will decouple. The earlier decoupling occurs, the more species of standard model particles will freeze out after the light relic decouples, heating the remaining species of standard model particles relative to the light relics. As a consequence, the more weakly a particle couples to the standard model, the higher its decoupling temperature and the smaller its contribution to $N_{\rm eff}$. Thus, precise measurements of $N_{\rm eff}$ provide an opportunity to detect light relic particles, or to constrain their properties.

Data from the Planck satellite provide constraints of $\Delta N_{\rm eff} < 0.4$ at 95% CL and imply that no light relic can interact with the standard model strongly enough to remain in equilibrium below 40 MeV. Future large scale structure surveys and Stage-3 CMB experiments are expected to improve this constraint to $\Delta N_{\rm eff} \lesssim 0.13$ at 95% CL and will exclude light relics that are in equilibrium with the standard model at temperatures below ~ 170 MeV. With the ability to constrain the contribution from light relics to $\Delta N_{\rm eff} < 0.06$ at 95% CL, CMB-S4 will provide the most precise and robust constraints and will extend our knowledge to temperatures beyond the QCD phase transition and exclude any light relics that decouple below 400 MeV. Particles with spin are more strongly constrained and the improvements are more dramatic. Whereas Stage-3 experiments would exclude particles with spin that decouple below 260 MeV, CMB-S4 will improve this constraint by over two orders of magnitude and exclude spin-1/2 particles below 30 GeV and has the potential to disfavor light relics of spin 1 entirely.

In general, these constraints are model-dependent and could be altered if there were degrees of freedom beyond the standard model that annihilate or decay into standard model particles after the light relics decouple. States that annihilate into standard model particles at temperatures below the electroweak scale are strongly constrained by the LHC, so that the limits on spin-0 and spin-1/2 particles are robust. CMB-S4 in combination with data from the LHC has the capability to exclude light relics of spin-1 that decouple after the electroweak phase transition even if they interact too weakly to be detected at the LHC. Late-decaying particles are less constrained and could dilute the effect of a relic particle if they exclusively decayed into standard model particles. However, in many models including models inspired by string theory in which the late-decaying particles are moduli, the effect is to increase $N_{\rm eff}$, so that CMB-S4 will place tight constraints on any such model. More

generally, CMB-S4 will tightly constrain any departure from the thermal history expected in the standard model, providing order-of-magnitude improvements on the decoupling temperature over Stage-3 CMB experiments as well as current and planned large-scale structure surveys.

2.4 Neutrino Masses

Neutrinos are the least explored corner of the Standard Model of particle physics. The 2015 Nobel Prize recognized the discovery of neutrino oscillations, which shows that neutrinos have mass. However, the overall scale of the masses and the full suite of mixing parameters are still not measured. A major effort is underway to study their properties in short- and long- baseline as well as neutrinoless double beta decay experiments. Short- and long- baseline experiments probe different mass-square differences, while neutrinoless double beta decay experiments probe a specific combination of the entries of the neutrino mass matrix. Cosmology offers a unique and complementary view of neutrinos, because they were produced in large numbers in the high temperatures of the early Universe and left a distinctive imprint in the CMB and the large-scale structure of the Universe. Cosmological probes are sensitive to the sum of the neutrino masses. Therefore, CMB-S4 and large-scale-structure surveys together will have the power to detect properties of neutrinos that supplement those probed by terrestrial experiments. In the simplest model with three active neutrinos, CMB-S4's lensing measurements are redundant and allow for interesting consistency checks. In models that contain additional sterile neutrinos, the measurements provide completely complementary information and can potentially distinguish between different models.

Baryon Acoustic Oscillation (BAO) surveys such as DESI, PFS, Euclid, and WFIRST complement CMB-S4's lensing measurements by breaking the degeneracy of the sum of neutrino masses with Ω_m. The precision of CMB-S4 is approximately a factor of two greater than that for Stage-3 experiments (with current BAO data) and will allow for a measurement of the sum of neutrino masses through weak gravitational lensing of the CMB even for the minimum mass in the normal mass hierarchy. The measurement will be limited by our knowledge of the optical depth to recombination τ, so that more precise measurements of τ by experiments like CLASS or LiteBIRD would further increase the sensitivity to the sum of neutrino masses through weak lensing. In addition, CMB-S4 will independently measure the sum of neutrino masses through cluster counts, with comparable sensitivity.

2.5 Evolution of Cosmic Structure

CMB-S4 is a unique and powerful tool to study the evolution of cosmic structure because of two key capabilities. It will be able to trace, with unprecedented precision, the distribution of normal, baryonic matter in its dominant, ionized phase, using the tSZ effect, and it will be able to map the total mass distribution (dominated by dark matter) through the lensing of CMB anisotropies.

Utilizing the redshift independence of tSZ measurements, CMB-S4 will trace the evolution of massive galaxy clusters, from the first objects formed at $z \sim 3$ to the present day. The abundance of massive clusters is very sensitive to the amplitude of matter perturbations as a function of cosmic time, $\sigma_8(z)$, and through this to a broad range of cosmological parameters, such as those describing the mean matter and dark energy densities, the dark energy equation of state, modifications to gravity, and the species-summed neutrino mass. Key to using clusters as cosmological probes is accurate mass calibration. Using CMB lensing, CMB-S4 will provide percent-level mass calibration, extending (and cross-checking) the work of LSST at $z < 1$ out to the highest redshifts at which clusters exist, thereby opening up the full redshift lever arm of clusters for cosmological work.

CMB-S4 will also constrain cosmological parameters by measuring the power spectrum of matter perturbations through CMB lensing. While CMB lensing alone provides an integrated measure of

the matter power spectrum, cross-correlation with other astrophysical tracers, such as the galaxy catalogs from LSST, WFIRST and Euclid, can yield both higher signal-to-noise ratios and enable reconstruction as a function of redshift. In this way, CMB-S4 will extend power spectrum measurements from cosmic shear lensing surveys to much higher redshifts, potentially to $z \sim 7$ with appropriate astrophysical tracers.

CMB-S4 will probe the relation between baryons and dark matter in the cosmic web. Cross-correlations of CMB lensing with galaxy catalogs will also determine the galaxy-mass cross power spectrum, establishing constraints on galaxy bias and informing galaxy formation models. By stacking observations of the tSZ and kSZ on objects found in other wide-field surveys (e.g., LSST, WFIRST, and Euclid), CMB-S4 will measure the thermal content and spatial distribution of the circumgalactic medium for galaxies, groups, and clusters as a function of redshift, out to redshifts extending beyond the peak of cosmic star formation ($z > 2$). CMB-S4 will quantify dust production and star-formation activity, linking such measurements to host dark matter halo mass and the local dark matter environment through CMB lensing measurements. As a result, CMB-S4 will provide unique information on astrophysical feedback mechanisms, the largest uncertainty in models of galaxy formation. These measurements are complementary to future X-ray observatories such as Athena and eROSITA. While X-ray observations will provide detailed thermodynamic measurements in the inner regions of individual groups and clusters, CMB-S4 will trace the mean properties of these systems out to many virial radii. Astrophysical feedback is also the largest source of theoretical systematic uncertainty for all Stage-4 weak lensing surveys, including LSST, Euclid, and WFIRST. Feedback effects can alter the matter power spectrum by 20–30% over the range of scales and redshifts that weak lensing surveys will study with (galaxy) lensing. CMB-S4 can help to resolve this issue by directly measuring the baryon profiles of the halos that dominate the signal, which are group-size and larger ($M > 10^{12.5} \mathrm{M}_\odot/h$).

Precise measurements of the kSZ effect will additionally enable us to map the momentum field of large-scale structure, providing independent constraints on cosmological parameters that are complementary to those from density fluctuation measurements such as cluster counts. The mean pairwise velocity of galaxy clusters, for example, will provide a precise probe of gravity on large scales, while the kSZ power spectrum will constrain the time and duration of reionization.

Finally, the CMB-S4 catalogs of galaxy clusters and high-redshift galaxies will provide the definitive target list for a broad range of astrophysics at other wavelengths, providing strong, natural synergies with leading contemporary observatories such as Athena, CHIME, DESI, eROSITA Euclid, LSST, PFS, SKA, and WFIRST. Here the value of the cluster catalog will be primarily driven by the clusters at the highest redshifts. CMB-S4 will further push millimeter galaxy counts to flux densities below 10 mJy to reveal and study the unlensed population. With sufficient angular resolution and sensitivity, CMB-S4 will also make the first detection of the polarized SZ signal sourced by the scattering of the remote CMB quadrupole off the free electrons located in galaxy clusters. This signal holds great promise for both future astrophysical applications (e.g., probing the ionized gas distribution in halos as a function of mass and redshift) and cosmological constraints (e.g., measuring the optical depth τ with an uncertainty smaller than the primary CMB cosmic variance limit via reconstruction of the E-mode power spectrum using the remote quadrupole information from the polarized SZ).

3 SCIENCE AND MEASUREMENT REQUIREMENTS

3.1 Measurement Challenges

The ambitious science goals laid out in the previous section will require significant advances, not only in raw sensitivity but also in control of foregrounds and instrumental systematics. As discussed in detail in the CMB-S4 Science Book (Abazajian et al. 2016) and in previous publications (Abazajian et al. 2015a,b), attaining the desired levels of sensitivity to the signatures of gravitational waves and light relics in the CMB necessitates first and foremost at least an order-of-magnitude increase in the raw number of detectors on the sky compared to Stage-3 experiments. These works also noted that foreground mitigation will be crucial for CMB-S4, particularly in the pursuit of the gravitational-wave signal. To minimize contamination from Galactic foregrounds, it is clear that multiple frequency channels are required. It is known from analysis of BICEP/Keck and Planck data (BICEP2, Keck Array and Planck Collaborations 2015) that, if unsubtracted and unmodeled, Galactic dust imparts a bias to the measurement of the tensor-to-scalar ratio r at 150 GHz at a level two orders of magnitude above the target $\sigma(r)$ for CMB-S4. Synchrotron emission is expected to impart a similar bias at lower frequencies. At least one frequency channel is required to remove each of these contaminants, and multiple channels will be needed for each component if the behavior of these foregrounds is not perfectly uniform across the survey. Galactic foregrounds are also reduced by observing a patch or patches of sky with as low as possible column density of Galactic material. This consideration—and optimizations of raw sensitivity—drive a survey for the degree-scale gravitational-wave signal to small patches of sky ($<$10%).

The other major foreground to gravitational-wave searches, one which cannot be mitigated with frequency coverage, is the signal from gravitational lensing. Scalar density perturbations in the early Universe produce only even-parity ("E-mode") patterns in the CMB polarization to first order, while gravitational-wave perturbations also produce odd-parity ("B-mode") patterns (Kamionkowski et al. 1997; Seljak & Zaldarriaga 1997). This fact is essential to our ability to detect a gravitational-wave signal in the presence of the much larger signal from scalar density perturbations. Gravitational lensing of the E modes, however, produces a secondary source of B modes, at an amplitude significantly larger than the target CMB-S4 sensitivity to the gravitational-wave signal at degree scales. The lensing process can effectively be inverted with sufficiently high-fidelity measurements of the E-mode signal and the gravitational lensing potential, and a map-level prediction of the lensing contamination can be accounted for in gravitational-wave searches (Knox & Song 2002; Kesden et al. 2002). The level of this so-called "delensing" required for a small-area survey and the CMB-S4 target $\sigma(r)$ requires deep, high-resolution data over the same patch of sky observed for the degree-scale B modes.

The CMB-S4 science target for the light relic search, meanwhile, drives the survey to large fractions of the sky ($\gg$10%). Expectations for the small-angular-scale polarized power spectra required for this science are that the Galactic foregrounds are less of an issue and that the extragalactic foregrounds are 1) largely unpolarized, and 2) amenable to being modeled and subtracted in power spectrum space. As a result, it is likely that less weight will need to be given to foreground-monitoring frequency bands than is required for the gravitational-wave search. The particular experimental systematics relevant to the light-relic search are also not necessarily the same as those most important for the gravitational-wave search. These considerations tend to drive the design of CMB-S4 to two surveys, one over a large fraction of the sky and one over $<$10% of the sky.

All of these considerations were taken into account in the optimization exercises described in the CMB-S4 Science Book, and the measurement and instrument requirements driven by the r and

$N_{\rm eff}$ goals have not changed significantly between that document and this report. The strawperson configuration resulting from the latest generation of those optimizations has been checked against map-level simulations for fidelity, including the effects of instrumental systematics, as described in Appendix A of this document.

Finally, to enable the cosmology and astrophysics from the CMB-S4 probes of large-scale structure discussed in the previous section, it is essential that CMB-S4 have sufficient sky area ($f_{\rm sky} \sim 0.4$), spatial resolution ($\sim 1'$ at 150 GHz), depth ($\sim 1\,\mu$K-arcmin at 90 and 150 GHz), and frequency coverage (sensitivity extending to at least 270 GHz) to detect and separate the relevant tSZ, kSZ, dust, and CMB cluster lensing signals. The measurement/instrument requirements for this science largely overlap with those for the light-relic search, but it will be necessary to keep the large-scale-structure science in mind independently as one explores the parameter space of the large-sky-fraction survey.

3.2 Science Traceability Matrix

The quantitative science goals in § 1, and the measurement challenges discussed in the previous subsection, are translated into science requirements, measurement requirements, and instrument requirements in the Science Traceability Matrix (STM, Table 1; see also Table 2). The translation is accomplished using simulations that incorporate the best present knowledge of foregrounds, instrument characteristics and systematic errors, and the relevant characteristics of the telescope sites, as described in Appendix A. Some points to note in the STM are:

- The requirements on r and on $N_{\rm eff}$ drive the design of the instrument. Given the instrument required by r and $N_{\rm eff}$, much other science can be done.
- The science requirement on r is written quantitatively in two different ways, as the statistics are different for upper limits than for detections. If r is not detected (i.e., $r = 0$ in simulations), we require CMB-S4 to be able to measure $\sigma(r) \leq 0.0005$ in four years, giving an upper limit on r of $r < 0.001$ at 95% confidence. Alternatively, with the same noise level from all sources, including foreground residuals and uncorrected systematic errors, CMB-S4 will be able to measure $r = 0.003$ at the equivalent of 4σ. If after four years such a detection has emerged, extending the observing time should be considered. An extension to eight years would increase the expected detection significance for $r = 0.003$ to 5σ. Beyond increasing the significance from 4σ to 5σ, extended observations would allow additional consistency tests (such as testing the detection with alternative sky coverage or experimental strategies) and a detailed characterization of the signal. For signals at the limit of the experiment's sensitivity, the extended observations would make the difference between a marginal detection and a detection that is universally regarded as conclusive. The additional operations cost for extending the observing time by four years would be about \$50 M.
- The science requirement on $N_{\rm eff}$, that $\Delta N_{\rm eff} \leq 0.06$ at the 95% confidence level, is for seven years of observing.
- Simulations show that multiple sets of measurements with somewhat different combinations of frequency, angular resolution, and noise satisfy the science requirements. In the Measurement Requirement column of Table 1, we give one such set for each of the driving requirements, chosen among alternatives generally for minimum cost.
- As choices between multiple sets of measurements that satisfy the science requirements are made, we require (lowest blue-shaded cell under Science Requirements) that galaxy cluster, neutrino, and other astrophysical and cosmological science return be maximized within the cost constraints and without compromising the performance for r and $N_{\rm eff}$.

Table 1: Science Traceability Matrix.

DOE & NAS SCIENCE GOALS (P5* Strategic Plan; New Worlds New Horizons 2010)	SCIENCE OBJECTIVE	SCIENCE REQUIREMENT	MEASUREMENT REQUIREMENT	INSTRUMENT REQUIREMENTS	EXPERIMENT REQUIREMENTS
"Undertand cosmic acceleration: dark energy and inflation."—P5 Support CMB experiments as part of the core particle physics program.	Test models of inflation by measuring or putting upper limits on r, the ratio of tensor fluctuations to scalar fluctuations.	If r= 0: $\sigma(r) \leq 0.0005$; r< 0.001, 95% confidence; in four years. If r= 0.003: measure at equivalent 4σ; in four years (see Section 3.2).	Measure Q and U over a sufficient fraction of the sky at a sufficient set of frequencies to control foregrounds with: a) sufficient control of low ell noise and systematics and b) sufficient resolution to allow delensing. One set of measurement specifications that satisfies this is: Measure Q and U over 3% of the sky at frequencies of 20, 30, 40, 85, 95, 145, 155, 220, and 270 GHz, with Q/U-map noise levels of 14, 8.7, 8.2, 1.6, 1.3, 2.0, 2.0, 5.2, 7.1 μK-arcmin, map noise 1/ell knees of ≤60 at the six highest frequencies and ≤60 at at least one of the three lowest frequencies, and angular resolutions of ≤15 arcmin at 150 GHz, scaled by wavelength to other bands except requiring <30 arcmin at at least one of the two lowest-frequency bands. For delensing, measure the same 3% of the sky at frequencies of 30, 40, 95, 145, 220, and 270 GHz, with angular resolutions of ≤3 arcmin at 95, 145, and 220 GHz and Q/U map noise levels of 7, 7, 1.0, 1.3, 7, and 7 μK-arcmin.	One instrument configuration that satisfies the measurement requirements is: 14 0.5-meter-aperture cameras with detectors distributed as Frequency: 30, 40, 85, 95, 145, 155, 220, 270 GHz # detectors: 260, 470, 17k, 21k, 18k, 21k, 34k, 54k plus higher-resolution channels at 20, 30, and 40 GHz from the delensing survey below. For delensing, one 6-meter-aperture telescope with detectors distributed as Frequency 20, 30, 40, 95, 145, 220, 270 GHz # detectors: 130, 250, 500, 25k, 25k, 8.7k, 8.7k	Four years of observing (wall-clock time). We do not explicitly specify here many characteristics of the instruments and experiment, such as bandwidth, optical efficiency, bad pixels, weather losses, Sun constraints, and so on. These are, however, built into the simulations that calculate the measurements that will be achieved (column D) by the hardware in column E, based on actual in-the-field performance achieved by BICEP/Keck and SPT at Pole, and ACT in the Atacama, over years of observing.
"What are the properties of neutrinos?"—NWNH2010 Discover the elementary constituents of matter and energy. "Explore the unknown: new particles, interactions, and physical principles."—P5	Determine the role of light relic particles in fundamental physics, and in the structure and evolution of the Universe.	$\Delta N_{\rm eff} \leq 0.06$, 95% confidence; in seven years.	Measure I, Q, and U over a sufficient fraction of the sky at sufficiently low noise and with sufficient frequencies to control foregrounds. One set of measurement specifications that satisfies this is: Measure 40% of the sky at frequencies of 40, 95, 150, 220, and 270 GHz, with angular resolution of <=1.5 arcmin at 150 GHz, and Q/U map noise levels of <=1.3 μK-arcmin when 95 and 150 GHz are optimally combined.	One instrument configuration that satisfies the measurement requirements is: Two six-meter-aperture telescopes with detectors distributed as Frequency: 20, 30, 40, 95, 145, 220, 270 GHz # detectors: 290, 640, 1.1k, 50k, 50k, 17k, 17k	Seven years of observing (wall-clock time). See additional information in the cell above.
		$\sigma(\Sigma m\nu)$ = 25 meV (lensing or clusters), with 0.006 < $\sigma(\tau)$ < 0.01 and (for lensing) errors on the BAO distance ratio parameter rs/DV as given by DESI Collaboration et al (2016).	Achieved with the same measurements as required for $N_{\rm eff}$, above.		
		CMB-S4 shall be designed to maximize the galaxy cluster, neutrino, and other astrophysical and cosmological science return without increasing the project cost or compromising the performance for r and $N_{\rm eff}$.			
"Understand cosmic acceleration: dark energy and inflation."—P5	Test models of dark energy and modified gravity by measuring the growth of cosmic structure	Measure $\sigma_8(z)$ in contiguous bins spanning redshifts 0<z<3 to a precision of 1–2% per bin, with at least 2 bins at z>1.5.	Achieved with the same measurements as required for $N_{\rm eff}$, above.		
"How do baryons cycle in and out of galaxies, and what do they do while they are there?"—NWNH2010 Cosmic Dawn: Origins: The origin of galaxies and large-scale structure	Understand the impact of feedback processes on the distributions of dark and baryonic matter in the Universe	Using the kSZ and tSZ effects, measure the baryon density and thermal energy profiles of halos of mass M > 10^12.5 Msun/h to 1% at z<1 and 2% (T redshift 1 < z < 3 on scales down to rmin = 300 kpc/h (kmax = 10 h/Mpc in Fourier space).	Achieved with the same measurements as required for $N_{\rm eff}$, above.		

