Report on the Inaugural CMB Analysis Summer School

Jeff McMahon, Renée Hložek, Bruce Partridge, Kasey Wagoner, Jesse Treu, the ACT collaboration

Department of Physics, University of Michigan, Dunlap Institute & Department of Astronomy and Astrophysics, University of Toronto, Haverford College, Princeton University

Abstract

The inaugural CMB Analysis Summer School was held by Jeff McMahon and Renée Hložek for the ACT Collaboration at the University of Michigan at the end of August 2016. The workshop was entirely interactive, with workbooks placed on an open-source Github repository for students to download, implement and modify. We present the motivation for the school, the topics covered, the feedback received and suggestions for future schools in this report.

hlozek@dunlap.utoronto.ca

1 Introduction and Motivation

The field of cosmology with the Cosmic Microwave Background continues to grow with the success of ground-based CMB experiments. The future of CMB cosmology is one in which maps of the microwave sky can be cross-correlated with measurements at other wavelengths (including large-scale optical surveys) to constrain a host of cosmological and astrophysical parameters. Cross-correlating involves spatially correlating the CMB temperature (or polarisation) maps with the positions (or fluxes) of other astrophysical objects to learn about how the microwave maps respond to objects along the same line of sight. This is work on the cutting edge of astro-statistics; it needs a large set of skills in a wide variety of areas.

Given the amount of work needed in any particular collaboration and the level of complexity of the instruments, there is often a tendency for students to become specialized in the more narrowly-focused fields defined by their own programs. Consequently, we see a need for training in the techniques of CMB analysis more broadly in the community. The CMB Analysis Summer School was started as a way to train graduate students and postdoctoral researchers (those working within CMB collaborations and those who are interested in performing the multi-wavelength correlations) in these techniques.

As a jumping off point, the inaugural CMB Analysis Summer School was targeted at members of the Atacama Cosmology Telescope, PolarBEAR and South Pole Telescope collaborations, advertised mainly by word of mouth. Prof. Jeff McMahon taught the La Serena Summer School (`http://www.aura-o.aura-astronomy.org/winter_school/content/2015-la-serena-school-data-science`). Based on the success of that program, he and Renée Hložek decided to adapt the lecture notes from the school into a two-day workshop.

2 Methodology

The workshop followed a blended learning style which combined lectures with an interactive framework of exercises. This helped achieve the goals of giving the students an understanding of the methods (and their implementation) of CMB analysis, rather than just the analytical formulation. This format required that all students arrive with laptops that they could use for the analysis, and that they had the requisite software installed in order to run the iPython notebooks. The iPython notebooks were added to a GitHub repository (`https://github.com/jeffmcm1977/CMBAnalysis_SummerSchool`) daily to ensure the students were able to focus on the tasks at hand while working at their own pace.

Before arriving at the summer school, the participants were given a workbook on Fourier analysis and convolutions. This served as a test of the software capabilities and provided an introduction to the structure of the workshop. During this time, they were offered an interactive video help session, hosted by Renée Hložek, to deal with any computational issues before arriving at the summer school. This not only prepared them for the content but also tested that they were able to run the iPython notebooks they would be downloading each day.

