CMB S4

Update on AoA Report

John Carlstrom & Colin Bischoff

May 3, 2024

Outline

This talk largely overlaps the AoA Analysis update presented at the March 25 Collaboration meeting

- What is the Analysis of Alternatives Study?
- Overview of the AoA Study and the work remaining
- Non-*r* forecasts for the alternatives
- Foreground models
- *r*-forecasts method, initial results, and next steps
    - Deaggregation of factors leading to achieved performance
    - Delensing validation
    - Map validation of Fisher-based r-forecasts
- Discussion

Reminder about the Analysis of Alternatives (AoA)

A DOE requirement for CD-1 is a reviewed AoA document for which the alternatives to be analyzed are agreed to in advance with the DOE. It covers the science reach, construction and lifecycle costs, schedule, and risks (both science and technical). See DOE G 413.3-22 AoA guide.

This talk is focused on a document that covers the scientific reach of the alternatives that will inform the DOE AoA document. It is entitled the ***"CMB-S4 Study to Support the Alternative Analysis and Selection."*** We refer to it as the ***AoA Study***.

We will also need the AoA Study for justification of our alternative to NSF at CDR, and we will likely get new additional CDR guidelines beforehand.

This means we should be prepared to complete the AoA Study in June.

Reminder of the 2022 Analysis of Alternatives Exercise

As you know, in Feb 2022 DOE and NSF requested that the CMB-S4 Project perform an analysis of alternatives (AoA) to assess options for configuration of the project that place reduced demands on infrastructure and logistics at the South Pole compared to those required by the design at that time and that can achieve all of the science goals.

Following an initial set of studies, three alternatives were selected for detailed analysis.

Alternative	South Pole Scope	Chile (Atacama) Scope
Alternative 1: SATs & LAT at South Pole	3 Small-Aperture Telescopes 1 Large-Aperture Telescope	2 Large-Aperture Telescopes
Alternative 2: Only SATs at South Pole	4 Small-Aperture Telescopes	3 to 5 Large-Aperture Telescopes
Alternative 3: Nothing at South Pole	None	≥9 Small-Aperture Telescopes ≥3 Large-Aperture Telescopes

Reminder of the 2022 Analysis of Alternatives Exercise

A great deal of the Project's and Collaboration's effort in 2022 was devoted to working on the AoA.

A tremendous amount of work was done by many of you (thank you!) and documented on confluence. It was a huge effort.

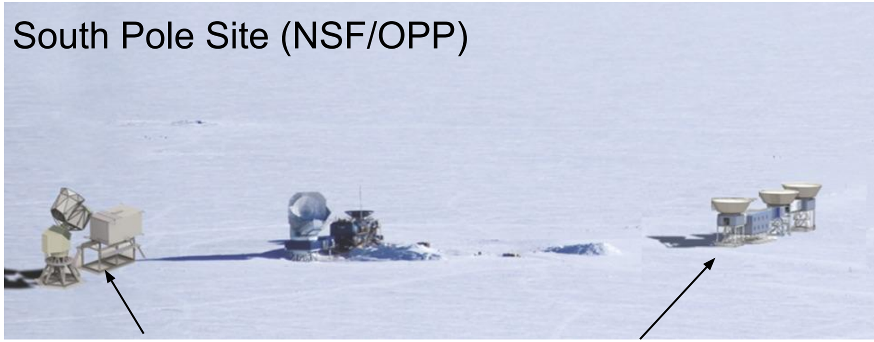
The process and results were vetted by the Collaboration on Oct 12 & 14, reviewed by an external panel of experts on Nov 4-5, and presented to the agencies on Dec 7, 2022.

The process resulted in the new conceptual design for CMB-S4.