Requirements that determine the design of the experiment are in boxes shaded blue.
Additional science that can be done with the instrument required by the blue boxes is shown in boxes shaded green.
*P5 = Particle Physics Project Prioritization Panel

- For r, at least one of the two lowest frequency bands should have angular resolution $< 30'$. The value of this is supported by simulations that include non-Gaussian, small-angular-scale structure in the synchrotron component.
- In the same way that multiple sets of measurements satisfy the science requirements, multiple instrument configurations can satisfy the measurement requirements. In the Instrument Requirements column, we give one such configuration for each of the driving requirements, chosen among alternatives generally for minimum cost.

4 INSTRUMENT DESIGN

4.1 Sites

State-of-the-art CMB observations require the highest and driest sites on Earth. The two best sites developed for CMB are South Pole in Antarctica and Cerro Toco in the Chilean Andes (Figure 1). Both of these sites have been in use for several decades and have hosted many CMB telescopes (Ogburn et al. 2012; Carlstrom et al. 2011; Fowler et al. 2007; Barron et al. 2014).

Figure 1: *Left:* the CMB telescope sector at South Pole Station in Antarctica. *Right:* CMB telescopes at Cerro Toco in the Chilean Andes.

At these sites, the atmosphere is almost transparent to microwave radiation in a series of windows bracketed by oxygen and water emission lines (Figure 2). These windows are 40 GHz and below, around 90 GHz, around 150 GHz, and from 200 to 300 GHz. Variation of the water content of the atmosphere leads to considerable variation of the residual atmospheric emission in these windows, especially at the higher frequencies. These fluctuations occur on all timescales: long-timescale variations lead to day-to-day changes in the white noise level, while short-timescale variations inject additional $1/f$ noise. Fortunately, atmospheric emission is almost completely unpolarized, so this $1/f$ (low frequency) noise can be suppressed by differencing pairs of detectors (Chiang et al. 2010) or modulation (Johnson et al. 2007; Kusaka et al. 2014). Figure 2 summarizes the situation.

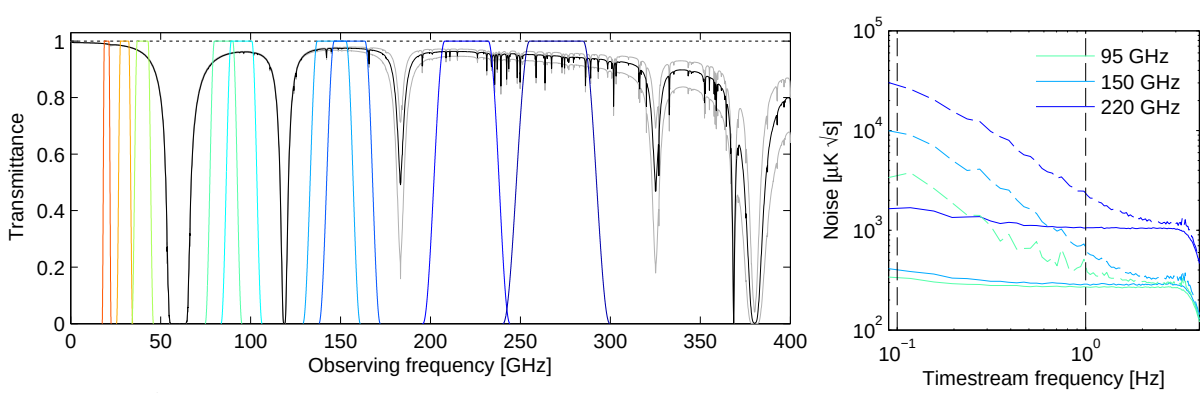

Figure 2: *Left:* Atmospheric transmission as a function of frequency showing the atmospheric windows and the passbands chosen for CMB-S4. *Right:* The power spectral density of detector timestream data showing that atmospheric $1/f$ noise is largely unpolarized (solid lines are polarization, dashed are total intensity).

away from the zenith the atmospheric column depth increases (as $\sec\theta$), and the atmospheric noise increases. In practice this means that observations from Pole can access only $< 10\%$ of the sky (excluding the plane of the galaxy). Due to its mid-latitude location (23° S), the Chilean site can observe $\sim 40\%$ of the sky with acceptable noise penalty. The combination of these two factors makes the use of both sites absolutely essential to achieve the core science goals of CMB-S4.

4.2 Heritage and Design Considerations

A wide range of instrument designs have been successfully used in previous CMB measurements, and several configurations have been studied by the CMB-S4 community (Abitbol et al. 2017; Hanany et al. 2013). The studied configurations use different combinations of telescope apertures from 0.3 m to 10 m. We categorize the instruments as "small-aperture cameras" with aperture diameters of order a meter or less, and "large-aperture telescopes" for instruments that are typically fed by cryogenic cameras positioned near a telescope focus. The smallest angular scale that can be resolved by a given aperture is in general proportional to the wavelength divided by the aperture size.

The instrument concept we focus on includes both small-aperture cameras to reconstruct large angular scales and large-aperture telescopes for higher-angular-resolution measurements. While there are no fundamental limits to the largest angular scales that can be measured with a large-aperture telescope, practical calibration and systematic control challenges exist. Thus far only small-aperture instruments, such as BICEP/Keck, have convincingly demonstrated the required systematic and calibration control on the angular scales required to constrain r. We adopt the conservative approach of assuming the large-aperture telescopes will not contribute directly to measuring the angular scales required to constrain r, which is the motivation for including both small and large apertures in the instrument concept.

To measure the large angular scales, approximately 0.5 m aperture cameras are required, which could use either refracting optics like BICEP/Keck or reflecting optics like ABS (Abitbol et al. 2017). In addition to lenses or mirrors, each of these cameras requires a vacuum window and several filters. The camera optical components require anti-reflection coatings to prevent reflections at percent or better levels. The BICEP3 and ABS camera designs as well as a BICEP Array mount design are shown as example concepts in Figures 3 and 4.

The level of measurement sensitivity required motivates use of multiple small-aperture cameras for CMB-S4. Using multiple cameras provides improved optical performance, and simplifies deployment of different frequency channels by reducing the bandwidth requirements of the optical components. If needed, the aperture size of these cameras can also be adjusted to maintain similar beam sizes at different frequencies.

To achieve the $1\farcm5$ measurement requirement at 150 GHz, a telescope aperture of at least 6 m is required. Several telescope designs for CMB-S4 have been studied by the community, including cross-Dragone (CD), three mirror anastigmat, Gregorian, and Cassegrain designs. A 6 m aperture CD design (Niemack 2016) has recently been adopted by both the CCAT-prime and Simons Observatory collaborations (Figure 5). The primary motivation for this design is its high optical throughput and low levels of polarization systematics, in comparison to existing large aperture CMB telescopes. These properties enable deployment of many more detectors on a single telescope than currently possible. In this design, the camera is aligned with the elevation axis, which enables an instrument rotator to be used to either track the telescope in elevation or to rotate the camera into different orientations for testing and mitigating camera systematics. Telescopes for CCAT-prime and the Simons Observatory are scheduled to be completed in 2021, which will considerably advance the technological readiness of this optical design for large-aperture telescopes.

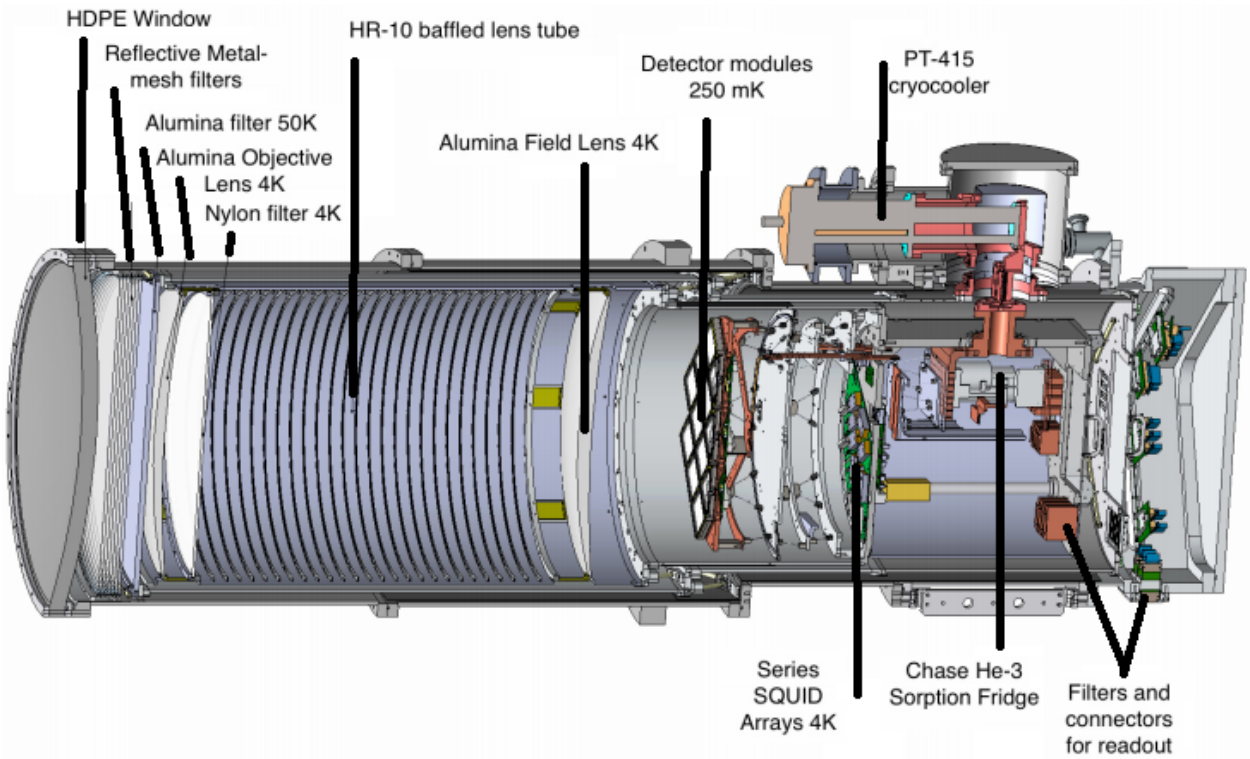

Figure 3: Design of the 0.5 m aperture BICEP3 refractor camera.

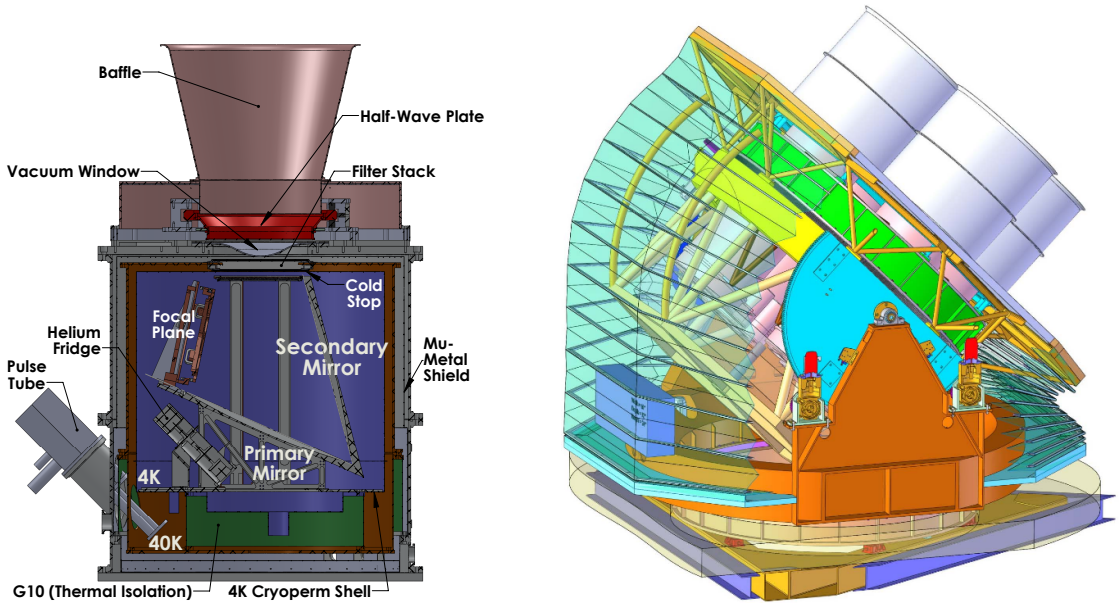

Figure 4: *Left:* Design of the 0.3 m aperture ABS reflector camera. *Right:* Preliminary design of four 0.5 m-aperture BICEP array cameras on a shared mount inside a ground screen.

We also consider similar telescope designs with apertures as large as 10 m when comparing different instrument configurations.

The most mature camera designs for the large-aperture telescopes are based on refractive optics. For example, the AdvACT, POLARBEAR-2, and SPT-3G cameras use three lenses (either silicon or alumina lenses) between the telescope focus and the detector arrays. These lenses serve

several important functions: 1) magnifying the telescope focus; 2) providing a cryogenic image of the primary, or stop; and 3) focusing the light onto detector arrays. Using several independent sets of lenses, or optics tubes, provides similar advantages to using several small-aperture cameras (improved optical performance and simplified deployment of different frequency channels). However, the limited field-of-view (FOV) on each large-aperture telescope motivates assembling these optics tubes as closely together as possible in a single camera. The practical constraints of assembling several optics tubes inside one cryogenic volume under vacuum near a telescope focus are being studied in detail and informed by the three-optics-tube AdvACT camera. A prototype camera design under development for a CD telescope is shown in Figure 5.

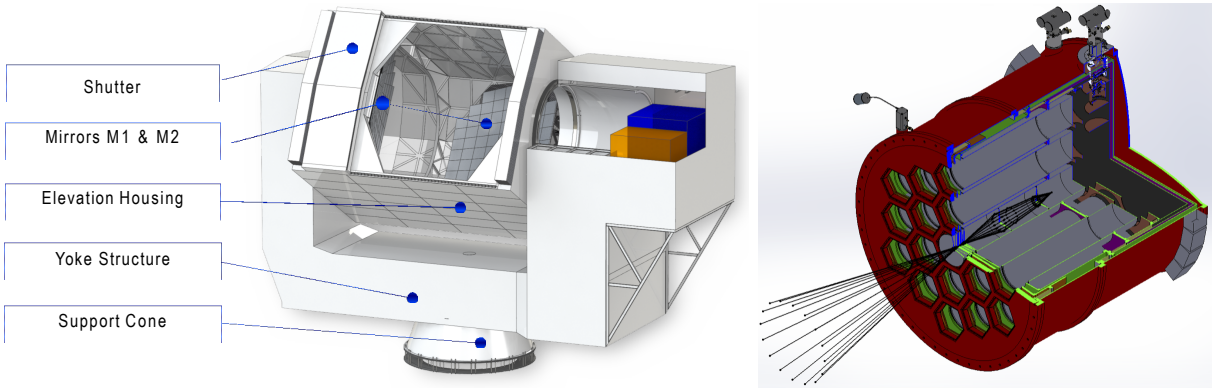

Figure 5: Candidate designs for the CMB-S4 large-aperture telescope and camera. *Left:* Preliminary design for the 6-m-aperture CD telescopes being built by CCAT-prime (www.ccatobservatory.org) and the Simons Observatory (simonsobservatory.org). *Right:* Preliminary design for a CMB-S4-scale camera developed by the Simons Observatory Collaboration, with 19 optics tubes (each with three or four 150 mm detector arrays) for the CD telescope shown on the left.

The detector technologies in all current ground-based CMB polarization projects are based on using superconducting antennas to couple the light from the telescope to transition-edge sensor (TES) bolometers. Three different techniques are being used to couple to the antennas (feedhorns, lenslets, and planar antenna arrays), and all three approaches are potentially viable for this concept design. To read out the TES arrays, three different techniques based on superconducting quantum interference devices (SQUIDs) are currently being used: time-division multiplexing (TDM), frequency-division multiplexing using MHz LC resonators (DfMux), and frequency-division multiplexing using GHz resonators (μMUX). Again, all of these are potentially viable detector readout options for CMB-S4. Examples of recently deployed multi-frequency arrays fabricated on 150 mm wafers are shown in Figure 6. For an overview of the detector array and readout techniques, see Abitbol et al. (2017).