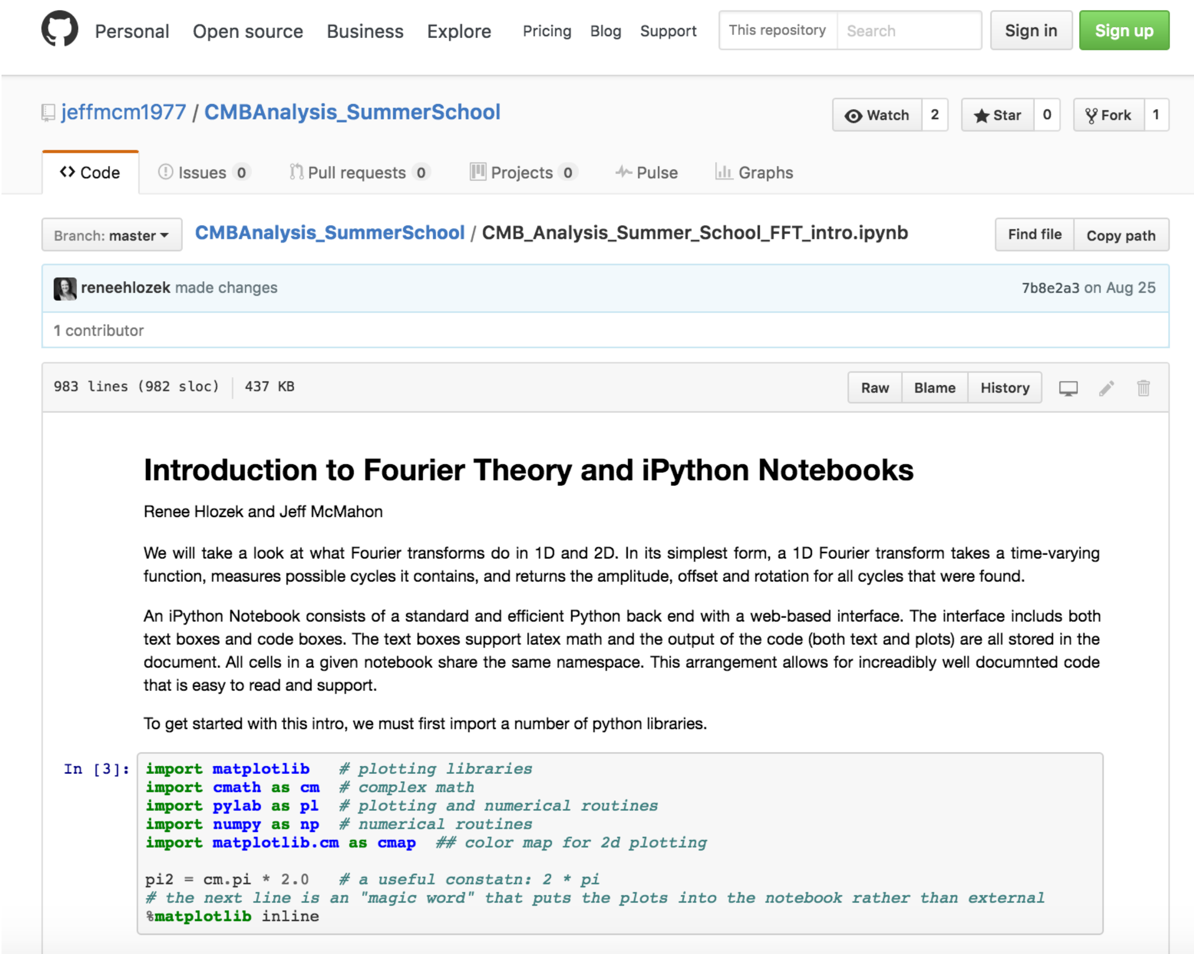

Figure 1: Fourier transformation and convolution practice workbook: this workbook was sent out in the week preceding the summer school as a test of the software required (and to recap Fourier theory that would be used in the workshop itself). It formed the basis of an interactive pre-workshop video call to address any computational issues the participants were experiencing.

The workshop was divided into two parts: an initial set of workbooks designed to teach content with exercises to work through and some extensions to prepare, and a series of project problems which were largely left to the participants to work on in groups.

2.1 Initial Project Exercises

The initial projects covered how to simulate realistic maps based on a cosmological power spectrum and identify assumptions about the properties of the astrophysical foregrounds that would contaminate the maps. Instrumental systematics and how these affect the maps were also discussed.

The initial workbook projects were a combination of introductory theory explanation, exercises to complete and demonstration code to examine (as shown in the code provided in Figure 2). The students were asked to modify the code for a particular situation, or to write additional code to supplement the exercise. Lecturers McMahon and Hložek presented the workbooks in an interactive way, and then guided the students as they attempted one section at a time. Questions by the students were encouraged and answered by the lecturers so that the whole group could hear and benefit. Figure 3 shows an example in which the learners were asked to modify the whole pipeline by changing the instrumental beam specifications, and to comment on the resolution of known experiments on the ground and in space.