The Conclusion of the 2022 AoA

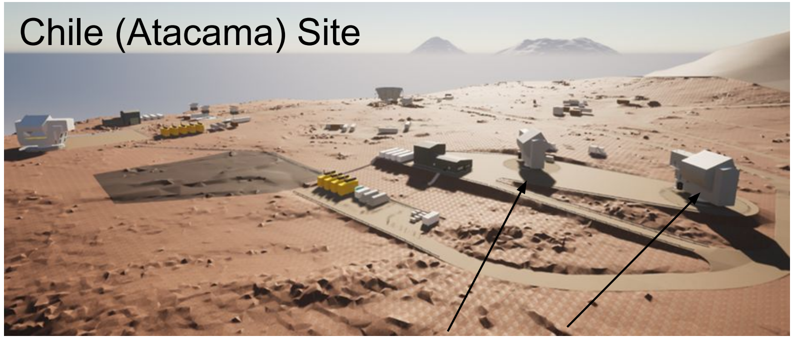
Alternative 1 was determined to be the best configuration because:

- The South Pole offers the best conditions for the ultra-deep survey focused on inflation science.
- The combination of small- and large-aperture telescopes observing the same patch of the sky provides unique checks on systematic errors.
- The Atacama site in Chile provides excellent conditions for the deep, wide-field survey with 2 large-aperture telescopes that addresses N_{eff} and many other science goals.



One Large Aperture (5 m) Telescope

3 Small Aperture Telescopes (9 0.56-m aperture optics tubes)



Two Large Aperture (6 m) Telescopes

CMB S4

So are we done with the AoA?

No, unfortunately.

We still need to write the DOE required AoA document and the supporting science study, i.e., the AoA Study.

In addition, we also need to document **Alternative 0, “status quo”** which is to do nothing new. We interpret this as continued operation of the South Pole Observatory and the upcoming Simons Observatory.

While a lot of the AoA Study effort has been completed over the last year and most of it is written up, there is still important work yet to be addressed.

Status the AoA Study

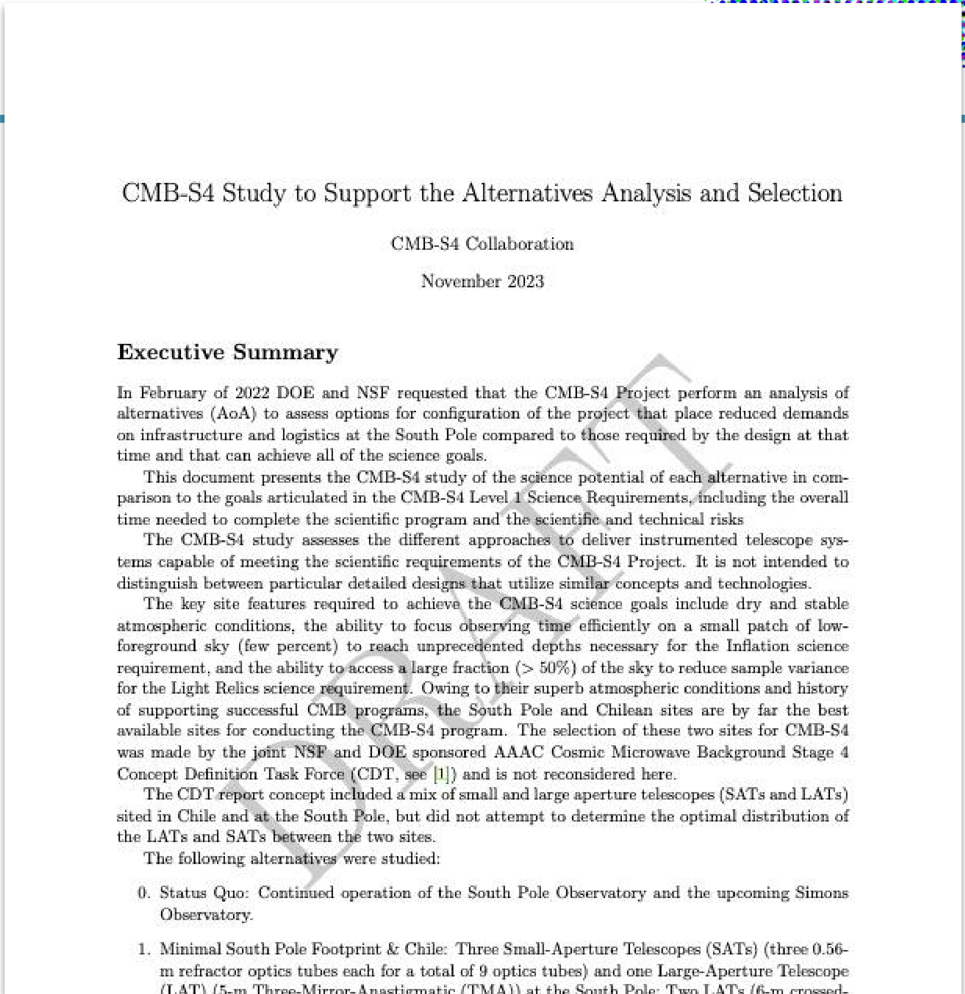
Builds on 2022 AoA effort which was reviewed and presented to NSF/DOE

It is the "write up" of that work which will support the required AoA document, which will be reviewed along with the AoA Study.