When using TES bolometers, a dominant fundamental noise source is thermal fluctuation noise between the bolometer and the thermal bath. This noise contribution is most easily suppressed by reducing the bath temperature, which is why all CMB instruments with TES bolometers use cryogenic coolers that provide a bath temperature near 0.3 K (using Helium-3 sorption refrigerators) or 0.1 K (using dilution refrigerators or sometimes adiabatic demagnetization refrigerators). The lower bath temperature of a dilution refrigerator is preferable, but the refrigerator cost is greater, and it was not until recently that commercial dilution refrigerators became sufficiently reliable to deploy on remote telescopes. With a careful and comprehensive approach to instrument design and implementation, either cryogenic system is capable of suppressing bolometer noise to levels sub-dominant to photon noise from the atmosphere; however, 0.1 K dilution refrigerators provide greater noise and performance margin and are assumed for this concept design study.

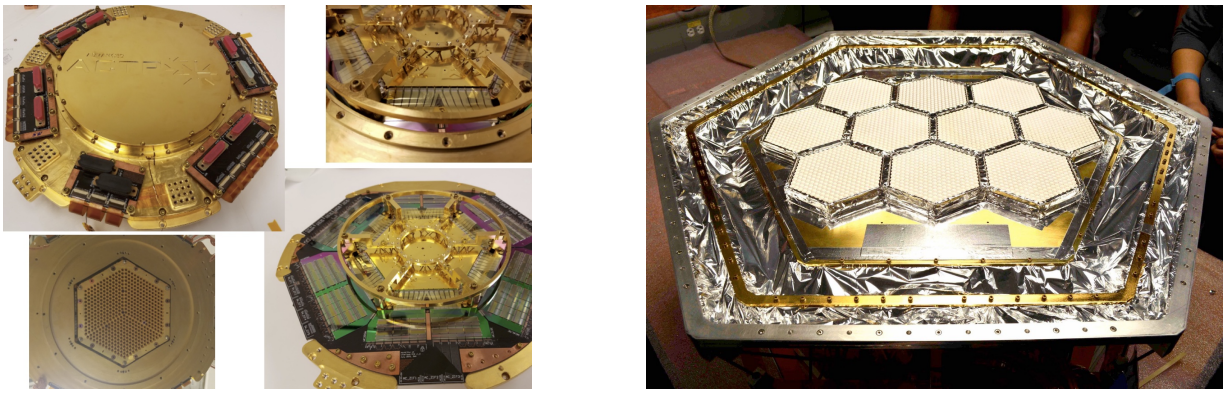

Figure 6: Examples of deployed multi-frequency detector arrays fabricated on 150 mm diameter wafers. *Left:* Different views of the first of four dual-frequency Advanced ACTPol detector arrays, which was deployed in 2016. This array is feedhorn coupled (feedhorns visible in lower left image) and uses TDM SQUID readout. *Right:* The three-frequency SPT-3G detector array, which was deployed in 2017. This array is lenslet coupled (white lenslets are visible on 10 detector wafers) and uses DfMux readout.

Polarization modulators are a compelling approach for mitigating low-frequency noise from the atmosphere as well as sources of instrument systematics that are on the detector side of the modulators. Polarization modulators have recently been deployed by the ABS, ACT, CLASS, and POLARBEAR collaborations. Results from measurements with Stage-3 experiments using modulators are needed to inform whether modulators will be used in CMB-S4, though the cost of modulators is small enough that this decision will not have a substantial effect on the strawperson concept presented here.

The observing strategy is strongly linked to the complementarity of the two sites. As described above, a major advantage of the South Pole is that the same sky regions are always accessible, which facilitates deep integration on small patches. A major advantage of Chile is that roughly 40% of clean sky is accessible, compared to roughly 10% from the South Pole (Figure 7).

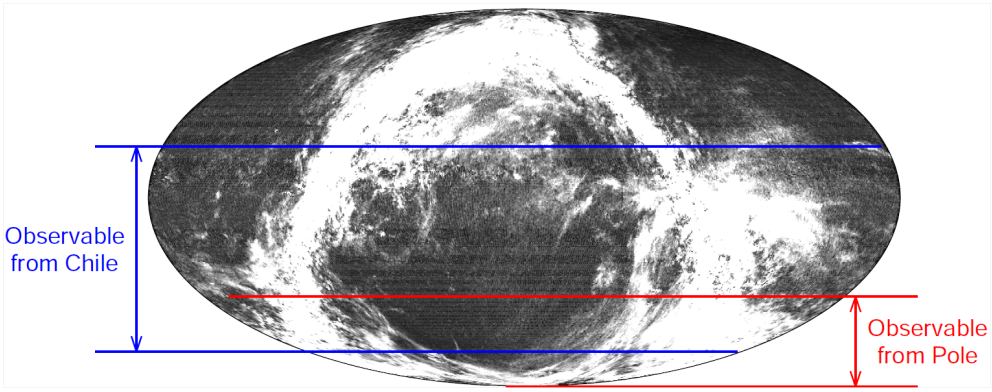

Figure 7: A Mollweide projection of the Planck 353 GHz polarized intensity map with the regions accessible by observing at elevation angles greater than 40° indicated. (The color scale is linear from 0 to 150 μK_{CMB}, and heavily saturated on the Galactic plane.)

4.3 CMB-S4 Strawperson Concept

A strawperson concept that meets the science and measurement requirements described in § 3 is shown in the Science Tracebility Matrix above. We envisage a deep survey for the inflation science over an effective 3%–8% of the sky made with both small-aperture cameras and large-aperture

telescopes, coupled with a shallower 40% sky area survey made only with large apertures for targeting $N_{\rm eff}$, neutrino mass, and large-scale structure science.

Table 2: Instrument configuration satisfying the measurement requirements.

		Frequency [GHz]									
Science	Item	20	30	40	85	95	145	155	220	270	Total
r	**14 x 0.5-m cameras**										
	# detectors	. . .	260	470	17 k	21 k	18 k	21 k	34 k	54 k	168 k
	Angular resolution [FWHM]		$77'$	$58'$	$27'$	$24'$	$16'$	$15'$	$11'$	$8\farcs5$	
	1 x 6-m telescope										
	# detectors	130	250	500	. . .	25 k	25 k	. . .	8.7 k	8.7 k	68 k
	Angular resolution [FWHM]	$11'$	$7\farcs0$	$5\farcs2$	. . .	$2\farcs2$	$1\farcs4$	. . .	$1\farcs0$	$0\farcs8$	
$N_{\rm eff}$											
	2 x 6-m telescopes										
	# detectors	290	640	1.1 k	. . .	50 k	50 k	. . .	17 k	17 k	136 k
	Angular resolution [FWHM]	$11'$	$7\farcs0$	$5\farcs2$	. . .	$2\farcs2$	$1\farcs4$	. . .	$1\farcs0$	$0\farcs8$	

Fourteen 0.5-m-aperture cameras will provide high sensitivity and stringent control of systematic errors on degree angular scales where we will search for a signal from primordial gravitational waves. To achieve additional foreground discrimination while maintaining high sensitivity, we propose to divide the 90 and 150 GHz atmospheric windows into two wide overlapping sub-bands. However, to make the most cost-effective use of focal plane area, we propose to use dichroic pixels (i.e., each pixel feeds detectors for two frequency bands and both linear polarizations—four detectors total). An approximate breakdown of the focal planes might thus be $1 \times 30/40$ GHz, $5 \times 85/145$ GHz, $5 \times 95/155$ GHz, and $3 \times 220/270$ GHz. At the same time, a single large-aperture "delensing" telescope (≥ 6 m) will scan the same patch of sky to provide the fine-angular-scale information required to reconstruct the lensing potential and to delens the large-angular-scale data. This telescope will be equipped with multiple optics tubes to cover the 30, 95/145, and 220/270 GHz bands. For the large-sky-area survey, two large-aperture telescopes will observe from Chile equipped with similar or identical cameras to the delensing telescope.

This strawperson has been derived using simulations detailed in Appendix A, which invoke current models of Galactic foregrounds and assumptions of instrumental performance that are directly scaled from sensitivities achieved so far by small-aperture instruments at South Pole and the combined experience of large-aperture telescopes operating at both sites. As more information becomes available over the next few years the situation will surely evolve. For instance, if evidence starts to emerge for a detection of r at higher levels, one would certainly want to expand the sky coverage of the deep survey to confirm on regions with a different realization of the foregrounds. Data from Stage-3 experiments and the knowledge of how CMB-S4 will be staged will inform the distribution of the telescopes among sites and the exact sky regions to target for the surveys.

5 OPERATIONS

The CMB-S4 project will operate under the auspices of a single community-wide collaboration, as described in §7. The basic operations model for CMB-S4 will be observations with multiple telescopes and cameras distributed across two sites, with observing priorities and specifications optimized for the CMB-S4 science goals, and data from all instruments shared throughout the entire CMB-S4 collaboration. Operational details for data management including science analyses

are described in § 6. Scientists working at DOE laboratories (funded by DOE) and at universities (funded by a mixture of NSF and DOE grants) will coordinate the observations, monitor the data, design and implement the data pipeline, and carry out science analyses. Instruments at both sites will collect data nearly continuously. These data will include calibrations as well as CMB observations. Nearly all observations will be automated, so that local operators on the sites will not be needed during routine observations.

The two CMB-S4 sites (§ 4.1), though remote, are both sufficiently well-established that fielding the CMB-S4 instrumentation does not represent a large risk, and costs and schedule estimates can be predicted from past experience. Although autonomous site management is required because there are many site-specific issues, the majority of the site costs scale with the number of cameras and telescopes. Examples include construction, shipping, power, and data storage and transfer. Therefore, operating two sites is not significantly more expensive than operating one.

In terms of operations, each site has advantages and disadvantages that will inform staffing needs, travel schedules, risk and safety planning, and pacing and phasing of installation and upkeep. For example, the Chile site is at very high altitude (5200 m, barometric pressure about half sea-level) and must deal with snow removal, but is accessible year round. Staff and visitors are housed in San Pedro de Atacama (2400 m, population ~ 2500) and drive ~ 1 hr up to the site as needed. The South Pole gets very cold (-90 C), is not accessible Feb–Nov, and suffers shipping restrictions on the size and weight of parts, but the small winterover staff is housed ~ 1 km from the telescopes and cameras, and is always available. At both sites, major maintenance is generally in summer when the weather is worst for observing.

6 DATA MANAGEMENT

6.1 Overview

The analysis of any CMB dataset is driven by the need to control statistical and systematic uncertainties to exquisite precision. Typically this involves an alternating sequence of steps, first mitigating the systematic effects native to the current data domain, and then compressing the cleaned data to a domain with higher signal-to-noise ratio.

Time domain systematics are dominated by experiment-specific instrument and observation effects, while the map and spectral domain systematics are dominated by experiment-independent foreground contamination, including lensing. As such, the core project deliverables are well-characterized maps at each observing frequency from which the community can then derive the full range of science results with appropriate uncertainties. This characterization can be explicit or implicit pixel covariance matrices, or well-matched Monte Carlo simulation sets. Note that the science results themselves are also an important source of feedback on the sufficiency of both the systematics mitigation, requiring close collaboration between these phases of the overall analysis, and potentially the instrument configuration and observing strategy. Given the unprecedented sensitivity of CMB-S4, there might also be the possibility of providing the community with nightly maps of a large fraction of the sky for identifying transient microwave sources. This could serve as a microwave complement to the LSST time domain astronomy program.

Beyond just the analysis pipelines, the overall data management program (Figure 8) must also include: getting the raw instrument data from the detector readouts to secure storage; generating the simulations required for experiment design, pipeline validation, and verification, and Monte Carlo uncertainty quantification and debiasing; making the legacy data and the associated software available to the community; and ensuring that sufficient and appropriate computational resources are available to execute all elements and stages of the project. Many of these issues are typically

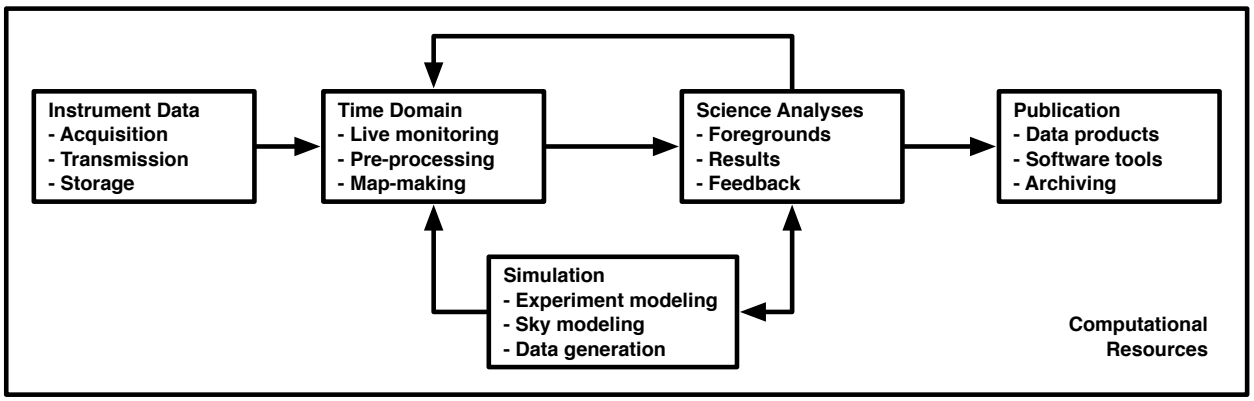

Figure 8: The broad elements of a CMB Data Management project and their interactions.

addressed by identifying one or more Data Centers (possibly tiered), which have the resources needed to meet the data management project requirements.

In what follows, all data rates and volumes correspond to the strawperson configuration and are taken to be their peak values. We further assume that the detectors are split equally between the Atacama and South Pole sites.

6.2 Computational Resources

As CMB experiments have targeted fainter and fainter signals, the quest for sensitivity has driven an exponential growth in the number of detectors fielded and the volume of data gathered (Figure 9). For suborbital experiments, this growth has paralleled Moore's Law, ensuring that High Performance Computing (HPC) and High Throughput Computing (HTC) have become integral elements of CMB data management. Particularly for projects with a long lifetime, it has also been essential to use the long-term development plans of national supercomputing facilities to take advantage of Moore's-Law growth in computing capability over the duration of the project. Over the last 20 years, the DOE's National Energy Research Scientific Computing (NERSC) Center has been the primary provider of HPC resources (cycles and storage) to the CMB community worldwide, annually allocating about 1% of its cycles to the CMB field. Continuing this support, we expect NERSC to serve as the primary CMB-S4 Data Center and to provide a significant fraction of the required computing resources, in conjunction with other HPC and HTC computing facilities including DOE's Argonne Leadership Class Facility (ALCF) and NSF's eXtreme Science and Engineering Development Environment (XSEDE), all connected over the Enegy Sciences Network (ESnet) high-bandwidth networks. Coupled with the instruments and international networks, the complete data infrastructure should then be conceived as a single, distributed Science Data Facility.

Key issues will be: securing guaranteed long-term access to the necessary HPC and HTC resources, including network bandwidth, spinning disk and archival storage, and compute cycles; managing internal data distribution, including identifying the need for and scope of any secondary data centers; achieving and maintaining the computational efficiency required to exploit these resources in the epoch of increasingly complex, increasingly energy-constrained, computing architectures; and developing and enforcing software standards across the collaboration.

6.3 Instrument Data

Acquisition—Each site will generate data at 3 Gbps, which can currently be reduced to 1 Gbps with lossless compression. Key issues will be provisioning fail-safe storage at the observing sites, improving compression algorithms, and developing optimal raw data file formats.

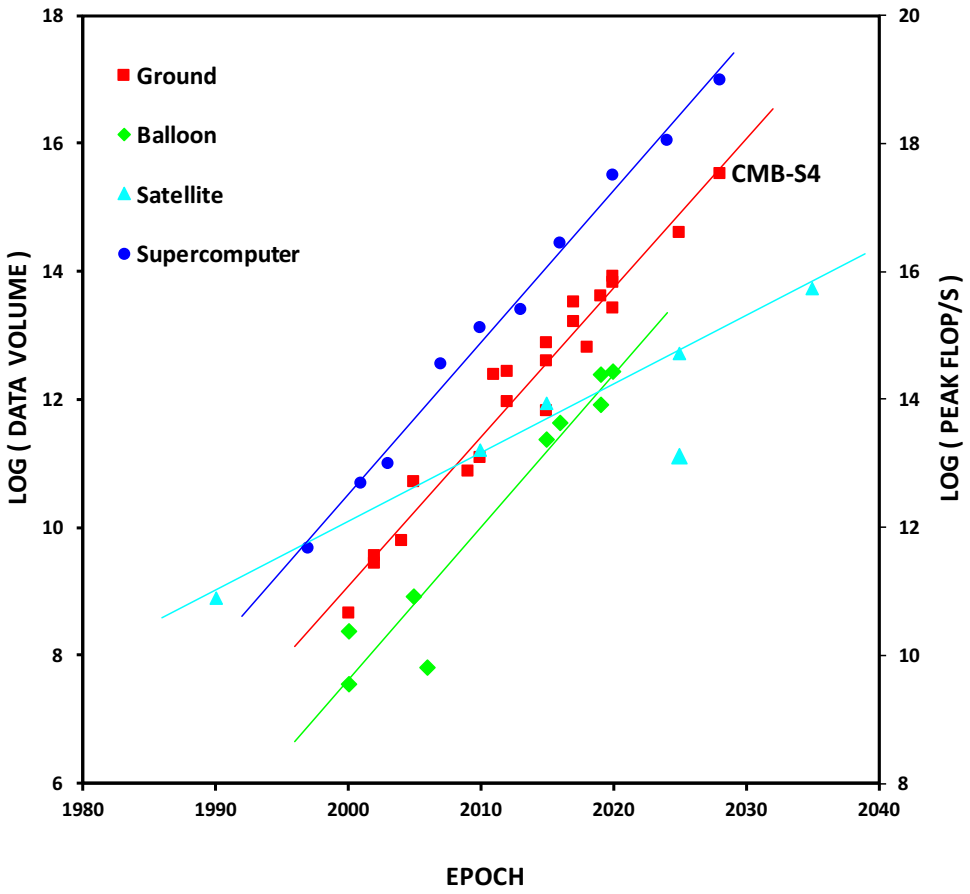

Figure 9: The growth in the volumes of ground-based, balloon-borne and satellite CMB datasets, and in the performance of HPC systems (using the peak performance of the flagship NERSC system at any epoch as a proxy), over a 40 year span, with CMB-S4 identified.