Section 1.2: Temperature Anisotropy Map

In this step we generate a simulated map of the CMB sky with the spectrum we read in above. Since the power spectrum is a function of ℓ we need to do much of the work in harmonic space. If we were generating a map on the full sky we would need to work with spherical harmonics. Here we consider a small patch of sky ($\sim 10^{\circ} \times 10^{\circ}$) were we cam is the 'flat-sky' approximation and replace ℓ with $k = \sqrt{k_x^2 + k_y^2}$. There is a linear dependance between these variables defined by $\ell = k * 2\pi$.

In the flat sky approximation we generate a CMB map by: (1) generating a 2 d power spectrum by revolving the above spectrum (properly normalized) about the axis in polar coordinates, (2) generating a gaussian random map with unit variance, (3) multiplying that maps from 1 and 2, and (4) Fourier transforming this to get a real space map. We provide a function to do this and a function to plot this (and other maps) with a uniform color scale.

```
In [ ]: ## variables to set up the size of the map
N = 2**10.  # this is the number of pixels in a linear dimension
            ## since we are using lots of FFTs this should be a factor of 2^N
pix_size  = 0.5 # size of a pixel in arcminutes

## variables to set up the map plots
c_min = -400  # minimum for color bar
c_max = 400   # maximum for color bar
X_width = N*pix_size/60.  # horizontal map width in degrees
Y_width = N*pix_size/60.  # vertical map width in degrees


def make_CMB_T_map(N,pix_size,ell,DlTT):
    "makes a realization of a simulated CMB sky map"

    # convert Dl to Cl
    ClTT = DlTT * 2 * np.pi / (ell*(ell+1.))
    ClTT[0] = 0.
    ClTT[1] = 0.

    # make a 2d coordinate system
    ones = np.ones(N)
    inds  = (np.arange(N)+.5 - N/2.) /(N-1.)
    X = np.outer(ones,inds)
    Y = np.transpose(X)
    R = np.sqrt(X**2. + Y**2.)

    # now make a 2d CMB power spectrum
    ell_scale_factor = 2. * np.pi / (pix_size/60. * np.pi/180.)
```

Figure 2: Screenshot of part of the first workbook. This workbook summarised the statistics of the CMB and how a power spectrum relates to maps of the CMB. In this case the power spectrum was generated online by running the `CAMB` software through NASA's LAMBDA website (`https://lambda.gsfc.nasa.gov/toolbox/tb_camb_form.cfm`).

The topics covered in the initial projects are listed below. These formed the basis for work for the entire first day and the first part of the second morning.

1. **CMB power spectra and simulation:** How to go from a theoretical CMB power spectrum to a 2D map.
2. **Microwave foregrounds and simulation:** How to simulate simple foreground models for a host of astrophysical sources.
3. **Instrumental effects, filtering, beams and noise:** How to incorporate real-world telescope systematics into the maps.
4. **Power spectrum estimation:** How to go from maps to an estimate of the 2D power spectrum, including bias correction and noise simulation.
5. **Matched filter techniques and stacking:** How to combine CMB maps with other probes, and to use map-based techniques for identifying astrophysical objects in the maps.

Once these workbooks were completed, the students had a general understanding of the map properties and a grounding in the types of analyses possible. The larger group projects that were to follow were less guided and allowed the students to focus more deeply on particular areas of research interest.

2.2 *Larger Group Projects*

The larger projects, which were tackled in groups on the second day (after the students finishhed the fifth project on matched filtering and statistical cross-correlations/stacking), covered the application of techniques to real data and the extension from