A draft AoA Study was reviewed at the November 2023 Directors Review

It needs to include that feedback and also anticipate additional requests from NSF/DOE

It is on the CMB-S4 Github repository and accessible via Overleaf.



CMB-S4 Study to Support the Alternatives Analysis and Selection

CMB-S4 Collaboration

November 2023

Executive Summary

In February of 2022 DOE and NSF requested that the CMB-S4 Project perform an analysis of alternatives (AoA) to assess options for configuration of the project that place reduced demands on infrastructure and logistics at the South Pole compared to those required by the design at that time and that can achieve all of the science goals.

This document presents the CMB-S4 study of the science potential of each alternative in comparison to the goals articulated in the CMB-S4 Level 1 Science Requirements, including the overall time needed to complete the scientific program and the scientific and technical risks

The CMB-S4 study assesses the different approaches to deliver instrumented telescope systems capable of meeting the scientific requirements of the CMB-S4 Project. It is not intended to distinguish between particular detailed designs that utilize similar concepts and technologies.

The key site features required to achieve the CMB-S4 science goals include dry and stable atmospheric conditions, the ability to focus observing time efficiently on a small patch of low-foreground sky (few percent) to reach unprecedented depths necessary for the Inflation science requirement, and the ability to access a large fraction (> 50%) of the sky to reduce sample variance for the Light Relics science requirement. Owing to their superb atmospheric conditions and history of supporting successful CMB programs, the South Pole and Chilean sites are by far the best available sites for conducting the CMB-S4 program. The selection of these two sites for CMB-S4 was made by the joint NSF and DOE sponsored AAAC Cosmic Microwave Background Stage 4 Concept Definition Task Force (CDT, see [1]) and is not reconsidered here.

The CDT report concept included a mix of small and large aperture telescopes (SATs and LATs) sited in Chile and at the South Pole, but did not attempt to determine the optimal distribution of the LATs and SATs between the two sites.

The following alternatives were studied:

0. Status Quo: Continued operation of the South Pole Observatory and the upcoming Simons Observatory.
1. Minimal South Pole Footprint & Chile: Three Small-Aperture Telescopes (SATs) (three 0.56-m refractor optics tubes each for a total of 9 optics tubes) and one Large-Aperture Telescope (LAT) (5-m Three-Mirror-Anastigmatic (TMA)) at the South Pole; Two LATs (6-m crossed-

Contents

iii

iv

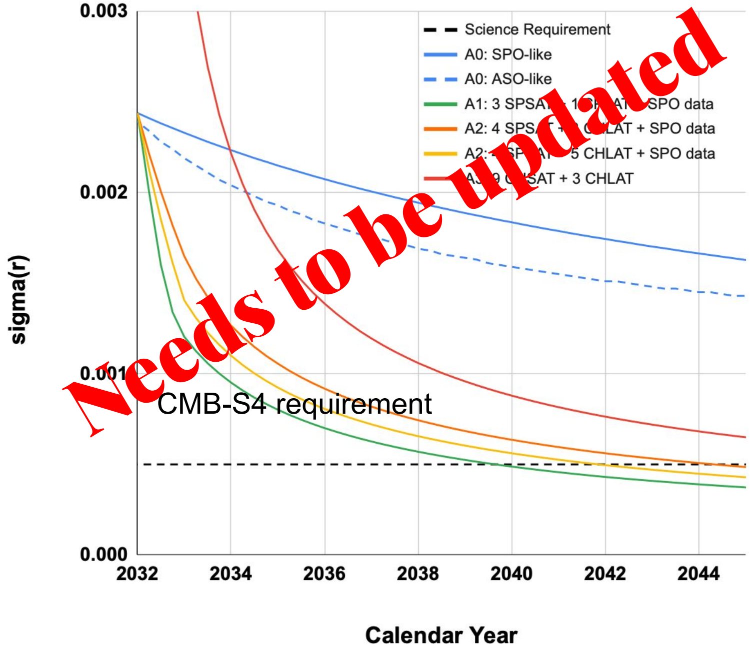
What's left to do for the AoA Study?