Transmission—From the Atacama, it will be possible to provision sufficient bandwidth over the REUNA/AMpath/ESnet networks to transmit the full data back to the US. Key issues will be installing the dedicated fiber needed to connect the site to REUNA (likely via ALMA/AURA, building on the planned Simons Observatory fiber), and obtaining sufficient bandwidth on the international networks (taking advantage of the significant upgrades due to LSST). At the South Pole, lossy compression and/or decimation will further reduce the data to a subset that can be transmitted over the TDRSS network back to the US, while the full data will be stored on disk and shipped. Key issues will be maximizing the TDRSS bandwidth available, and optimizing the information content of the compressed/decimated data that this then carries (including by onsite reduction to maps). Note that the two sites will support very different data cadences, both of which will need to be respected by downstream processing.

Storage—CMB-S4 will need 2 PB of disk to keep a single season's good data spinning,[2] and 50 PB of archival storage in at least two separate locations to preserve the entire dataset. Key issues will be provisioning the necessary disk and tape space and efficiently managing the timely transfer of data back and forth between archive and disk.

[2]Here we assume 20% observing efficiency, broadly consistent with experience to date. Note that the forecasts and simulations detailed in Appendix A rely on direct scaling from achieved noise levels, rather than assumptions about observing efficiency.

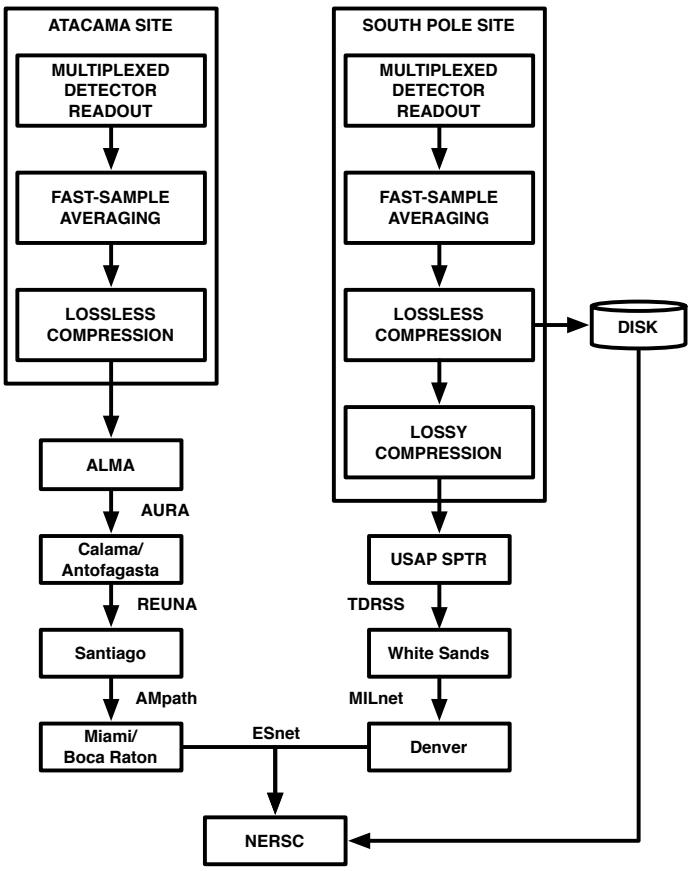

Figure 10: Raw data acquisition and transmission from the two observing sites to the primary Data Center.

6.4 Time Domain

Live monitoring—Near-real-time monitoring of the raw data is necessary for site and instrument quality assurance. Live monitoring tools could also be adapted to enable time domain microwave astronomy, since large fractions of the sky will be observed nightly. Key issues will be automating this monitoring to manage the data volume, providing timely alerts to allow issues to be addressed with minimal loss of observing time, and optimally combining information from multiple telescopes.

Pre-processing—The raw time-ordered data are first made to match the data model assumed by the map-making algorithm they are to be fed to, including quality control and flagging, absolute and relative calibration, and mitigation of time-domain systematics. This will be the ultimate limiting factor on the CMB-S4 science, and an ongoing area of research and development throughout the lifetime of the project. Key issues will be developing a computational- and human-efficient framework for data exploration, mitigating systematic effects to unprecedented precision, and addressing the challenges of the inevitable unknown unknowns.

Map making—Some or all of the pre-processed data at a given frequency are reduced to a well-characterized map, using either the binning, destriping, or maximum likelihood algorithms. Key issues will be computational tractability and efficiency, assessing the viability of different types of map for different science goals, and sufficiently well-characterizing the correlations of the signal, noise, and residual systematics in any map.

6.5 Science Analyses

Foregrounds—Science results are derived from the maps and their 2-point and higher correlation functions, and their quality depends on the removal of foreground residuals in each of these data representations, through component separation and masking in the map domain and cleaning (e.g., faint point source subtraction, delensing) in the spectral domain. Key issues will be validation and verification of the algorithms and implementations in the absence of exact knowledge of the foregrounds, and the propagation of uncertainties from the frequency maps through to the cleaned data representations.

Results—While we have focused on the primary science goals in order to define the project requirements, the CMB-S4 dataset will support a wealth of science investigations. Wherever possible, these scientific results will be derived from multiple analyses to ensure robustness to algorithm choice. Key issues will be algorithm validation, verification, and comparison, algorithm optimization given the added complexity of a variety of instruments and platforms and a hybrid survey strategy, and ensuring an appropriate level of control wherever analyst-bias might be a concern.

Feedback—The various science analyses can also provide vital feedback. Residual instrument- and observation-specific systematics identified in a reduced domain (such as the signature of beam sidelobes in maps) are fed back to the time domain processing for further mitigation, while the science results themselves (including component-separated foreground maps and cosmological parameter values) are fed back to the input sky model to provide more representative simulations. Key issues will be coordination and consistency across the wide range of disparate science investigations, and minimizing the introduction of residual noise and systematics into the sky model.

6.6 Simulations

Experiment modeling—An experiment comprises the instruments (optics and electronics) and the observations (scanning strategy and environment). An experiment model must characterize each of these sufficiently to generate a synthetic dataset. This characterization is derived from a parametric description of each of the experiment elements (including any time-variation), and estimates of the parameter values and distributions from laboratory and site measurements, including from the real data during operations. Key issues will be developing a parameterization that captures the experiment to the required level of detail, and deriving the values of these parameters from limited data.

Sky modeling—A sky model includes the scalar, tensor, and non-Gaussian CMB signals, the extragalactic foregrounds (and the lensing of the CMB that they induce), and the Galactic foregrounds. Key issues will be the development of representative Galactic foreground models that capture the complexity of the real sky, and of extragalactic foreground models that are both representative and consistent with those being used by large-scale-structure survey projects such as DES, DESI and LSST to enable cross-correlation studies to inform the CMB-S4 detailed design phase.

Data generation—A simulation applies an experiment model to a sky model to generate a synthetic data realization. This can be in the map or time domain, depending on the representation of the experiment. Map-domain simulations are less computationally expensive, but only time-domain simulations can capture both the full complexity of the experiment and the impact of the algorithm chosen to reduce the data to maps. Key issues will be computational tractability and efficiency, especially when Monte Carlo methods are used to characterize the maps.

6.7 Publication

Data products—The CMB-S4 project will need to deliver both its own legacy science results, and all of the data products necessary for the community to continue the scientific exploration of this extraordinary data set. Key issues here will be determining the schedule(s) for releasing project deliverables and their derived data products and science results, providing a tractable characterization of correlation structure of any data product, and deciding what (if any) time-domain data to release and how.

Software tools—Along with the derived data products, the collaboration will need to release the software used to derive them in order to enable the community both to validate and to extend every step in the analysis pipeline. Key issues here will be documenting all of the detailed experiment-specific knowledge, particularly with respect to the time-ordered data processing, and distributing legacy code that will continue to be useable in an evolving computing environment.

Archiving—Legacy maps, power spectra, and likelihood functions will be distributed through NASA's Legacy Archive for Microwave Background Data Analysis (LAMBDA), which has become the *de facto* standard for both satellite and suborbital data. However, much larger data objects, such as the time-ordered data or massive suites of Monte Carlo simulations, are too large to serve from LAMBDA, and furthermore are of no value unless they are alongside sufficient computational resources to manipulate them. Key issues will be the provisioning of the spinning disk and of the cycles necessary to serve the largest legacy data products in a way that is useful to the community.

7 COLLABORATION STRUCTURE

A single community-wide collaboration is needed to realize the CMB-S4 project and to extract the science. Traditionally, the experimental efforts of the US CMB community have been distributed over several highly competitive, small, university-led projects. To meet the increasingly challenging sensitivity goals required to advance the science, the complexity and scale of the experiments have increased, which in turn has required larger teams and significantly more resources. The result has been a natural trend to reduce the number of experiments, with teams merging and technologies and sites being shared. Currently the ground-based experimental effort in the US is distributed in five collaborations with experiments deployed at the South Pole (SPT and the BICEP/Keck program) and the Atacama Plateau in Chile (ACT, POLARBEAR, CLASS). Consolidation is continuing, with the POLARBEAR and ACT teams merging into the Simons Observatory, and with the SPT and BICEP/Keck teams coordinating their measurements for joint analysis of the same patch of sky. The substantial scaling—in detectors and readout, data management, precision of phenomenology and simulations, and data analysis—to achieve the science goals targeted for CMB-S4 requires further consolidation as well as the expertise and resources of the National Laboratories. The management and logistical challenges are far beyond the capabilities of the traditional university CMB research groups.

The US CMB community came together in the 2013 Snowmass Physics planning exercise to consider the future of the CMB experimental program. Two influential papers were produced, *Inflation Physics from the Cosmic Microwave Background and Large Scale Structure* (Abazajian et al. 2015a) and *Neutrino Physics from the Cosmic Microwave Background and Large Scale Structure* (Abazajian et al. 2015b), which in addition to articulating an exciting science case with clear goals, presented an experimental concept, CMB-S4, to achieve the required measurements. This science case and CMB-S4 concept were input to the deliberations of the Particle Physics Project Priority Panel (P5), which in its 2014 report recommended CMB science be considered part of particle physics, and that the CMB-S4 project should be pursued under all budget scenarios considered.

Figure 11: Group photo taken at the September 2016 CMB-S4 workshop. Roughly 150 scientists participate in the workshops, which are held roughly every six months. There have been six workshops to date.

Building on the Snowmass efforts, the informal CMB-S4 collaboration, including participation from all US ground-based teams and many international groups, was created and has been working effectively since to advance the science case, technology, and design of CMB-S4. The informal collaboration has held workshops on a six month cadence (six to date, each with approximately 150 scientists, see Figure 11), organized science and technical working groups, and produced the CMB-S4 Science Book and CMB-S4 Technology Book. These efforts have led to the creation of the joint agency CDT.

Going forward with multiagency and other support requires a formal CMB-S4 collaboration. The collaboration is currently working toward this goal, having recently appointed a 15-member Interim Collaboration Coordination Committee (ICCC) charged with setting up working groups to propose a governance model for CMB-S4 for review by the next workshop in March 2018.

The CDT endorses this plan, and encourages that a premium on career development of young scientists be explicitly included in the structure of the formal collaboration. To maintain the necessary expertise, the collaboration must provide attractive and rewarding career paths through the long process of designing, prototyping, deploying, commissioning, and operating CMB-S4.

The collaboration structure will also need to anticipate working with the yet-to-be-established CMB-S4 project, and eventually CMB-S4 operations management. While the community itself, coordinated by the ICCC, will set up the formal CMB-S4 collaboration, we provide an example structure in Figure 12 that incorporates aspects of existing collaborations and illustrates how the CMB-S4 Project and Collaboration might interact. Note that prior to the start of the project, the CMB-S4 collaboration Technical Working Group Board will guide the continued technical development of CMB-S4. During the construction Project phase, the function of these groups, and many of the collaboration members, would be incorporated into the Project Technical Board and the work breakdown structure (WBS) Level 2 management (yellow boxes).

8 COMPLEMENTARITY WITH MEASUREMENTS FROM SPACE AND BALLOONS

The history of CMB measurement is punctuated by breakthroughs from experiments on the ground, high-altitude balloons, and satellites. Penzias and Wilson discovered the CMB from the ground, anisotropy was discovered with the COBE satellite, the first acoustic peaks were discovered from balloons complemented by ground measurements, and polarization anisotropy was discovered from the ground.

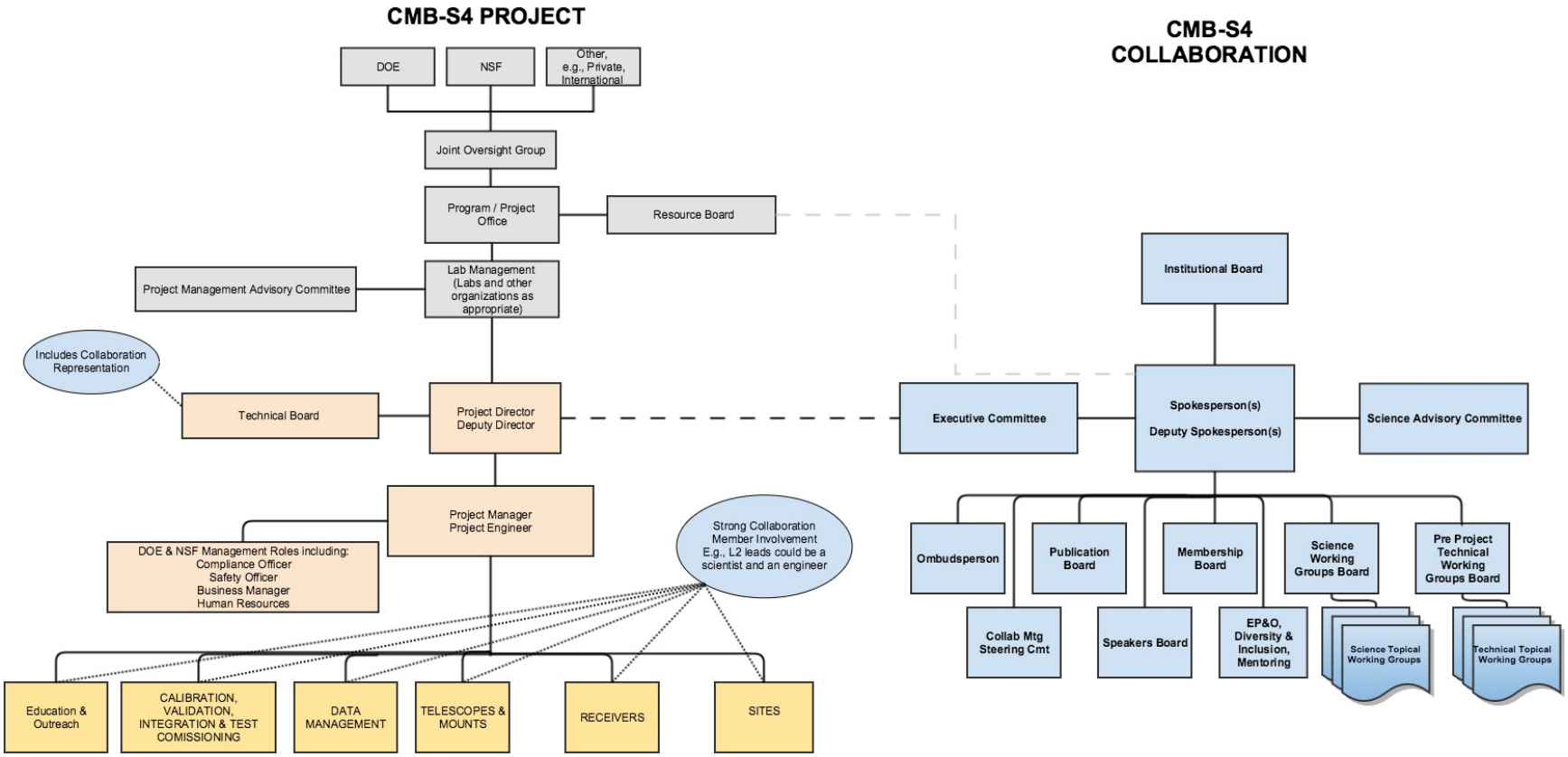

Figure 12: Example of a CMB-S4 collaboration structure (right side) that incorporates aspects of existing collaborations and shows the possible interplay with the CMB-S4 project (left side). This figure is for illustration only—the detailed structure will be determined by the CMB-S4 collaboration. Prior to the start of the project, the CMB-S4 collaboration Technical Working Group Board will guide the continued technical development of CMB-S4. During the construction project phase, the function of these groups, and many of the collaboration members, would be incorporated into the project Technical Board and the WBS Level 2 management (yellow boxes).

Measurements from ground and space have been complementary in the angular scales and the range of frequencies measured. Ground-based experiments have focused on degree to arcminute angular scales, whereas space-based measurements have focused on scales as large as the dipole to five-arcminute angular scales. The 10-meter class telescopes required to resolve arcminute scales are more practical in cost on the ground, whereas the clear view from space with no atmospheric interference and access to the full sky is optimal for measuring the largest angular scales. Recently, WMAP and Planck in space and ACT and SPT on the ground together have made the state-of-the-art temperature anisotropy measurements from dipole to arcminute angular scales.

There is also a complementarity in the frequency coverage of ground-based and space-based observations. Ground-based measurements are sensitive in atmospheric windows up to $\sim$300 GHz, which simulations suggest is adequate to reach the science goals of CMB-S4. Space-based measurements can give additional information at frequency bands across the entire spectrum that are not accessible from the ground, and at frequencies $>$300 GHz, which are important for understanding the properties of Galactic dust.