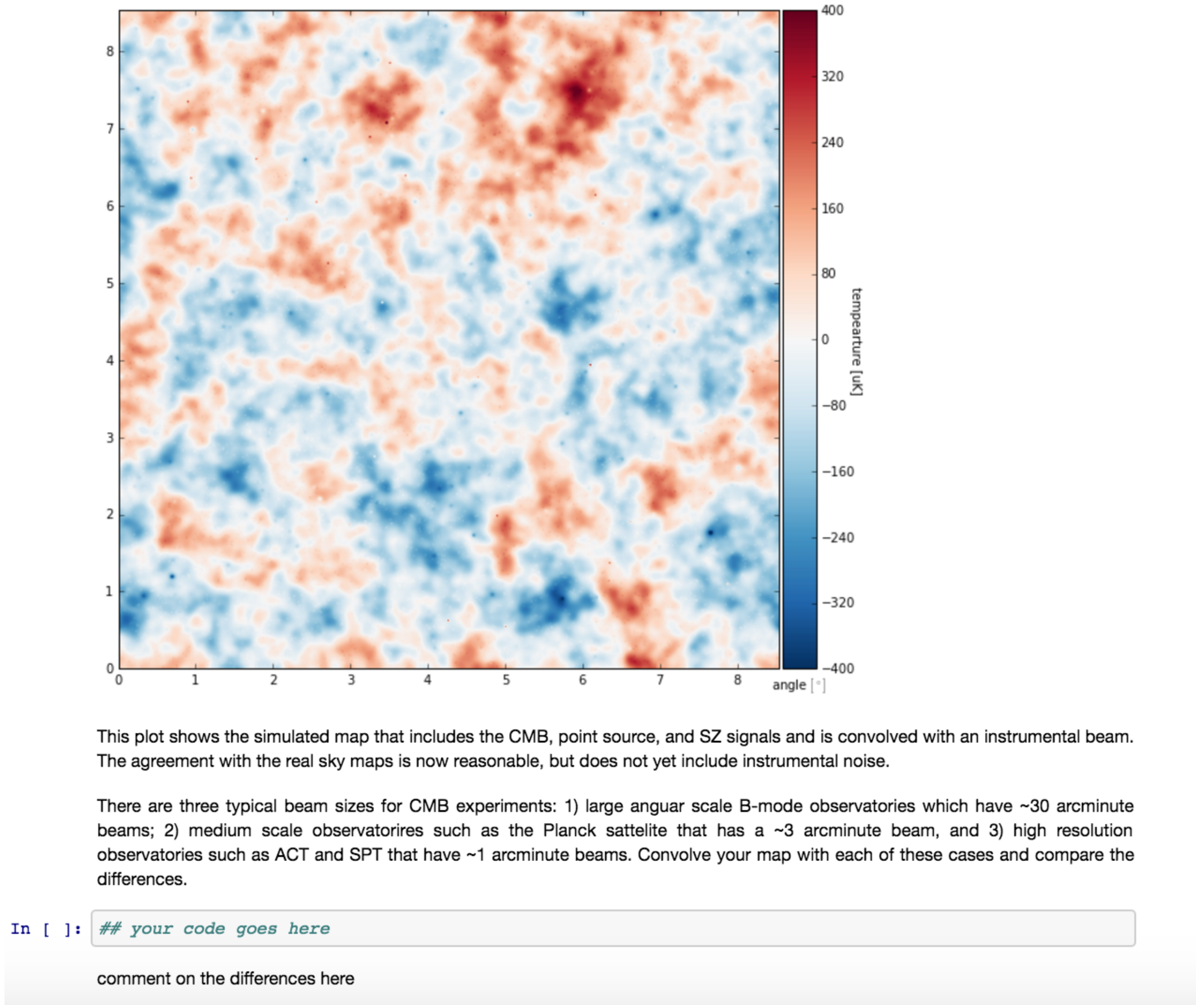

Figure 3: Screenshot of part of the third workbook. The workbook demonstrated how to include instrumental parameters such as the telescope beam size and sensitivity into any map of the sky. After a series of exercises using the code provided, the participants were asked to modify the code for different beam sizes and to comment on how these relate to actual experiments with which they might be familiar.

temperature to polarisation. They also dealt with how to perform analysis on power spectra derived from the data, including, for example, a demonstration of the Markov Chain Monte Carlo algorithm.

The students worked in small groups of 3-5 and could discuss and collaborate on methodology; however the second day of the workshop, everyone was coding and implementing their strategies independently, while contributing to their group's focus on their respective project.