- The **non-*r*** AoA analyses are essentially done.
- The AoA Study currently provides a solid ***comparative analysis*** of the *r* forecasts of the various alternatives, which are based on scaling from achieved performance from the South Pole, which are in turn 'transferred' to Chile.
- We have considered three foreground models of varying complexity for the forecasts; for the final AoA Study results, we should probably assume only one benchmark model for comparison.
- We should probably report the results of only one benchmark forecasting methodology if we understand why it works best. We would like to show map-based validation of the Fisher-based forecasts.

What’s left to do for the AoA Study?

- Based on feedback and anticipated requests from NSF for CDR, we need to provide simple comparative plots of $\sigma(r)$ v time for the alternatives.
- To aid in comparing our results with those for other experiments, we need to provide the breakdown of the major factors that lead from idealized *r* forecasts to those using achieved performance.

Inflation Science With Time



Reminder of our Science Requirements

SR 1.0: CMB-S4 shall test models of inflation by putting an upper limit on r of $r \lesssim 0.001$ at 95% confidence if $r = 0$, or by measuring r at a 5σ level if $r > 0.003$. [$\sigma(r) \lesssim 5 \times 10^{-4}$]

SR 2.0: CMB-S4 shall determine N_{eff} with an uncertainty ≤ 0.06 at the 95% confidence level.

SR 3.1: CMB-S4 shall detect at $\geq 5\sigma$ all galaxy clusters at $z \geq 1.5$ with an integrated Compton $Y_{SZ,500} \geq 2.4 \times 10^{-5}$ arcmin^2 over at least 50% of the sky.

SR 3.2: CMB-S4 shall detect at $\geq 5\sigma$ all galaxy clusters at $z \geq 1.5$ with an integrated Compton $Y_{SZ,500} \geq 1.2 \times 10^{-5}$ arcmin^2 over at least 3% of the sky.

SR 4.1: CMB-S4 shall detect GRB afterglows brighter than 30 mJy at 90 and 150 GHz over at least 50% of the sky and enable followup by issuing timely alerts to the community.

SR 4.2: CMB-S4 shall detect GRB afterglows brighter than 9 mJy at 90 and 150 GHz over at least 3% of the sky and enable followup by issuing timely alerts to the community.

General Considerations for Non-r Science

Alt 0: Status Quo: Higher noise levels - would require ~35 years of observing with advanced SO.

Alt 1. 3 SATs and 1 LAT at South Pole & 2 LATs in Chile: Same high-ell as pre-AoA design, which was constructed to meet Science Requirements

Alt 2. Only SATs at the South Pole and increased LATs in Chile: Wide survey unchanged compared to Alternative 1, Delensing survey wider (f_{sky} ↑) and shallower (N_{ell} ↑) with smaller beam (θ_b ↓)

Alt 3. All Telescopes in Chile: Very similar to Alternative 2 for non-r science, since SATs do not contribute to other Science Requirements

AoA Study Results for Non-r Science

Science Case	Requirement	Alt 1	Alt 2a	Alt 2b	Alt 3
Light Relics	MR 2.0	✓	✓	✓	✓
	SR 2.0	✓	✓	✓	✓
Clusters	MR 3.1	✓	✓	✓	✓
	SR 3.1	✓	✓	✓	✓
	MR 3.2	✓	✗	✓	✗
	SR 3.2	✓	✓†	✓	✗‡
Transients	MR 4.1	✓	✓	✓	✓
	SR 4.1	✓	✓	✓	✓
	MR 4.2	✓	✓	✓	✓
	SR 4.2	✓	✓	✓	✓

