CASPEN Report: CCA Visit

Alexandre Adler & Nadia Dachlythra

April 2024

Context

The Cosmic Microwave Background (CMB) is a rich source of information about the Universe's beginning and its evolution. In recent years, the big prize for CMB experiments has been the detection of tensor mode perturbations, a signature of cosmic inflation. Cosmic inflation describes a period of dramatic expansion in the very early Universe, and involves energy scales far greater than those reachable by particle accelerators. The signature of the tensor modes, primordial B-mode polarization, is however very weak, and can easily be obscured by other polarized microwave sources in the sky, like the emission of our Galaxy. Those foregrounds have two main components whose emission scales in opposite ways with frequency. Synchrotron radiation is very bright at low frequencies but dims gradually. Around 90 GHz, the polarized foreground emission starts being dominated by thermal dust, whose emissivity increases with frequency until 1 THz. CMB experiments build maps of the sky at different frequencies and then try to use the frequency scaling of foregrounds to separate them from the CMB at the level of the angular power spectrum (the two-point correlation function).

Component separation at the very core of CMB data analysis. The end product is either component maps or a series of parameters that describe the amplitude (A_i) and the angular (α_i) and frequency (β_i) scaling of the foregrounds' angular power spectra. In the presence of primordial B-modes, the power spectrum scales in proportion of the tensor-to-scalar ratio r. With two foreground components, this means seven parameters to estimate. Estimating the likelihood for those parameters requires a complex set of assumptions on systematic errors. Likelihood-Free Inference methods could sidestep that problem.

Dr Duivenvoorden and ourselves are experts in simulating CMB experiments. Therefore, for two weeks at the end of January/beginning of February 2024 we visited Dr Duivenvoorden at the CCA to start a project evaluating how Simulation-Based Inference (SBI) can be applied to CMB component separation.

Goals

During our time in CCA, we laid the ground for this project and produced a set of preliminary results. We started by generating sky simulations including a primordial B-mode signal corresponding to different values of the r-tensor, varying lensing amplitude, and foreground model parameters. By the term 'lensing amplitude' we refer to the amount of E-mode signal we expect to see

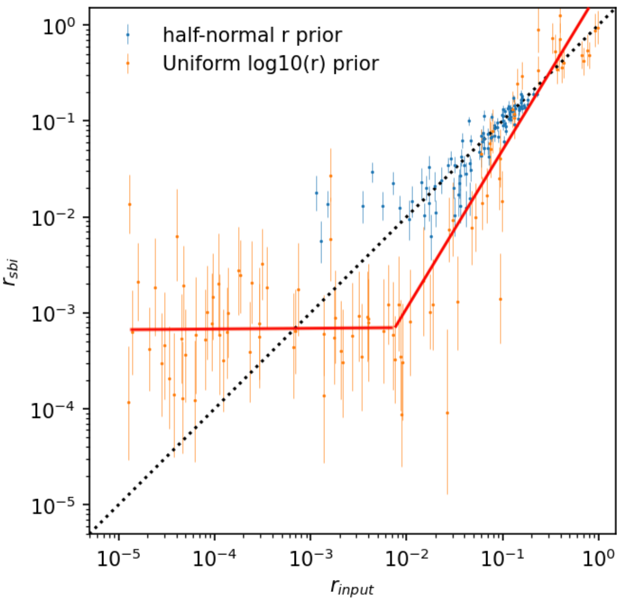

Figure 1: The r value estimated by `SBI` vs the input r for two versions of the model with different priors on r.

mixed with the B-modes due to gravitational lensing of the CMB photons by large-scale structure. The angular power spectra of these sky simulations were then summed with white noise spectra corresponding to the Noise-Equivalent-Power (NEP) values that were calculated for the Simons Observatory (SO) Small Aperture Telescopes (SATs). The values for the r-tensor, lensing and foregrounds were drawn from sampling priors using `SBI`. This is the software employed throughout the project and is used both to inform the priors of the parameters and analyze the angular spectra to find the best-fit values for - in total - eight parameters.

Learning to work with the `SBI` software was, thus, the first goal. This code allows for a likelihood-free estimation of the desired parameters by inferring their full posterior distributions from the simulations. Understandably, the quality of the results depends on the number of simulations provided to the code. This was our second challenge: to develop a pipeline capable of producing simulations both efficiently and fast. Once this short milestone was achieved, we generated the first set of results which were indicative of the algorithm's performance. We then, scaled the complexity of the simulations by truncating the observed region of the sky and, finally, applying an 'observation matrix' to the simulated maps. The observation matrix's purpose is to connect the observed, pixelized version of the sky with the true, continuous, signal from the sky. Our last goal was to re-produce the posterior likelihoods for the tensor-to-scalar ratio and the rest of the desired parameters under these additional challenges.

Results and next steps

With Dr. Duivenvoorden's help, were able to install and use `SBI`. Nadia focused on creating the pipeline that would generate the simulated auto and cross-spectra at all frequencies, including noise splits, with assistance from Dr. Susanna Azzoni (Princeton University). Alexandre explored the different training parameters of `SBI` by training the model on different data and performing tests of performance, like for instance the coverage plot in Figure 1). We managed to converge on a choice of Kernel Density Estimation and learning rate, but struggled with populating our r-prior for small values of r as the allowed values span several orders of magnitude. We did not have the time to explore the effect of the observation matrix in detail.

In the near future, we will work to complexify our simulations further with Dr. Duivenvoorden and Azzoni and other CCA members like Pr. Colin Hill, Kristen Suarro, and Dr. Adrian Bayer. We will also study the observation matrix's effects more rigorously.

Acknowledgements

This work was supported by collaborative visits funded by the Cosmology and Astroparticle Student and Postdoc Exchange Network (CASPEN).