The CMB-S4 deep search for inflationary physics and rich science goals for $N_{\rm eff}$, neutrino masses, and dark energy are self-contained in that they do not require auxiliary data from a future CMB satellite. But CMB-S4's science would be enhanced if satellite data are available. JAXA is supporting a phase-A study for the LiteBIRD mission, and NASA has recently commissioned a study for a 'Probe' class mission. The combination of data from CMB-S4 and a satellite would increase confidence in the inflation measurements through the complementarity of CMB-S4's measurements at degree scales and the satellite's measurements at degree and larger angular scales. High-resolution CMB-S4 observations can be used to delens coarser resolution space data, improving the depth of the space-based inflation search. Also highly complementary, the space mission's measurement of the optical depth τ could greatly increase the precision of the measurement of the sum of neutrino masses from CMB-S4 lensing analysis.

Balloon-borne measurements have been critical in the past and will continue to play an important role in the future. They have the potential to provide complementary data on foregrounds well before a space experiment is flown. They are less prone to atmospheric fluctuations, opening the prospect for measurements on large angular scales, and possibly improving the constraint on τ.

9 TECHNOLOGY DEVELOPMENT

The CMB-S4 strawperson instrument can be developed based on existing technology that is being deployed in Stage-3 experiments. CMB experiments that are either operating or in construction have successfully demonstrated most of the characteristics required for CMB-S4, but at a much smaller array scale. Investment in technology development will reduce the cost, schedule, and performance risk associated with the project, and investigation of alternatives has the potential to reduce the baseline cost.

The Stage-3 technology currently being deployed is described in § 4.2. Appendix B describes details of the technology development status. We further refer the reader to the CMB-S4 Technology Book (Abitbol et al. 2017) for a detailed description of the current state of technological readiness of four main technology areas: telescopes; camera optics; focal-plane optical coupling; and focal-plane sensors and readouts. We leave site logistics to § 5. In this section, we provide an overview of the recommended investment in the pre-project period.

The plan for pre-project investment was derived from a coordinated plan developed by the four DOE labs (ANL, FNAL, LBL, and SLAC), and does not include important work being done in universities for existing experiments. This plan was prioritized based on the development of a risk assessment for the CMB-S4 strawperson instrument configuration.

Table 3: Assessment of most significant project risks across subsystems. Risks are ordered from highest (top) to lowest (bottom). Pre-project investment is prioritized by risk and schedule.

Subsystem	Risk	Mitigation
Readout	Integrated performance (MUX factor, noise)	Develop multiple readout technologies with orthogonal technical risks, and downselect.
Detectors	Array production	Develop and validate processes, yield, and throughput at multiple fabs.
Data management	Simulations, especially those based on time-ordered data.	Develop simulation framework to evaluate instrument designs and systematics.
Optics	Half-wave plates and anti-reflection coatings	Demonstrate ETU of half-save plate and anti-reflection coating.
Cryostats	Complexity, cryogenics	Build Stage-3 prototypes; execute early designs and procurement.
Telescopes	Ground pickup	Early design and analysis of ground shields.

As part of a risk-assessment exercise, the top project risks were identified across subsystems. These are shown in order of decreasing risk in Table 3, along with research items to mitigate the risks. The prioritized plan for pre-project investment was developed to mitigate project risks at the appropriate times, given the severity of the risk and the anticipated schedule of the straw-person instrument configuration. The associated effort over a two-year period is $4, 4, 4, 2$ full-time equivalents (FTEs) per year for scientists, engineers, postdocs, and technicians, respectively, for readout, $2, 2, 2, 2$ for detectors, $1, 1, 3, 1$ for testing, and $1, 1, 2, 0$ for the development of a simulation framework. The cost of these investments is described in § 11. Further design studies and systems engineering work will also be needed to optimize technological choices in terms of performance, schedule, and cost.

In summary, CMB-S4 will require a large increase in the number of detectors and associated readout components and camera optical elements as compared to Stage-3 experiments, so improving reliability and throughput is important. Technology development should be targeted to mitigate cost, schedule, and performance risk, as well as to exploit opportunities to reduce cost. The program of investment in the pre-project period described in this section is budgeted in § 11, with details provided in Appendix B.

10 STAGING AND SCHEDULE

Although the science goals of CMB-S4 are signficantly broader in scope than any of the existing CMB experiment teams have attempted to meet, many of the technical challenges for CMB-S4 are ones of scale: fabricating and testing hundreds of thousands of detectors; building large cryostats; fielding multiple cameras and telescopes at two remote sites; and collecting and analyzing massive amounts of data in multiple survey regions with multiple frequency bands. Happily, the formation of the CMB-S4 collaboration from a half-dozen or more experiment teams with many years of field experience, largely from universities, with a growing influx of DOE lab researchers with extensive experience with large projects, means those challenges can be addressed with parallel efforts. Thus, the project schedule for CMB-S4 can be short, seven years from start of project (CD0) to the start of operations (CD4).

The CMB-S4 project could start immediately. A project start (CD0) in 2019 with construction (CD3) three years later has been suggested. The CDT supports and recommends this timeline. Project start in 2019 is also aligned with the ongoing Stage-3 experimental efforts. They have recently completed or will soon complete their construction phases, so resources regarding personnel

and university groups with crucial expertise are already becoming available for CMB-S4. Instruments such as Simons Observatory and BICEP Array that, in many ways, serve as prototypes for CMB-S4 instrument, will come on line and help inform the CMB-S4 design.

Funding considerations and the specific prioritization requirements for the agencies impact the timeline, however, with DOE and NSF having different constraints. It is therefore important to develop a staged implementation plan that leads to the full CMB-S4 program.

As the CMB-S4 science goals lead to the requirements for two sites, two sizes of telescopes, and two surveys, there are several staging options for its implementation, and even more when the use of legacy instrumentation located at the two sites is included. Once the formal CMB-S4 collaboration is established, it could develop memoranda of understanding to allow optimal use of existing or privately funded telescopes and other instrumentation. In this scenario, an example of an attractive early implementation plan would be for DOE to proceed with detector development and fully outfit the focal plane(s) of one or more large-throughput 6-m telescopes in Chile, as well as adding several small-aperture cameras at the South Pole. NSF could continue to provide modest operations support. This would jump-start the full CMB-S4 program and mitigate the highest technology risks. Scientifically it would enable the start of the CMB-S4 science goals on primordial gravitational waves, with the large telescope(s) in Chile used to delens the large-angular-scale data from the deep survey targeted with the small-aperture cameras. For the portion of the day when the deep field is not observable from Chile, the large telescope could be used to start the large $N_{\rm eff}$ survey. This staging would allow substantial early progress toward CMB-S4 science goals.

11 COST

11.1 Approach

The overall approach in the cost estimate is to capture all aspects of the experiment, with fidelity appropriate for a concept design, taking care not to miss any scope that might drive the cost. The estimate includes cost vs. experiment configuration, which supports preliminary design vs. cost trades, identifies the major cost drivers, and informs the planning of pre-construction technology development. A preliminary estimate of the operations cost is also included to support agency planning for the post-construction phase of the experiment. The strawperson concept for CMB-S4 captures the overall scope of the experiment needed to meet the science requirements, but not the technology details. The cost estimate is based on parametric scaling from existing and prior projects, with contingency based on design maturity. A panel of DOE experts reviewed the costing approach on 12 September 2017 (Appendix C), and recommendations from that review have been incorporated in the cost estimate.

The cost estimate has three parts: pre-project technology development; the construction project; and post-project operations. Of these, the construction project costs are the most detailed. Cost estimates for technical tasks are based on the total effort to complete the work (FTE-years); management activities are level of effort (FTE/year), assuming a 7-year construction project (Figure 13). Construction costs are based on a WBS (Figure 14), with a simple dictionary that describes the work at WBS level 1 (e.g., telescopes, cryostats, detectors, and readout) and level 2 (e.g., large telescopes). The work includes design, pre-production prototypes, management, reviews, fabrication, integration, and commissioning. Costs are estimated at WBS level 3 or lower, with a contingency for each item based on maturity of design. The contingency captures risks that flow from the design uncertainty in a simple way that is appropriate for preliminary cost estimates (Table 4). Contingency is assigned uniformly across the WBS, with 40% for design, most fabrication tasks, assembly, and travel, 50% for optics-tube fabrication, wafer-test setups, wafer testing, software, integration,

and commissioning, 10% for management tasks, and 10% for overhead on procurements.

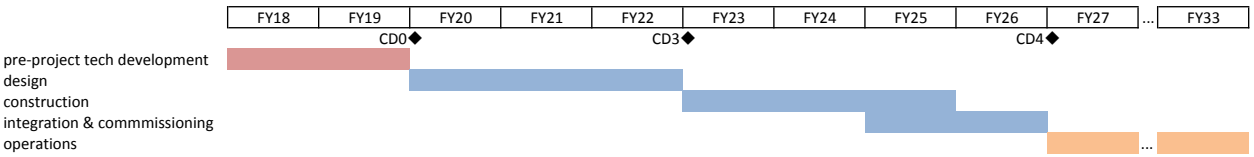

Figure 13: CMB-S4 strawperson schedule.

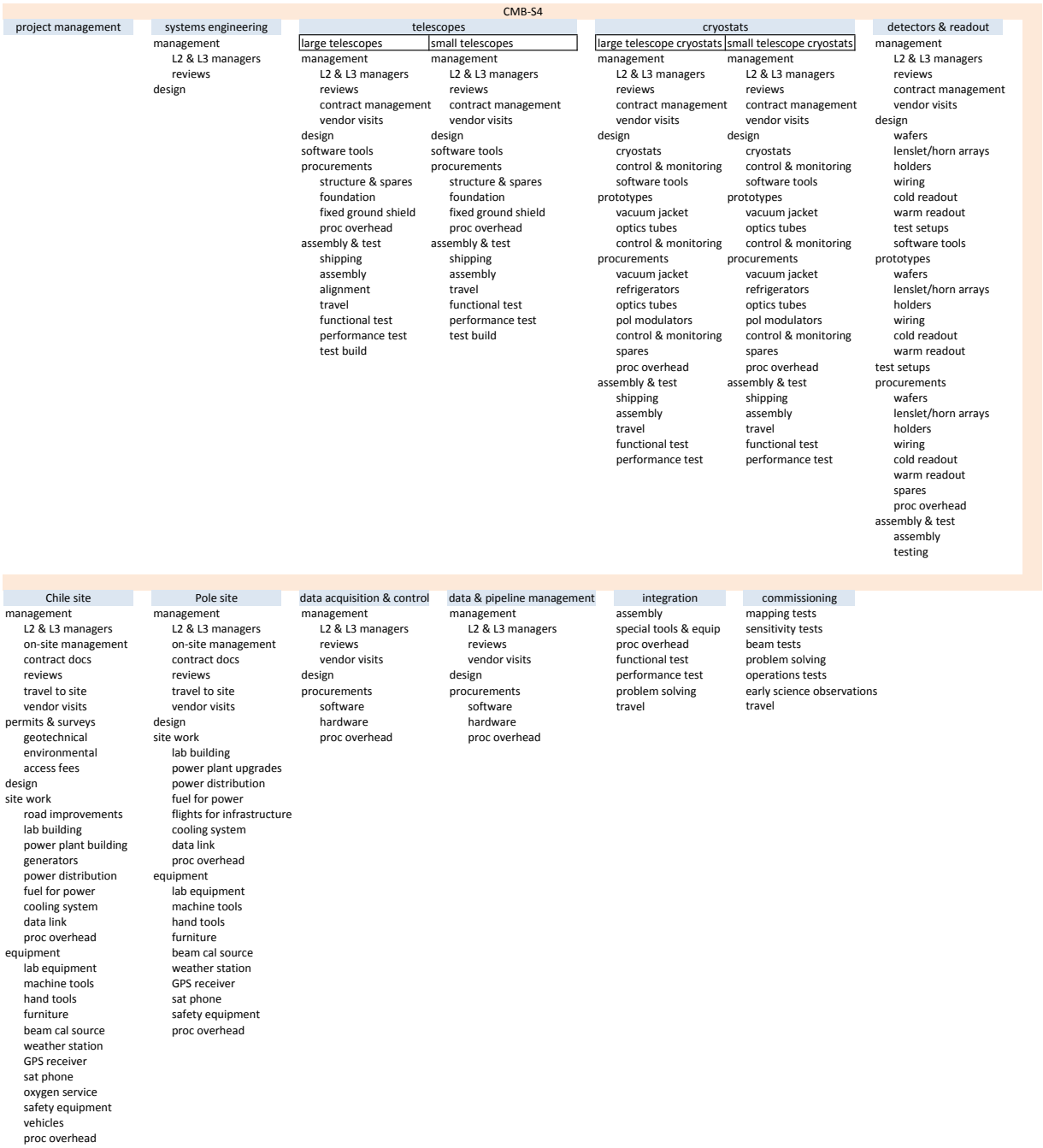

Figure 14: CMB-S4 strawperson WBS.

Table 4: Contingency allocation.

Type of estimate	Contingency [%]
Advanced	10–20
Preliminary	20–40
Conceptual	40–60
Pre-conceptual	60–80
Rough	80–100
Beyond state of the art	> 100

11.2 **Basis of Estimate**

Each entry in the cost estimate is based on costs for deployed instruments, estimated costs for instruments that are being designed now, or bottom-up estimates by experts. The origin of each entry is recorded to provide traceability. Stage-2 CMB experiments provided costs for telescopes, circa 2006, but Stage-2 cameras were too small ($\sim$ 1 k detectors), and the technology too old, to support reasonable scaling. Stage-3 experiments provided costs for cameras with appropriate technology, but an order of magnitude smaller detector count than for CMB-S4. No new, large telescopes were built in the Stage-3 era, because the Stage-3 projects all use existing Stage-2 telescopes. Experiments currently being designed, e.g., Simons Observatory and BICEP Array, provided estimates and quotes for telescopes and cameras that are close to CMB-S4 in scope and technology.

Costs are scaled to CMB-S4 using simple parametric models, e.g., the cost of a wafer assembly scales with number of wafers, and the cost of a telescope scales roughly with mass, which in turn scales with (diameter)$^{8/3}$. The result is a bottom-up estimate based primarily on many top-down subsystem costs from previous and current projects. Cost models of this type are useful for small perturbations relative to recent, similar projects.

The CMB-S4 cost estimate has multiple cost models for many items, reflecting different implementations in previous projects. In most cases, the CMB-S4 cost estimate takes the mean of the cost models, after removing obvious outliers, but some specific models have been adopted, e.g., the cost of the large telescopes is based on a recent quote for the Simons Observatory, and the cost of detector wafers is based on a detailed analysis of fabrication steps for Advanced ACTpol wafers.

The pre-project technology development plan is described in §9. Most of the development work is on readout electronics and detectors, with a smaller effort on simulations to support experiment design trade-offs. The associated effort is 16, 16, 22, and 10 FTE-years for scientists, engineers, postdocs, and technicians, respectively, at a cost of \$9 M, excluding the 16 FTE-years of scientist effort because that is supported by DOE research funds. The cost does not include important work being done in universities for existing experiments, technology development supported by e.g., NASA and NIST, or R&D in the Simons Observatory.

The operations cost is based on a preliminary bottom-up estimate that includes management, site staff, utilities, instrument maintenance, data transmission, data products, pipeline upgrades, collaboration management, and science analysis. The effort is roughly 30 FTE/year for data products and pipeline upgrades, 15 FTE/year for science analysis, and 25 FTE/year for management and site support. The annual operations cost is \$32 M in 2017 dollars, excluding 20 FTE/year of scientist effort supported by DOE research funds. If the DOE-supported scientist effort is included, annual operations are 10% of the construction cost, which is typical for an observatory.

11.3 Cost for Strawperson Concept

The strawperson concept for CMB-S4 has three 6-m diameter large-aperture telescopes, each equipped with a single 2.8-m-diameter cryostat containing nineteen optics tubes, and fourteen 0.5-m-diameter small-aperture cameras in two 2-m-diameter cryostats, each on its own small mount. The instruments are deployed at two sites, Chile and the South Pole. In order to simplify the cost model, the detector counts are converted to the equivalent at 150 GHz, based on focal plane area. This conversion gives a total of 512 k 150-GHz detectors, 96 k on each large-aperture telescope and 224 k on the fourteen small-aperture cameras.

The construction project cost is $412 M in 2017 dollars, including 45% contingency. Figures 15 and 16 show the distribution of cost by WBS item, cost by type of work, and effort by type of personnel. The cost does not include 170 FTE-years of scientist effort that is traditionally supported by DOE research funds at labs and universities. One quarter of the work is assumed to be done in NSF-funded universities, and the cost of this is included. The construction project effort for scientists, engineers, and technicians is 21, 32, and 10% of the cost; 30% engineering effort is typical for projects of the complexity of CMB-S4. The top-level management effort, which includes the project manager, project controls, agency reviews, safety, quality assurance, and outreach, is estimated at 10% of the construction cost, with 10% contingency. Top-level management does not include systems engineering, which is estimated in a separate WBS item. Costs for the strawperson concept are summarized in Table 5.

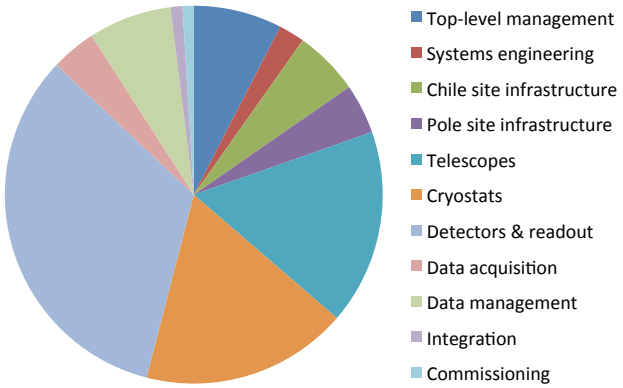

Figure 15: Fraction of the total cost for each WBS item.