SR 2 Light Relics: Benefits from wider sky coverage at fixed effort

SR 3 Clusters: Benefits from smaller beam, negatively impacted by higher noise

SR 4 Transients: Benefits from smaller beam, negatively impacted by higher noise

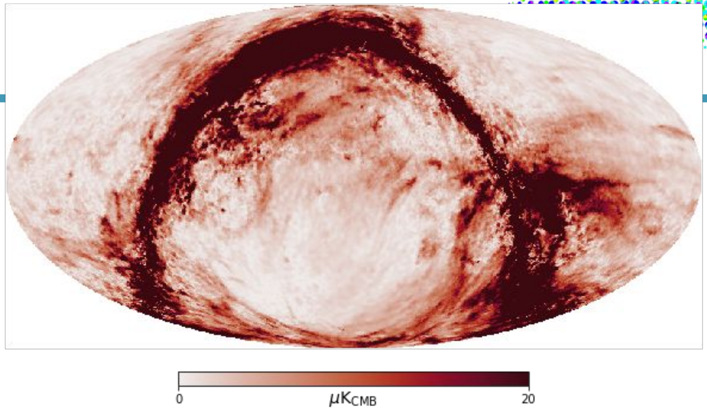
†: Despite failing Measurement Requirement 3.2 due to the higher noise in the delensing patch, the smaller beam of Alternative 2 compared to Alternative 1 allows Science Requirement 3.2 to still be met in this case.

‡: Alternative 3 fails both Measurement Requirement 3.2 and Science Requirement 3.2 due to the higher noise in the delensing survey; in this case the integrated Compton $Y_{\mathrm{SZ}_{500c}}$ will be slightly larger than that of Science Requirement 3.2, though the wider sky coverage implies that a greater number of clusters will be detected.

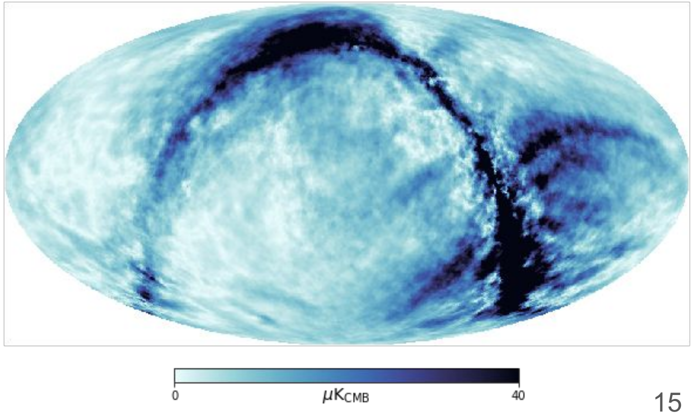
Foreground models

- New suite of PySM models developed by PanEx Group
  - Improved emission templates based on latest component separation analyses
  - Small-scale fluctuations in amplitudes as well as spectral parameters
  - “Layer model” (MKD) with line-of-sight frequency decorrelation (2018 MNRAS, 476, 1310)
- Three sky models, all consistent with current data:
  - **Medium Complexity:** Parameter maps based on component separation with extrapolation to small scales in both amplitudes and spectral parameters
  - **Low Complexity:** Small-scale fluctuations in amplitudes only, no decorrelation
  - **High Complexity:** Near maximum-allowed decorrelation for dust emission, line-of-sight dust SED variations, AME polarization, synchrotron curvature

353 GHz P



30 GHz P



r-forecasts method, initial results, and next steps

Forecasting σ(r) is challenging, and comparison between the alternatives intersects with many of these challenges:

- **Foregrounds**: The alternatives have different sky coverage, so we need to understand how foreground complexity varies between the observing patches.
- **Delensing**: Each of the three alternatives has a different configuration of LAT(s) to achieve the necessary level of delensing.
- **Systematics**: Impact of instrumental systematics is hard to evaluate. BICEP has proven performance for a particular site and instrument design. Variations like HWP for SATs in Chile could help in some ways but hurt in others.

To compare between alternatives, we tried to simulate the differences that we understand: detector counts/integration time, geometric sky coverage, foreground models, observing efficiency, atmospheric effects on NET

Low-ell noise levels were normalized based on scaling from achieved performance, so effects that matter are ones that differ between alternatives.

We have not attempted to simulate differences in systematics.

r-forecasts method, initial results, and next steps

Step 1: Forecasting yields estimated noise levels vs time for SATs and LATs in the three alternative configurations.

Step 2: LAT noise is used to estimate residual A_{lens} after delensing [R. Flauger]

Step 3: Low-ell SAT/LAT noise, delensing residual, and foreground model(s) are used to forecast σ(r) vs time.