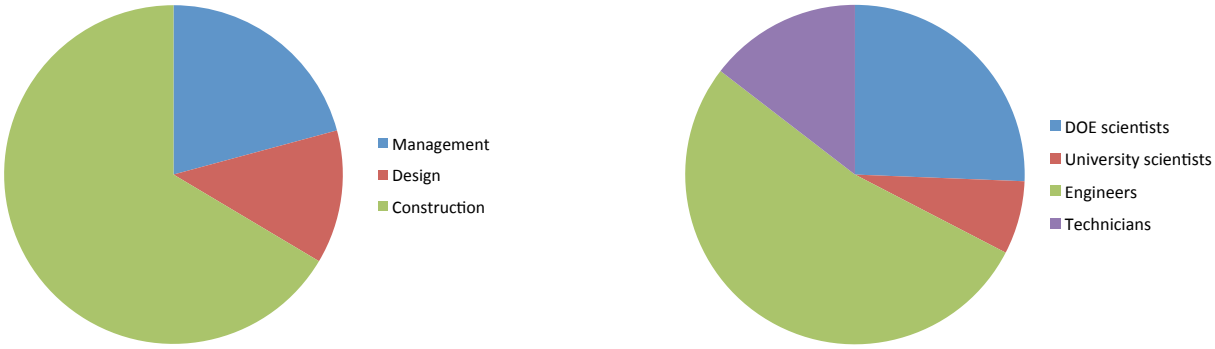

Figure 16: *Left*—Distribution of cost by type of work. Management includes levels 1–3. *Right*—Distribution of effort by type of personnel. DOE scientists are not included in the total project cost; university scientists are.

Table 5: Strawperson cost summary.

Estimate	Cost including contingency [$M in 2017]	Contingency [%]
Pre-project development	9	
Construction project	412	45
Annual operations		
Sites	18	100
Science analysis	14	150

11.4 **Cost Range**

Table 6 shows some options for the use of existing telescopes (Simons Observatory, CCAT-prime, and BICEP Array) that will be deployed prior to CMB-S4. The table also includes cost estimates for alternative configurations with larger telescopes that would increase the legacy value of the high-resolution survey. The higher resolution provided by a 10-m diameter telescope, for example, would roughly double the number of high-redshift galaxy clusters, for an additional cost of about $50 M. The cost range for the various options is $354–470 M.

Table 6: Options and Alternatives.

Option	Cost [$ M]
Strawperson ≡ 3×6 m, 14×0.5 m, 512 k 150-GHz detectors	412
2×10 m; 14×0.5 m; 480 k 150-GHz detectors	462
2×8 m; 1×6 m; 14×0.5 m; 512 k 150-GHz detectors	462
2×6 m; 1×10 m; 14×0.5 m; 512 k 150-GHz detectors	470
Strawperson but use one existing 6 m	393
Strawperson but use 2 existing 6 m	375
Strawperson but use 4 existing 0.5 m	391
Strawperson but use 2 existing 6 m and 4 existing 0.5 m	354

REFERENCES

Abazajian, K. N. et al., Inflation Physics from the Cosmic Microwave Background and Large Scale Structure. 2015a, Astropart. Phys., 63, 55, arXiv:1309.5381

Abazajian, K. N. et al., Neutrino Physics from the Cosmic Microwave Background and Large Scale Structure. 2015b, Astropart. Phys., 63, 66, arXiv:1309.5383

Abazajian, K. N. et al., CMB-S4 Science Book, First Edition. 2016, arXiv:1610.02743

Abitbol, M. H., Ahmed, Z., Barron, D., et al., CMB-S4 Technology Book, First Edition. 2017, ArXiv e-prints, arXiv:1706.02464

Barron, D., Ade, P., Anthony, A., et al., The POLARBEAR Cosmic Microwave Background Polarization Experiment. 2014, Journal of Low Temperature Physics, 176, 726

BICEP2 and Keck Array Collaborations, Improved Constraints on Cosmology and Foregrounds from BICEP2 and Keck Array Cosmic Microwave Background Data with Inclusion of 95 GHz Band. 2016, Phys. Rev. Lett., 116, 031302, arXiv:1510.09217

BICEP2 Collaboration, Ade, P. A. R., Aikin, R. W., et al., BICEP2. II. Experiment and three-year Data Set. 2014a, *Astrophys. J.*, 792, 62, arXiv:1403.4302

BICEP2 Collaboration, Ade, P. A. R., Aikin, R. W., et al., BICEP2. II. Experiment and three-year Data Set. 2014b, *Astrophys. J.*, 792, 62, arXiv:1403.4302

BICEP2, Keck Array and Planck Collaborations, Joint Analysis of BICEP2/Keck Array and Planck Data. 2015, Phys. Rev. Lett., 114, 101301, arXiv:1502.00612

Carlstrom, J. E. et al., The 10 Meter South Pole Telescope. 2011, Publ. Astron. Soc. Pac., 123, 568, arXiv:0907.4445

Challinor, A. et al., Exploring cosmic origins with CORE: gravitational lensing of the CMB. 2017, arXiv:1707.02259

Chiang, H. C. et al., Measurement of CMB Polarization Power Spectra from Two Years of BICEP Data. 2010, Astrophys. J., 711, 1123, arXiv:0906.1181

Day, P. K., LeDuc, H. G., Mazin, B. A., Vayonakis, A., & Zmuidzinas, J., A broadband superconducting detector suitable for use in large arrays. 2003, Nature, 425, 817

Eriksen, H. K., Banday, A. J., Gorski, K. M., & Lilje, P. B., Foreground removal by an internal linear combination method: Limitations and implications. 2004, Astrophys. J., 612, 633, arXiv:astro-ph/0403098

Essinger-Hileman, T., Kusaka, A., Appel, J. W., et al., Systematic effects from an ambient-temperature, continuously rotating half-wave plate. 2016, Review of Scientific Instruments, 87, 094503, arXiv:1601.05901

Fowler, J. M. et al., Optical Design of the Atacama Cosmology Telescope and the Millimeter Bolometric Array Camera. 2007, Appl. Opt., 46, 3444, arXiv:astro-ph/0701020

Ghosh, T., Boulanger, F., Martin, P. G., et al., Modelling and simulation of large-scale polarized dust emission over the southern Galactic cap using the GASS Hi data. 2017, *Astron. Astrophys.* , 601, A71, arXiv:1611.02418

Grace, E., Beall, J., Bond, J. R., et al. 2014, in Proc. SPIE, Vol. 9153, Millimeter, Submillimeter, and Far-Infrared Detectors and Instrumentation for Astronomy VII, 915310

Hanany, S., Niemack, M. D., & Page, L. 2013, in Planets, Stars and Stellar Systems. Volume 1: Telescopes and Instrumentation, ed. T. D. Oswalt & I. S. McLean (Springer), 431

Henning, J. W. et al., Measurements of the Temperature and E-Mode Polarization of the CMB from 500 Square Degrees of SPTpol Data. 2017, Submitted to: Astrophys. J., arXiv:1707.09353

Hensley, B. 2015, PhD thesis, Princeton University

Irwin, K. D., An application of electrothermal feedback for high resolution cryogenic particle detection. 1995, Applied Physics Letters, 66, 1998

Irwin, K. D. & Lehnert, K. W., Microwave SQUID multiplexer. 2004, Applied Physics Letters, 85, 2107

Johnson, B. R., Collins, J., Abroe, M. E., et al., MAXIPOL: Cosmic Microwave Background Polarimetry Using a Rotating Half-Wave Plate. 2007, *Astrophys. J.*, 665, 42, arXiv:astro-ph/0611394

Kamionkowski, M., Kosowsky, A., & Stebbins, A., A Probe of primordial gravity waves and vorticity. 1997, Phys.Rev.Lett., 78, 2058, arXiv:astro-ph/9609132

Kesden, M., Cooray, A., & Kamionkowski, M., Separation of gravitational wave and cosmic shear contributions to cosmic microwave background polarization. 2002, Phys. Rev. Lett., 89, 011304, arXiv:astro-ph/0202434

Knox, L. & Song, Y.-S., A Limit on the detectability of the energy scale of inflation. 2002, Phys. Rev. Lett., 89, 011303, arXiv:astro-ph/0202286

Kritsuk, A. G., Ustyugov, S. D., & Norman, M. L., The structure and statistics of interstellar turbulence. 2017, New Journal of Physics, 19, 065003, arXiv:1705.01912

Kusaka, A., Essinger-Hileman, T., Appel, J. W., et al., Modulation of cosmic microwave background polarization with a warm rapidly rotating half-wave plate on the Atacama B-Mode Search instrument. 2014, Review of Scientific Instruments, 85, 024501

Kusaka, A. et al., Modulation of cosmic microwave background polarization with a warm rapidly rotating half-wave plate on the Atacama B-Mode Search instrument. 2014, Rev. Sci. Instrum., 85, 024501, arXiv:1310.3711

Niemack, M. D., Designs for a large-aperture telescope to map the CMB 10x faster. 2016, *Applied Optics*, 55, 1688

Ogburn, IV, R. W. et al., BICEP2 and Keck Array operational overview and status of observations. 2012, Proc. SPIE Int. Soc. Opt. Eng., 8452, 1A, arXiv:1208.0638

Planck Collaboration, Planck intermediate results. L. Evidence of spatial variation of the polarized thermal dust spectral energy distribution and implications for CMB B-mode analysis. 2017, *Astron. Astrophys.* , 599, A51, arXiv:1606.07335

Seljak, U. & Zaldarriaga, M., Signature of gravity waves in polarization of the microwave background. 1997, Phys.Rev.Lett., 78, 2054, arXiv:astro-ph/9609169

Sheehy, C. & Slosar, A., No evidence for dust B-mode decorrelation in Planck data. 2017, ArXiv e-prints, arXiv:1709.09729

Smith, K. M., Hanson, D., LoVerde, M., Hirata, C. M., & Zahn, O., Delensing CMB Polarization with External Datasets. 2012, JCAP, 1206, 014, arXiv:1010.0048

Thorne, B., Dunkley, J., Alonso, D., & Naess, S., The Python Sky Model: software for simulating the Galactic microwave sky. 2016, arXiv:1608.02841

Vansyngel, F., Boulanger, F., Ghosh, T., et al., Statistical simulations of the dust foreground to cosmic microwave background polarization. 2017, Astron. Astrophys., 603, A62, arXiv:1611.02577

Yoon, J., Clarke, J., Gildemeister, J., et al., Single superconducting quantum interference device multiplexer for arrays of low-temperature sensors. 2001, Applied Physics Letters, 78, 371

APPENDICES

A SIMULATIONS

Here we present our case that the strawperson concept achieves the science requirements. The concept evolved from the process we developed to flow down science requirements to measurement requirements and instrument requirements, a process that began with our work on the Science Book. We have extended that work to improve our understanding of the impact of astrophysical foregrounds, instrumental systematics, delensing non-idealities, and analysis methodology.

We begin with a high-level summary of the key points, first regarding the surveys motivated by gravitational waves, and second regarding the survey motivated by light relics.

Key points regarding the gravitational wave surveys of the strawperson concept:

- The allocation of detector effort across frequency for the strawperson concept largely follows from the optimization process described in the Science Book, with some important adjustments we describe below.
- Analysis of multi-frequency map-based simulations reproduce the semi-analytic forecasts of $\sigma(r)$ used in the Science Book in the case of what we will call the "Science Book Configuration" of the CMB-S4 instrument and survey (defined below) and Sky Model 0, a Gaussian realization of the foreground model assumed in the Science Book.
- Using multi-frequency map-based simulations with the Science Book Configuration and a suite of foreground models, two different analysis methods find bias in recovered r at acceptable levels. For sky models with more complicated foreground behavior, both analysis methods find an increase in $\sigma(r)$ by a factor between 1.1 and 3 over the result using Sky Model 0.
- To achieve the science requirement on r for all baseline sky models, we increased the total effort for the gravitational wave surveys in the strawperson concept by a modest amount compared to the Science Book Configuration.
- The primordial B-mode signal (the signal of interest) is dominated by foreground emission at all frequencies, even in the cleanest regions of the sky. Unexpected properties of the foregrounds remain a risk. For example, although we expect the thermal dust emission to be highly correlated across frequency, extreme levels of decorrelation (that are not favored by current data and appear unlikely based on physics considerations, but that do not violate any direct empirical constraints) as implemented in Sky Model 5, lead to an increase in $\sigma(r)$ by another factor of 2 for the Science Book Configuration. We emphasize that the aspect of this model that leads to these results is unlikely to be realized in nature, but the model serves as an existence proof that unexpected foreground properties can degrade our expected results significantly.
- One of the baseline sky models (Sky Model 6), analyzed in the most highly conservative manner, shows unacceptably large bias with the Science Book Configuration. The bias arises due to synchrotron emission residuals at $\ell \sim 100$ to 150. The strawperson concept thus has the 20 GHz channel on the 6-m telescope to improve sensitivity in this ℓ range.

Key points regarding the light relics survey of the strawperson concept:

- The primary drivers for detection of light relics are raw sensitivity and sky fraction. The requirement of $\Delta N_{\rm eff} \leq 0.06$ at the 95% confidence level drives the large-area survey to a map noise level of $\sim$1 μK-arcmin (in temperature) over $\sim$40% of the full sky.

- Foregrounds are not expected to have a major impact on the $N_{\rm eff}$ constraint, but foreground monitor channels are included in the strawperson concept of the light relics survey, particularly to preserve the CMB lensing and large-scale-structure-science aspects of this survey.
- The most important systematic effect for the $N_{\rm eff}$ constraint is uncertainty on the effective instrument beam, particularly as a function of position in the map and as affected by pointing variations. Simulations of beam characterization using planets and sources in the field indicate that these systematics can be adequately controlled in the strawperson configuration.

A.1 Small-Area Survey Targeting the Degree-Scale Signature of Gravitational Waves

Using simulations to optimize the design of a CMB experiment inevitably involves a trade-off between the degree of detail that the simulations are able to capture and the computational (and human) cost of generating and analyzing them. This trade-off includes the choice of domain in which the simulation is generated, ranging from the most detailed but most expensive time domain through the map domain to the most simplified but most flexible spectral domain. Inclusion of additional detail can help to validate general results, to explore their sensitivity to assumptions about foreground models, sky coverage, and instrumental noise and systematics, and in more mature stages of design can inform specific instrument and survey strategy choices.

For the Science Book work an r-forecasting machinery was assembled based on scaling the bandpower covariance matrices of published BICEP/Keck analyses. Given a defined set of bandpasses, and assumptions about foreground power spectra, this semi-analytical approach is capable of optimizing detector allocation across frequency, and across delensing and degree-scale surveys, for the lowest $\sigma(r)$ at fixed effort, where 'effort'= total number of 150-GHz-equivalent detector-years of observation. (It makes sense to define 'effort' in these units since it is equivalent to focal plane area, which is in turn the strongest driver of overall project cost.)

For this report we have extended this work to map-domain simulations in order to be able to capture additional complexities that cannot be represented in the spectral domain, while remaining computationally tractable. These complexities include:

- Non-Gaussianity and statistical anisotropy of the Galactic foregrounds.
- Instrumental systematic effects.
- Inconsistency between the data and the assumptions (either explicit or implicit) of any given analysis method.
- Foreground contamination in the delensing survey.

We also use these simulations to validate the spectral domain forecasts for configurations where the approaches are directly comparable.

Here we review the methods used to explore parameter space for the small-area survey, including map level noise simulations, sky models, and observation strategy. We also describe our approach to modeling instrumental systematics, the delensing survey, and the analysis methods.

A.1.1 **Map noise simulations** The Science Book assumed 1×10^6 150-GHz-equivalent detector-years of observation in 30, 40, 85/95, 145/155, 220 & 270 GHz bands (where the split bands are assumed to be overlapping). For this work we experimented with allowing a 20 GHz band in the semi-analytic optimization and found a small benefit, as we will describe below, so this is now part of the strawperson concept. The optimization specifically includes the need to delens, and assigns a fraction of the detectors to a high-resolution band (typically about a third of them). We use the BICEP/Keck published noise power spectra (N_ℓ's) at 95 and 150 GHz, together with the

actual calendar years of observation and total number of deployed detectors for those datasets, and scale these to other bands using only the *ratios* of ab initio NET projections. The resulting net observing efficiency is typically 20% or less, but the information loss and noise spectrum depend in detail on angular scale. This automatically builds in all real world inefficiencies, including (but not limited to) imperfect detector yield, non-uniform detector performance, read-out noise, observing inefficiency, losses due to timestream filtering, beam smoothing, and non-uniform sky coverage.

To produce map-level simulations it is necessary to translate the BICEP/Keck noise bandpowers into a prescription for map noise. We do this by fitting the N_ℓ's to a white $+\,\ell^\gamma$ model accounting for beam smoothing, etc. For the small-aperture BICEP/Keck data, we find $\ell_{\rm knee} = 50$–60 with γ of -2 to -3. To translate to map noise levels, we must pick a specific effective sky area. As in the Science Book, we have mostly used 3% with some simulations also done for 1% and 10%. We then generate Gaussian noise realizations at each band. Small-aperture cameras have a very wide instantaneous field of view and hence the observed sky region necessarily has a large edge taper. For the nominal 3% sky coverage simulations, we assumed a circular sky patch with full coverage at $r < 12^\circ$ and "relative hits" tapering to zero with a cosine-squared shape for $12^\circ < r < 27^\circ$. The noise realizations are divided by the square-root of this coverage pattern such that the noise "blows up around the edge" as it does in real maps.

We refer to the simulations as described above, including the extension to a 20 GHz channel on the small-aperture telescope, as simulations of the "Science Book Configuration."[3] This is to distinguish it from the strawperson configuration, which has both a larger number of 150 GHz-equivalent detector-years and a 20 GHz channel with a resolution (and $\ell_{\rm knee}$) appropriate for a 6m aperture.

A.1.2 **Foreground models** To make simulated sky maps we add realizations of lensed CMB both without and with an r component to models of the Galactic foregrounds. So far we have run simulations with seven foreground models:

0. Simple Gaussian realizations of synchrotron and dust with power-law angular power spectra at amplitudes set to match the observations in the BICEP/Keck field, and simple uniform SEDs (power law for synchrotron, greybody for dust).

1. The PySM[4] model `a1d1f1s1`, where the letters refer to anomalous microwave emission, dust, free-free and synchrotron respectively, and the numbers are the base models described in Thorne et al. (2016).

2. The PySM model `a2d4f1s3`, where the models have been updated to variants that are also described in Thorne et al. (2016). Note that these include 2% polarized AME, a curvature of the synchrotron SED, and a two-temperature model for dust.

3. The PySM model `a2d7f1s3`, where the dust model has been updated to a sophisticated physical model of dust grains as described in Hensley (2015). This model is interesting in that it does not necessarily conform to the greybody SED.

4. The dust model in 3 is replaced by a model of polarized dust emission that incorporates HI column density maps as tracers of the dust intensity structures, and a phenomenological description of the Galactic magnetic field as described in Ghosh et al. (2017). The model is expanded beyond that described in the paper to produce a modest amount of decorrelation

[3] Note that this configuration differs slightly from the one that appears in the 2016 edition of the Science Book, due mainly to updates in frequency band definitions, but the total assumed effort is the same.

[4] https://github.com/bthorne93/PySM_public

of the dust emission pattern as a function of frequency motivated by the analysis of Planck data in Planck Collaboration (2017).

5. A toy model where the dust decorrelation suggested in Figure 3 of Planck Collaboration (2017) is taken at face value ($\mathcal{R}_{80}^{217\times353} = 0.85$). While such a model is not ruled out by current data it appears to be very hard to produce such strong decorrelation in physics-based models. We also note that Sheehy & Slosar (2017) have re-analyzed the same Planck data and, while they find that the high level of decorrelation in this model is consistent with the data, their best fit to that same data has no decorrelation.

6. A model based on MHD simulations (Kritsuk et al. 2017) of the Galactic magnetic field, which naturally produces non-Gaussian correlated dust and synchrotron emission.

Models 1 to 4 use the actual large-scale modes of the real sky as measured above the noise in the Planck data. This means that these models are intrinsically "single-realization," and this must be borne in mind when interpreting the results. The PySM models fill in the small-scale structure with power-law Gaussian extrapolations, while models 4 and 6 naturally produce non-Gaussian small-scale structure. However, all of these models are consistent with current data, and we should be careful not to necessarily associate nominal sophistication with greater probability to more closely reflect reality.

A.1.3 **Instrumental systematics** As described in § 3, control of instrumental systematics is a critical design consideration. However, predicting and modeling these effects realistically is a difficult task that is dependent on actual instrument and survey design details, and in any case their impact on an actual result comes not through the modeled effects but through unmodeled residuals. For this study we have taken the first steps in simulating various generic classes of additive systematic by injecting additional noise-like components into the maps, and then re-analyzing them without knowledge of what was put in. We have experimented with components that are both correlated and uncorrelated across frequency bands, and which have white, $1/\ell$, and white + $1/\ell$ spectra, at varying levels compared to single-frequency map noise or, for correlated cases, combined map noise. Examples of mechanisms that might produce map residuals within this class, after modeling them and either correcting or filtering their leading-order effects, include bandpass mismatches, beam and pointing variations, calibration variations, cross-talk effects, half-wave-plate leakages, ground pickup, and readout irregularities.

Other classes of systematics can be simulated by manipulating the reanalysis procedure only. Examples of such effects include bandpass, polarization angle, calibration, and beam shape uncertainties.

A.1.4 **Delensing** As described in § A.1.1, we assume a separate, high-resolution instrument dedicated to measuring the medium- and small-scale information necessary to construct a template of lensing B modes, so that their effect can be removed. In the semi-analytic optimization process, this instrument is assumed to have 1-arcminute resolution and detector weight at a single frequency. The translation between detector effort and map noise in the delensing instrument is based on the method used for the low-resolution instrument, but with certain non-idealities specific to low-resolution instruments and low-ℓ analysis (such as mode removal and non-uniform coverage) removed. Following the formalism in Smith et al. (2012), we convert the map noise in the delensing survey to a delensing efficiency, or equivalently a fractional residual in lensed B-mode power. We have started to generate high-resolution simulated maps on which we can run explicit lensing reconstruction and then include that information in the analysis. However, that process is not yet

converged, and so for the present we approximate delensing by scaling down the ΛCDM lensing signal by the appropriate factor.

A.1.5 **Analysis methods** To make simulated maps the noise realizations described in § A.1.1 are added to the sky models described in§ A.1.2. For each realization one then has a stack of multi-frequency $I/Q/U$ maps containing non-uniform noise, foregrounds and signal, and the challenge is to re-analyze them to recover the parameter of interest (in this case r). This can be done by different teams using different methods, and could be done in a blind manner, although we have not done this yet.

So far we have experimented with two methods. The first is a map-based ILC cleaning method (e.g., Eriksen et al. 2004), which seeks the linear combination of maps that minimizes the remaining CMB signal, followed by a marginalization over residual foregrounds. This method has the advantage that it does not need to know the bandpasses of the frequency channels.

The second method is an evolution of the parametric multi-component fit to the ensemble of auto- and cross-spectra as used for the BICEP/Keck analysis to date (BICEP2, Keck Array and Planck Collaborations 2015; BICEP2 and Keck Array Collaborations 2016). This method fits the observed bandpowers to a model composed of the lensing expectation plus dust and synchrotron contributions and a possible r component. Dust and synchrotron each have an amplitude (A_d and A_s), a spatial spectral parameter (α_d and α_s), and a frequency spectral parameter (β_d and β_s). We also allow dust/synchrotron correlation (ϵ), and decorrelation of the foreground patterns over frequency (δ_d and δ_s).

Both of these analysis methods are only close to optimal when the foreground behavior is close to uniform across the observing field. For analysis of larger fields, algorithms that fit, for example, the frequency spectral indices individually in (large) pixels, will be required.

A.1.6 **Results** Table 7 summarizes the results of the analysis for simulations of the Science Book Configuration (1.0×10^6 150-GHz-equivalent detector-years) and residual lensing power $A_L = 0.1$. We see that for $r = 0$ the simple Gaussian foreground model 0 gives $\sigma(r) = 5 \times 10^{-4}$, exactly as expected from the semi-analytic code. As we progress to the more complex foreground models, $\sigma(r)$ is generally in the range 7–8$\times 10^{-4}$. To satisfy the science requirement, we have therefore increased the number of detector-years in the strawperson concept, and checked in one of the map based pipelines that this does indeed return $\sigma(r)$ to 5×10^{-4} (see below). For at least one analysis method, the bias remains at $\lesssim 0.3\sigma$ for all the models. (These simulations are sets of 500 realizations, so the statistical uncertainty on the bias is $\approx 0.04\sigma$.) However, the strong decorrelation model 5 does significantly increase $\sigma(r)$. While the parametric method is able to account for the decorrelation, by construction information is lost, and in fact if one believed in such a scenario, re-optimization to concentrate the sensitivity at closer-in frequencies would be called for.

Table 8 shows results for the strawperson configuration for sky model 6 and the ILC analysis method. For $r = 0$, the 95% upper limit is about $2.1\sigma(r)$. The value of the tensor-to-scalar ratio for which we expect a 5σ detection is expected after 4 years of operation is $r = 0.004$. For a tensor-to-scalar ratio of $r = 0.003$, the median detection significance after 4 years is expected to be 4σ. If a detection were to be emerging at this point, extending the run time to 8 years would be justified to reach a 5σ detection.

While for given assumptions $\sigma(r)$ can be precisely forecast, the achieved detection level for r depends on the realization of the B-mode field in the observed patch of sky and the instrument noise. Therefore we can only forecast a distribution of detection levels. For a tensor-to-scalar ratio of $r = 0.003$ and the strawperson concept with 8 years of observing we expect to achieve more than 3σ detection with a probability of 0.99, more than 4σ with a probability of 0.93, more than 5σ with a probability of 0.53, and more than 6σ with a probability of 0.14. For simplicity we focus on $\sigma(r)$,

Table 7: Results of two analysis methods applied to map-based simulations assuming the Science Book Configuration and our suite of sky models. All simulations assume an instrument configuration including a (low-resolution) 20 GHz channel, a survey of 3% of the sky with 1.0×10^6 150-GHz-equivalent detector-years, and $A_L = 0.1$. Note that this configuration is not the final strawperson concept, and in particular has fewer detector-years.

		ILC		Parametric	
r value	Sky model	$\sigma(r) \times 10^4$	r bias $\times 10^4$	$\sigma(r) \times 10^4$	r bias $\times 10^4$
0	0	5.7	0.0	6.7	0.2
	1	7.0	0.3	7.8	5.8
	2	7.7	0.8	7.1	3.1
	3	5.6	0.8	8.1	1.8
	4	7.5	5.0	9.3	−3.4
	5[a]	16	18	14	−2.5
	6	5.8	−1.1	7.3	1.1
0.003	0	7.2	−4.0	10	0.3
	1	9.1	0.0	9.0	6.2
	2	9.6	−1.9	9.4	3.5
	3	7.2	−0.3	10	1.6
	4	10	5.8	11	−1.8
	5[a]	20	20	15	3.0
	6	8.3	−1.1	9.9	1.1

[a] An extreme decorrelation model—see § A.1.2. The parametric analysis includes a decorrelation parameter. No attempt is made in the ILC analysis to model decorrelation.

and on median detection levels as well as median 95% confidence upper limits to state the typical outcome.

Table 8: Results on detection significance for the strawperson concept selected for CMB-S4, using the ILC analysis method. Note that this has an increase in detector-year effort versus the configuration in Table 7.

r value	Duration	Sky model	$\sigma(r) \times 10^4$	r bias $\times 10^4$	95% CL UL	Detection Significance
0	4 years	6	4.7	0.5	1.0×10^{-3}	. . .
0.003	4 years	6	6.9	−1.2	. . .	4.0
	8 years	6	5.9	0.4	. . .	5.1

The numbers in Table 7 clearly show dependence on the foreground model used in the simulation. If the actual foregrounds are substantially different from any of these cases, then the biases could be larger. To get some understanding of how large the biases could be, and what instrument modifications might help to reduce them, we have also looked at ILC biases in the extreme case that the foreground residuals are not modeled, but simply absorbed into the estimated B-mode power spectrum. Doing so with simulations based on Sky Model 6 increases the magnitude of the bias on r from 0.4×10^{-4} to 4×10^{-4}. The dominant contribution to the bias comes from synchrotron residuals. To reduce this bias, we explored the possibility of placing one or two lower-frequency channels on the large-aperture dish. Doing this with 20 GHz improves the angular resolution from 76.6' to 11', while also increasing the noise power $\ell_{\rm knee}$ from 50 to 200. We find that this change reduces the magnitude of the bias to 1.3×10^{-4}.

Table 9 summarizes the results of analysis of simulations including additive systematic effects, in different combinations of uncorrelated and correlated contamination with varying spectra, added

on top of foreground model 3. The levels of systematic contamination for these simulations were chosen to predict biases on r of 1×10^{-4} in semi-analytic forecasts. We can see that the different combinations explored increase biases on r by amounts that typically vary from 0.5–1.5$\times 10^{-4}$ for the two different analyses, over the different cases. We find that to restrict bias on r to this level, the sum of additive contamination effects needs to be controlled to 3–7% of the single-frequency survey noise, or (in the case of correlated systematics) 6–11% of the total combined noise levels. Such percentages are consistent with the upper limits currently achieved for residual additive systematic contamination compared to survey noise by small-aperture experiements (e.g., BICEP2 and Keck Array Collaborations 2016). Assuming CMB-S4 will include a sustained effort to continue to control, understand, and model systematic effects down to levels limited by survey noise, these percentages provide reasonable benchmark requirements.

Table 9: Map-based simulation results for simulations containing systematics. Simulations are as in Table 7 for sky model 3 and $r = 0$, with additive systematic effects in varying combinations, the amplitudes of which are specified as percentages of survey noise.

	Uncorrected		Corrected		ILC		Parametric	
Systematic	A [%]	B [%]	A [%]	B [%]	$\sigma(r) \times 10^4$	r bias $\times 10^4$	$\sigma(r) \times 10^4$	r bias $\times 10^4$
None	0	0	0	0	5.3	0.0	7.2	0.0
Uncorrelated white	3.3	0	0	0	6.0	0.84	8.0	0.63
Uncorrelated $1/\ell$	0	6.8	0	0	5.0	0.99	7.0	0.85
Correlated white	0	0	5.8	0	6.3	1.2	7.3	1.41
Correlated $1/\ell$	0	0	0	10.5	5.2	1.0	6.7	0.97
Uncorrelated white + $1/\ell$	1.6	3.5	0	0	5.6	0.89	7.5	0.76
Correlated white + $1/\ell$	0	0	2.9	5.3	5.5	0.98	6.9	1.04
Both, white + $1/\ell$	0.8	1.7	1.5	2.6	5.6	1.1	7.9	0.98

Results of simulating systematic errors in the determination of bandpasses vary by analysis method. The construction of the ILC method makes it largely insensitive to such errors. The parametric analysis, which includes parametrized models of the frequency spectra of different foregrounds, shows biases on r at the 1×10^{-4} level for uncorrelated random deviations in bandcenter determination of 0.8%, or for correlated deviations of 2%, which we adopt as reasonable benchmark requirements to accommodate a variety of both blind and astrophysical foreground modeling approaches.

In the current simulations, the treatment of delensing is still somewhat crude. We have explored alternate methods of translating detector effort to map noise, including scaling directly from the noise in a fielded, high-resolution experiment (SPTpol), but with some assumed modifications in per-detector sensitivity. The scaling using this alternate method is slightly more pessimistic than the default scaling, and we adopt this more pessimistic scaling for the numbers in the Science and Measurement Requirements section (§ 3). We have also included multiple frequency bands in the delensing instrument in § 3, because we cannot conclusively rule out the possibility that non-Gaussianity in small-scale Galactic foregrounds will cause a bias in a single-frequency delensing survey. The level of delensing assumed in the optimization code is reproduced with the configuration in § 3 if the information from the two "CMB channels" (95 and 145 GHz) can be combined optimally for CMB sensitivity. If instead we assume we have to combine frequency bands to explicitly project out a dust component and a synchrotron component, the delensing efficiency degrades by 5–10%. Effort is currently underway to include delensing with fully non-Gaussian small-scale foregrounds in the Data Challenges. Separate simulations with one model of fully non-Gaussian dust (Vansyngel

et al. 2017) indicate that biases to delensing are negligible (also see Challinor et al. 2017), but a larger parameter space of foreground models will need to be explored.

A.2 Large Area Survey

In the future, the simulated maps described above will be generated on larger sky areas and re-analysed to recover $N_{\rm eff}$ and neutrino mass. However, for the present we rely on the forecasting techniques for these parameters that are described in the Science Book. As indicated in the Science Book, the primary drivers for $N_{\rm eff}$ are raw sensitivity and sky fraction. For fixed detector effort, and neglecting foregrounds and instrument systematics, $\sigma(N_{\rm eff})$ decreases monotonically as sky fraction is increased, so the large-area survey for all forecasts and for the strawperson instrument configuration described in § 3 is assumed to cover the largest area of clean sky obtainable from Chile, or approximately 40% of the full sky. The map noise level required to achieve the target $\Delta N_{\rm eff} \leq 0.06$ at the 95% confidence level is approximately $1\,\mu$K-arcmin ($1\,\mu$K rms in a map with $1'$ pixels). To determine the detector effort needed to achieve this map noise level, we use the same scaling from achieved SPTpol results as used in the delensing survey.

The $N_{\rm eff}$ constraint is expected to be relatively insensitive to foregrounds. The primary observables contributing to this constraint are the TE and EE power spectra at moderate to high ℓ (medium to small angular scales); at these scales the dominant polarized foreground is expected to be extragalactic radio sources. Contamination from these sources is trivially modelable in ℓ space, and the overall amplitude is constrained to be less than the EE power out to $\ell \simeq 4000$ with moderate source masking (Henning et al. 2017). There is no mean contamination to TE from extragalactic sources; they do add variance to TE, but that variance is small compared to the noise in a $1\,\mu$K-arcmin map. On the other hand, the same large-area survey will be used for reconstructing the CMB lensing potential and for extracting the legacy large-scale-structure and astrophysical information discussed in § 2. The lensing information comes from higher-order correlations in the CMB map, which could be contaminated by the very faint but non-Gaussian small-scale structure in Galactic foregrounds. Meanwhile, multiple frequency bands are crucial for teasing apart the tSZ and kSZ signals that underlie the legacy large-scale-structure impact of CMB-S4. For these reasons, foreground-monitor channels are added above and below the primary CMB frequencies of 90 and 150 GHz in the strawperson discussed in § 4, though we emphasize that under baseline foreground assumptions, information at all bands can be optimally combined for CMB sensitivity for the $N_{\rm eff}$ constraint.

Investigations of the impact of multiplicative systematics on parameters extracted from the large (40% sky fraction) survey indicate that knowledge of the beam is the most critical, particularly for $N_{\rm eff}$ and astrophysical sources such as galaxy clusters, and simulations of beam characterization from planets and sources in the field have been conducted. These simulations indicate that the levels of beam knowledge required are achievable, even in the presence of pointing variations over time and beam variations across the field, for the strawperson configuration in § 4.

B TECHNOLOGY DEVELOPMENT

A high-level overview of a plan for technology investment in a pre-CD0 phase is presented in § 9. This plan is prioritized to reduce risk in the strawperson instrument concept for CMB-S4. This appendix provides more details of status and development in different technology areas: focal-plane sensors and readout, focal-plane optical coupling, telescopes, camera cryogenics and optics, and data management. We refer the reader to the CMB-S4 Technology Book for a more detailed description of the current state of technological readiness.

B.1 Focal-Plane Sensors and Readouts

Existing experiments in the field have achieved background-limited noise performance in representative CMB frequency bands at both the South Pole (BICEP2 Collaboration et al. 2014a) and the Atacama (Grace et al. 2014). These experiments use superconducting Transition-Edge Sensors (TES) (Irwin 1995) read out with multiplexed Superconducting Quantum Interference Device (SQUID) amplifiers. While this core capability has been fully demonstrated on multiple instruments, investment will reduce schedule, performance, and cost risk in the development and production of integrated detector modules for CMB-S4.

Sensors—CMB polarimeter wafers for CMB-S4 must be produced with high wafer-fabrication yield and low sensor parameter variation across the wafer (e.g., uniform saturation power, normal resistance, transition temperature, etc.). The fabrication yield must be sustained in order to produce the large number of arrays needed for CMB-S4. The polarimeters must maintain their high yield and performance when integrated with readout elements in detector modules. Investment is needed in improving process uniformity and reproducibility in TES polarimeter fabrication. In addition, cost reduction could potentially be achieved either by maturing Microwave Kinetic Inductance detectors (MKIDs) (Day et al. 2003), which have simpler integration with readout elements, or by exploring commercial fabrication options.