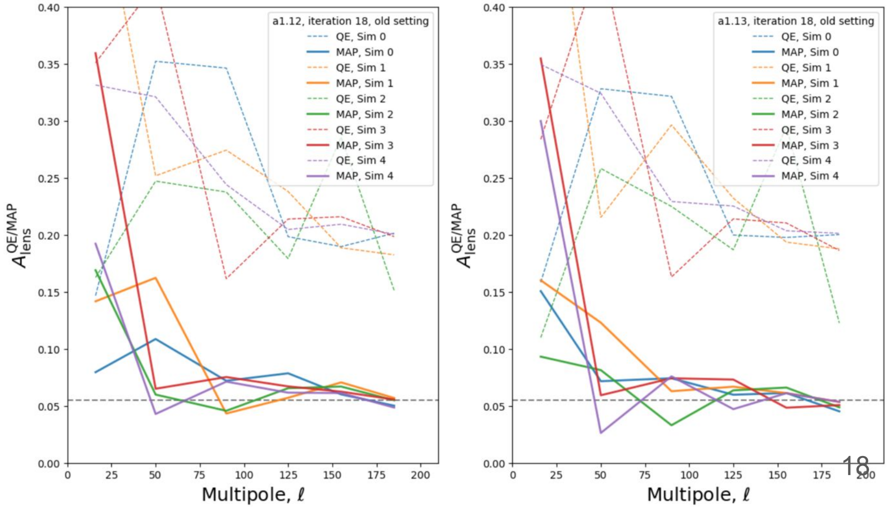
- Parametric likelihood analysis of auto/cross spectra (BICEP style analysis) [C. Bischoff]
- Harmonic-space internal linear combination [R. Flauger]

Both forecasts agree that alternative 1 reaches the science target fastest.

We are working to understand disagreements between the two forecasts, especially the impact of varying foreground complexity. This process involves *validating* the forecasts by carrying out analysis of map-based sims to test whether the delensing and foreground separation perform as predicted. Analysis of sims can reveal biases on r that Fisher forecasts gloss over.

Validation of delensing for *r* forecasts

- Initial AoA results used forecasts for the efficacy of delensing. We have now tested these forecasts by analyzing simulated SPLAT maps (alt.1) including foregrounds, anisotropic noise, etc.
- Analysis involves map-based foreground cleaning [S. Ghosh] followed by iterative internal delensing pipeline that works on curved sky Healpix maps [S. Belkner, J. Carron]
    - Software described in recent CMB-S4 delensing paper: https://arxiv.org/abs/2310.06729
- So far this validation has only been carried out on a small number of realizations (5) for alt. 1, due to computational complexity. However, results are in *excellent agreement* with forecasts for all foreground models and we expect similar results for the alt.2,3 configurations.
- **At Right**: Residual A_{lens} for five realizations with medium complexity (left) and high complexity (right) foreground models. [S. Belkner]



CMB S4

Map validation of Fisher-based r forecasts

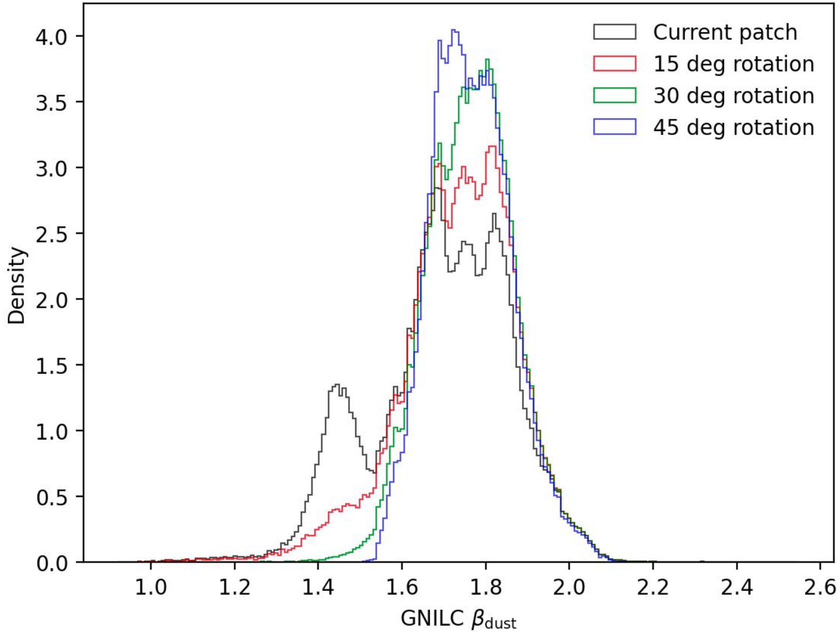
- Demonstrating the compatibility of forecasts with (semi-)realistic map-based simulations is required to strengthen the reliability of the conclusions of the AoA.
- In particular, forecasts predict errors, but not biases.
- Recent work on sky modeling recommends three foreground model options. Differences between those three models are representative of our lack of knowledge about the real-sky foreground emission.
- For two of these models (“medium” and “high” foreground complexity), map-based methods that were successful in previous work yield model-dependent, method-dependent biases that range from no detectable bias, to biases of up to r~0.002, which we do not fully understand yet.