Readouts—The final yield of the CMB-S4 detector modules is affected by the multiplexed readout, the integration complexity, and the wiring yield. The multiplexer components are one of the most challenging elements of Stage-2 and Stage-3 experiments. These risks can be mitigated by implementing cold multiplexer components with higher multiplexing factors, higher fabrication yield, more consistent noise performance, and more reliable interconnects and room-temperature electronics. Investment is needed in the area of multiplexer cryogenic components. The two leading candidates for the multiplexed readout of CMB-S4 are microwave SQUID multiplexers (μMUX) (Irwin & Lehnert 2004) operated at GHz frequencies, and Frequency Division Multiplexers (FDM) (Yoon et al. 2001) operated at MHz frequencies. Investment is needed to mature both of these technologies. In addition, cost reduction could potentially be achieved either by maturing MKIDs (which have integrated readout elements), or exploring commercial fabrication options.

B.2 Focal-Plane Optical Coupling

Development is needed in focal-plane optical coupling. Optical coupling elements, with functions including beam forming, polarization separation, and band definition, need to be developed for all instrument frequencies. Improvements are needed in high-throughput fabrication and testing, and trade studies are needed to consider different technological options. We refer the reader to the CMB-S4 Technology Book for details.

B.3 **Telescopes**

Multiple telescope architectures have been successfully employed in the current generation of CMB polarization experiments. Arcminute-scale telescope resolution is enabled by large reflector (mirror-based) optics while both refractor (lens-based) and reflector systems have been utilized in lower angular resolution settings. Refractor systems are also used for camera optics. Both reflector and refractor technologies are well established and at a high level of technical maturity. For reflector-based optical telescopes, both standard Gregorian and crossed Dragone designs have successfully been implemented. Refinements of the candidate low- and high-resolution optical system designs that accommodate large detector arrays and achieve high polarization purity represent ongoing activities in the community. In particular, continuation of these developments and focused design efforts to improve optical baffling/sidelobe rejection, calibration of angular response and polarization orientation, and optical system stability over the state of the art would add value in preparation for CMB-S4.

B.4 **Camera Cryogenics and Optics**

Camera cryogenics—For the large-aperture telescope discussed in Sect. 4.2, a large $\sim$ 3-m-scale cryostat is necessary in order to take full advantage of the telescope throughput. Cryostats with large vacuum jackets have been used in other instruments, but need to be developed for the CMB-S4 receivers. This is a significant engineering challenge, particularly in the areas of vacuum windows with large total area, optical thermal loading through the window, and the large volume of the sub-kelvin cryogenic stage. Design and prototyping studies in these areas will mitigate the project risks both in cost and schedule.

The small-aperture cameras do not require a cryostat as large as the large-aperture telescopes, because there is no strong driver toward densely packed vacuum/optical windows. This flexibility enables different optimization of the cryostat configuration to mitigate instrument complexity, project cost, and risk. Design and prototyping studies should also be conducted for small-aperture camera cryostats.

Stage-2 and Stage-3 experiments use either 3He/4He dilution refrigerators or 3He/4He sorption refrigerators for sub-K cooling. These technologies are commercially available, but CMB-S4 would benefit from possible progress by commercial vendors to reduce the cost and to increase the cooling power. Early investment in sub-kelvin cooling technology is not required, however, to meet the technical specifications of the CMB-S4 project, and the focus of the development should be toward the system design of the cryostat equipped with commercially available cooling technologies.

Variations in either the sub-kelvin focal plane or camera optics temperature will lead to long-timescale and common-mode noise in the science data. Suppression of this undesired signature at levels commensurate with the sensitivity of the CMB-S4 focal planes is needed. Stability and monitoring of the cryogenic stage temperature thus presents another challenge and investments in this area would mitigate performance risk.

Camera optics—Investment in camera optics can improve system performance and component manufacturability, and reduce overall cost. The required sub-kelvin operating temperature of the multichroic focal planes and the large throughput required at millimeter wavelengths necessitates the further development and validation of appropriate broad-band, large-aperture refractive cryogenic optics ($\gtrsim$ 40 cm), antireflection coatings, optical and wide stop-band, on-chip, band-pass filters (also see Focal-Plane Optical Coupling section above), and thermal blocking filters. Attention to high optical throughput in the CMB-S4 signal bands in these efforts has the potential to improve system margins and reduce risk.

Polarization modulators—Polarization modulation can be employed as a tool to mitigate beam-induced systematic errors and $1/f$ noise. The BICEP/Keck collaboration demonstrated white Q, U noise at 90 and 150 GHz for $\ell > 30$ at the pole without an active modulator (BICEP2 Collaboration et al. 2014b). Currently no CMB B-mode results, either at Chile or the Pole, have been demonstrated in the higher frequency bands (220 and 270 GHz), where increased atmospheric noise and reduced stability are anticipated. A modulator may turn out to be required to operate at these frequencies at the noise levels targeted by CMB-S4. With a modulator, very low ℓ values are accessible, as demonstrated by ABS (Kusaka et al. 2014). The BICEP/Keck collaboration demonstrated systematic discrimination of I to Q/U at a level of $r = 0.003$ without a modulator. This level was limited by noise, not residual systematics. As low as it is, it is not yet a demonstration at the levels required for CMB-S4. The ABS collaboration demonstrated discrimination of beam induced systematics to a level of $r < 0.001$ with a continuously rotating half wave plate (HWP) (Essinger-Hileman et al. 2016).

Optical elements intrinsically create instrumental polarization through polarized emission, differential transmission of the intensity signal, and diffractive scattering. Spurious instrumental polarization leads to low-frequency excess signatures, which must ultimately be mitigated through the detailed instrument design and observing strategy. Data from ABS, Polarbear, and EBEX suggest that if a continuously rotating modulator is used, it is desirable to be implemented as the first element in the optical path. However, use of a modulator as a first element in the optical path is not deemed practical for large-aperture telescope designs. Implementing a HWP with any optical system requires careful attention to the modulator bandwidth and polarization purity. Active efforts to improve the performance of polarization modulators are underway.

In summary, a modulator has been shown to mitigate low frequency noise and mitigate beam-induced systematics. But today's best constraints on r have been obtained without a modulator. At this time, it is unknown whether implementing a continuously rotating modulator will be (i) absolutely necessary, (ii) necessary in some cases, (iii) desirable to have, or (iv) not necessary. The prudent course of action is to continue the development and field evaluation of modulators, retain them in cost models and as a design option, and defer a final decision pending a comparative technical evaluation of data from ongoing efforts.

B.5 **Data Management**

Any proposed CMB-S4 experimental configuration (comprising both instrument and observation) must be validated by first simulating a synthetic dataset that reflects that configuration, and then demonstrating that it is consistent with the science requirements. While approximate methods provide essential guidance at low computational cost, ultimately the only way to capture the full impact of the statistical and systematic uncertainties introduced by the experiment and the data processing is to simulate the raw observations and pass them through the full reduction and analysis pipeline. The risks associated with failing to do this are of deploying an experiment that exhibits excess statistical uncertainty, requiring longer observations, or excess systematic uncertainty, requiring new mitigation algorithms and possibly limiting the achievable science, and of deploying an analysis pipeline that is unable to meet the challenges of the real data, requiring new pipeline development and deployment during operations. Reducing these risks requires the early development of detailed and realistic parametrized models of the instrument and observation scaled to hundreds of thousands of detectors over tens of telescopes, an accurate simulation and reduction pipeline running on state-of-the-art HPC and HTC systems with sufficient computational efficiency to scale to CMB-S4 data volumes, and analysis pipelines able to derive the core science from the reduced simulated data to unprecedented precision.

C THE CONCEPT DEFINITION TASKFORCE

C.1 Formation and Purpose

The US Department of Energy (DOE) Office of High Energy Physics (HEP) and the National Science Foundation (NSF) Divisions of Astronomical Sciences (AST), Physics (PHY) and Polar Programs (PLR) requested that the Astronomy and Astrophysics Advisory Committee (AAAC) establish the Cosmic Microwave Background Stage 4 Concept Definition Task force (CMB-S4 CDT) in November 2016 as a subcommittee of the AAAC to develop a concept for a CMB-S4 experiment. CMB scientists from universities, DOE laboratories and NASA centers were chosen from across the CMB field in the US to ensure the CDT membership included broad experience in CMB experiments, leadership roles in astronomy and physics experiments from small to large, using ground-, balloon- and space-based platforms. The charge to the CDT is given in Appendix D.

C.2 Process

The CDT held weekly telecons, with agendas circulated beforehand and notes recorded on a private CDT wiki. The CDT wiki was also used as a common logbook for the CDT. The CDT also held six two-day face-to-face meetings. The timing and venue of three of them were tied to workshops of the entire CMB-S4 collaboration. The CDT made use of the substantial amount of work that been done by the CMB-S4 collaboration to explore the science capability of a ground-based CMB experiment with a scale an order of magnitude greater than experiments now being built, as summarized in the CMB-S4 Science Book. The task of the CDT, as reflected in its charge, was to develop a concept design driven by science requirements flowed down to measurement and technical requirements. The level of interaction between the CDT and the ever-growing CMB-S4 collaboration, which is open to anyone, remained high throughout the process. Volunteers were plentiful, requests for help were never refused, and the level of interest and enthusiasm was high. The level of unity and cooperation in the community is both remarkable and gratifying, and its contribution to this report is large.

Simulations are used to connect the science to the measurements and to the hardware that makes the measurements; they must capture the essential features of both the Universe to be observed and the experiment that will observe them. A simulations team led by Lloyd Knox, comprising members of both the CDT and the broader CMB-S4 collaboration, and supplemented by several others brought in for particular expertise, had their own weekly telecons. The main areas of concentration, as described in Appendix A, were on improving the complexity and fidelity in the simulations of foregrounds and instrument systematics.

The legacy science that can be done with a survey of nearly half the sky from 20 to 270 GHz has many facets. Identifying the most important, and in particular determining whether any of them should impose additional design requirements on CMB-S4, was a major effort. One full day of discussion and work at each of two CMB-S4 collaboration workshops, with roughly 150 scientists participating, were devoted to this task. After the first of these workshops sessions, a team of 10–15 volunteers from the CMB-S4 collaboration, led by Jim Bartlett and Steve Allen, continued the work, reporting to the CDT at its weekly telecons.

CMB-S4 can be developed based on existing technology (§ 9; also see the CMB-S4 Technology Book). Rather than an "R&D" plan, the CDT therefore provides a "D" plan. The goal is to reduce the risk in the cost, schedule, and performance associated with the project. This plan was developed by a group of 10 people, half from the CDT and half from the CMB-S4 collaboration not on the CDT who were leading authors and assemblers of the Technology Book, and was aided by an expert from one of the DOE labs in the assessment and management of risk in major projects. This activity was concentrated over the last months, starting once the strawperson design was stable.

Development of a cost estimate with accuracy appropriate for a concept design was another area of intense activity for the CDT. The CDT itself, comprising the leadership of essentially all of the CMB experiments now taking data or under construction, has the best existing knowledge of CMB-experiment hardware, data management, and data analysis costs, and this knowledge was fully incorporated in the cost model described in § 11. For project costs (e.g., management, systems engineering), CDT members with extensive experience with ground-based experiments worked closely with a team of experienced engineers and cost modelers from the four DOE Labs. The members of this team were Karen Byrum, Jennifer Fortner, and Steve Kuhlmann (Argonne), Marc Kaducak and Mike Lindgren (Fermilab), Kelly Hanzel (LBNL), Nadine Kurita (SLAC), and Stew Harris (Berkeley SSL). They were coordinated by Jim Yeck, who has managed multiple large projects for both the NSF and the DOE. On 12 September 2017 we held a review of our cost methodology with a panel of five senior DOE Lab project managers, including Chris Bebek (LBNL, retired), Gil Gilchriese (LBNL, chair), Jim Kerby (Argonne), Paul O'Connor (Brookhaven), and Ron Ray (Fermilab). Recommendations from that panel, most importantly about appropriate levels of contingency in various stages of a project, were incorporated in the costing.

Comments and advice were solicited from several senior members of the field at several points in the process. Amber Miller (USC), Lyman Page (Princeton), and John Ruhl (Case Western Reserve) provided helpful comments, and Lloyd Knox, in addition to leading the simulations team, was invaluable.

D CHARGE TO THE CDT

Professor Buell Jannuzi
Department of Astronomy/Steward Observatory
933 North Cherry Avenue, Rm. N204
Tucson, AZ 85721-0065

Dear Professor Jannuzi:

The US Department of Energy (DOE), Office of High Energy Physics (HEP), and the National Science Foundation (NSF) Divisions of Astronomical Sciences (AST), Physics (PHY), and Polar Programs (PLR) request that the Astronomy and Astrophysics Advisory Committee (AAAC) establish a Cosmic Microwave Background Stage 4 Concept Definition Task force (CMB-S4 CDT) as a subcommittee in order to develop a concept for a CMB-S4 experiment.

Background

Ground-based CMB generally falls under the purview of NSF and DOE, while NASA supports CMB projects within its long-duration balloon program and space missions. In its 2016 report, the interagency AAAC recommended the following: "We encourage DOE, NSF, and the university community to continue working toward a plan for a future (Stage 4) ground based CMB experiment." Following that recommendation and other community input, NSF and DOE are requesting that the AAAC establish this CDT. The creation of the CDT also is in response to the favorable comments or recommendations on CMB science that have been made by community advisory groups over the past decade:

- 2010, Astronomy and Astrophysics Decadal Survey (charged by NASA Astrophysics, NSF-AST, and DOE-HEP): CMB projects are among the "projects thought compelling for a [competed] Mid-Scale Innovations Program" in NSF and "The committee recommends… continuing steps consistent with the DOE mission"
- 2014, Particle Physics Project Prioritization Panel (P5) of the High Energy Physics Advisory Panel (HEPAP) (charged by DOE-HEP and NSF Directorate for Mathematical and Physical Sciences, which includes NSF-PHY and NSF-AST): "Support CMB experiments as part of the core particle physics program. The multidisciplinary nature of the science warrants continued multi-agency support"
- 2015, National Academies report on "A Strategic Vision for NSF Investments in Antarctic and Southern Ocean Research" (charged by NSF-PLR): Identified continuation of studies of the Cosmic Microwave Background as one out of three strategic priorities

These community reports identify CMB as an important scientific priority for consideration by DOE-HEP, NSF-AST, NSF-PHY, and NSF-PLR, hence providing additional rationale for the CDT activity.

The Snowmass 2013 process (http://science.energy.gov/hep/research/snowmass-p5-process/) brought together the U.S.-based CMB community, including many of the current experimental teams, and began to define a coordinated next generation experiment which was termed CMB-S4. The CMB-S4 community-based collaboration has held semiannual meetings and produced a substantial CMB Science Book (see arXiv:1610.02743 [pdf, other]) justifying the CMB-S4 science case. DOE-HEP national laboratories and some university groups are already actively engaged in technology

development for CMB-S4.

Purpose and Charge

The CMB-S4 CDT is asked to develop a concept for implementing a ground-based CMB-S4 experiment. The CDT will take as input the community CMB-S4 Science Book and any further community information as appropriate, and will consider the global landscape of CMB experiments (including ground, balloons, and space).

Specifically, the CDT is asked to deliver:

- The Science Requirements and their rationale
- Measurement and Technical Requirements derived from the Science Requirements
- Project Strawman Concept
- Options and Alternatives (prioritized to the extent possible) for:
    - Concept design (e.g. sites, telescopes, detectors)
    - Concept staging and schedule
    - Collaboration and Data models and interfaces
- R&D development needed, with priorities, to demonstrate technical readiness
- Cost ranges for the strawman concept, including explanations for how they were developed.

The CDT should provide a report on the Science and Measurement Requirements to the AAAC by June 2017 and a final report to AAAC by October 2017 for consideration. In accordance with Federal Advisory Committee Act (FACA) rules, the reports will be discussed and approved by the AAAC before formal transmittal to the agencies.

We appreciate your effort in establishing this subcommittee. Its deliberations and recommendations will inform the agencies on a concept for the CMB-S4 and contribute to the agencies' planning activities. The formation of the CDT does not imply any commitment by the agencies to specific funding or project status for CMB-S4.

We look forward to working with you in this important endeavor. The point of contact for each of the agency participants is listed below.

Sincerely,

Denise Caldwell Division Director Division of Physics NSF	Eric Saltzman Section Head, Ant. Sciences, Div. of Polar Programs, NSF	James Siegrist Associate Director of Science for High Energy Physics, DOE	James Ulvestad Division Director Div. of Astronomical Sciences, NSF
17 November 2016 Date	11/21/2016 Date	11/18/16 Date	11/21/2016 Date

2

Agency Points of Contact

For NSF: Dr. Richard Barvainis
Program Officer
Division of Astronomical Sciences
Directorate for Mathematical and Physical Sciences
Phone: 703-292-4891
Email: rbarvai@nsf.gov

For DOE: Dr. Kathleen Turner
Program Manager
Office of High Energy Physics
Office of Science
Phone: 301-903-1759
Email: Kathy.Turner@science.doe.gov

THE UNIVERSITY OF ARIZONA
COLLEGE OF SCIENCE
Astronomy
& Steward Observatory

Director's Office
Steward Observatory
URL: www.as.arizona.edu

933 North Cherry Avenue
P.O. Box 210065
Tucson, AZ 85721-0065
Telephone: (520) 621-6524
buelljannuzi@email.arizona.edu

November 24, 2016

Dr. Charles R. Lawrence
Chief Scientist for Astronomy and Physics
Jet Propulsion Laboratory
M/S 169-327
4800 Oak Grove Drive
Pasadena, CA 91109

Dear Charles Lawrence:

I am pleased to appointment you chair of the Cosmic Microwave Background Stage 4 Concept Definition Task force (CMB-S4 CDT), a subcommittee of the Astronomy and Astrophysics Advisory Committee (AAAC). The charge for CMB-S4 CDT is detailed in the attached letter I received from representatives of the DOE and the NSF. Please note that you and your task force are asked to provide a report to the AAAC on the Science and Measurement Requirements by June 2017 and a final report by October 2017.

I thank you for your willingness to take on the leadership of the task force and I stand by to assist you in any way I can.

Sincerely,

Dr. Buell T. Jannuzi
Chair, Astronomy and Astrophysics Advisory Committee

Cc: Dr. Denise Caldwell, NSF
Dr. Eric Saltzman, NSF
Dr. James Siegrist, DOE
Dr. Jim Ulvestad, NSF
Dr. Kathy Turner, DOE
Dr. Richard Barvainis, NSF