Map validation of Fisher-based r-forecasts

- The origin of the biases, and ways to mitigate them, are being investigated. Several recent posts in the CMB-S4 logbook address various aspects of this issue and new work is actively discussed on weekly Low-ell BB / AoA study telecons (Mondays, 9:30 Pacific).

- Ongoing work includes:
  - Identifying the origin of the biases and the reason for the difference with earlier results
  - Developing alternative foreground cleaning and mitigation tools, and testing their performance
  - Validation / Consolidation of the foreground models in patches of specific interest

- Recent progress includes:
  - Exploration of how biases changes with observing patch selection. [Figure right from S. Ghosh]
  - Studying the difference between dust spectral index maps derived from Planck Commander vs GNILC data products.



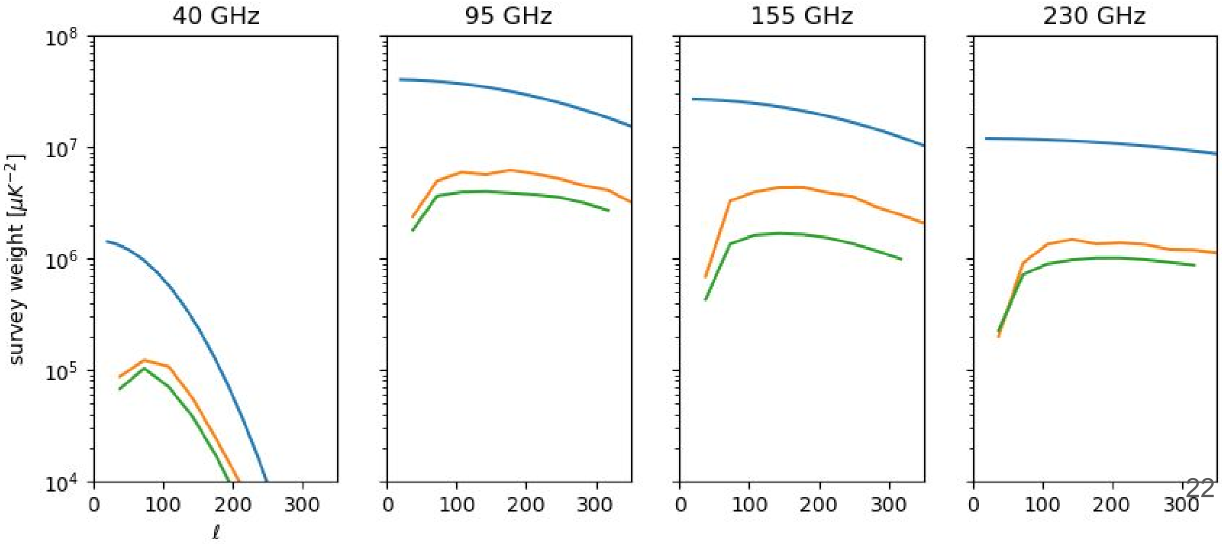
Deaggregation of factors leading to achieved performance

- We have tried to understand, and incorporate into forecasts, factors that differ between the alternatives.
- There is still a need to better understand the differences in performance between *any* CMB-S4 configuration (like Alternative 1) and past/current experiments (BICEP/Keck) that are used for performance-based forecasting.
- We are currently working to deaggregate the factors that lead to achieved BB spectrum sensitivity.
  - Some factors, like detector yield and uniformity, are intended to be better for CMB-S4 than for past experiments.
  - We have little control over other factors, like weather cuts, so should assume that CMB-S4 will suffer similar impact as past experiments.
  - It is important that our model of observing efficiency is still grounded in achieved performance / can explain achieved BB spectrum error bars.

Deaggregation of factors leading to achieved performance

- Figure below shows the difference in survey weight (increases linearly with integration time) between alt. 1 (7-year) forecasts:
  - **Hyper-idealized forecast** (blue): All detectors operating at design sensitivity 24/7/365.
  - **AoA sims** (orange): Serious attempt at *ab initio* sensitivity simulation including efficiency factors based on current experimental efforts [S. Simon, R. Keskitalo]
  - **Scaled performance** (green): Scaling from BICEP/Keck achieved performance with limited modifications (NET impact of bandpass differences, 100 mK bath temp)

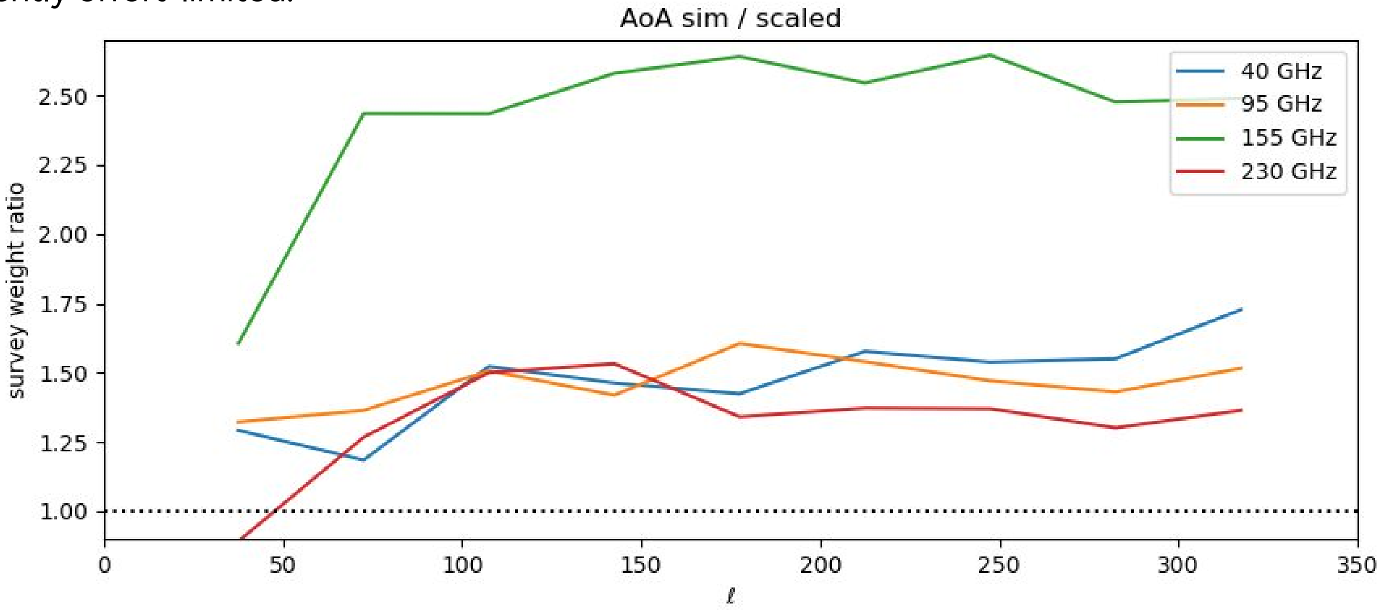
- The AoA sims include specific assumptions about detector yield, observing time on sky, data cuts and other weather impact, etc. By comparing these assumptions against BICEP historical record we will learn about the source of the discrepancy and understand where we might improve with CMB-S4.



Deaggregation of factors leading to achieved performance

- Figure below shows the survey weight ratio between the AoA sims and scaled performance forecasts.
    - Note that, in the limit of optimal distribution of effort, $\sigma(r)$ scales as 1 / survey weight so these forecasts show ≳50% differences.

- Lots of existing data to study year-to-year and wafer-to-wafer variation in performance, but we are currently effort-limited.



Discussion

To help with the AoA Study:

- Please consider participating in the joint low-ell BB analysis working group and AoA Study group meetings held Mondays at 9:30am PT / 11:30 am CT / 12:30pm ET (see the CMB-S4 calendar)
- Read and comment on the AoA Study document