Here are the larger projects:

6. **Analysis of real ACT maps:** In this project, publicly available ACT maps were used and the topics from the first five projects were applied to these maps.

7. **CMB Polarization:** This project contained the basics of how to move from intensity to polarisation simulations.

8. **Parameter estimation and MCMC:** In this project the learners were introduced to Markov Chain Monte Carlo methods and had to compute constraints using `pyCAMB`.

9. **Map-making from TODs:** This project contained code from Sigurd Naess on how to produce maps based on time ordered data (TODs). The learners went through the code and documented it based on their understanding.

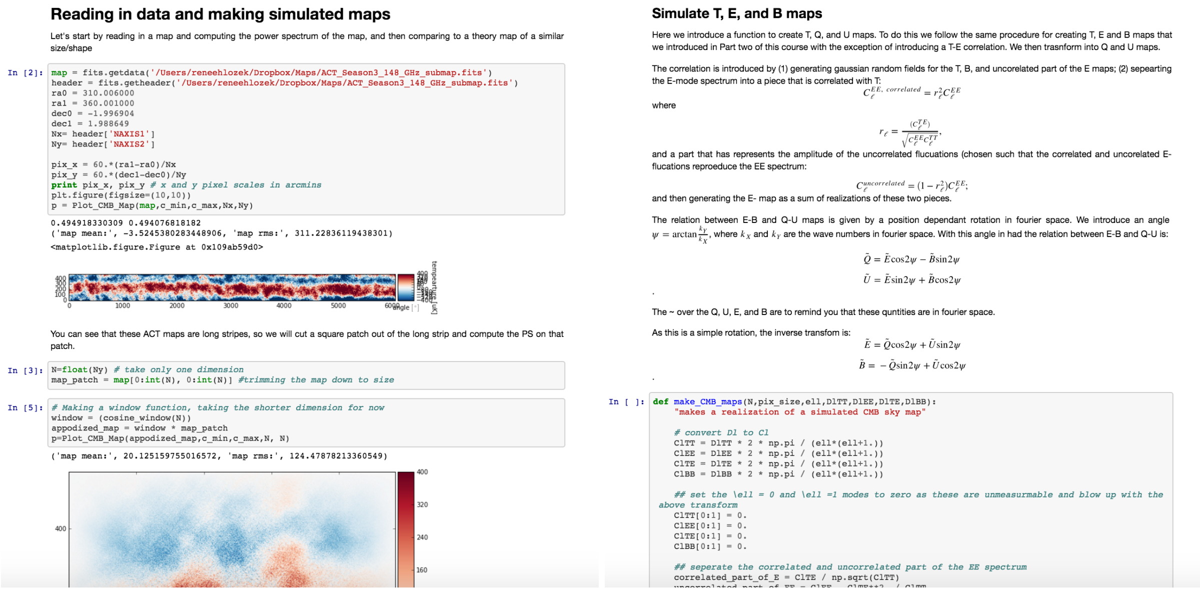

Figure 4: Examples of the workbooks used in the larger group projects for the second day of the workshop. The left panel shows the workbook which used ACT data from the public NASA respository (`https://lambda.gsfc.nasa.gov/product/act/act_prod_table.cfm`), while the right hand panel describes the simulation of polarisation maps.

It was understood that the groups would not tackle more than one or two of the larger projects during the school. The workbooks were available to all (indeed, the repository for all workbooks is open source and publicly available); the students' progress was then set by their own pace and interest. While the students worked, the instructors moved through the room and answered particular questions related to the individual workbooks.

3 Participation

The summer school was attended by 24 students and postdocs from seven institutions, with female participants making up almost a third of the total school. For this inaugural workshop, the advertising was targeted within the ACT, PolarBEAR and SPT collaborations; future schools will advertise more broadly with an aim to reach students who are not familiar with CMB analysis techniques and are not working within any given experiment. The attendees are shown in Figure 5.

For this workshop, there was a focus on graduate students and early postdoctoral researchers. There were a few students that were early-career (for example the school was attended by one senior undergraduate and a student who was in the first weeks of graduate school). They served as a test case for the utility of the course materials to those who are less well versed in the technical language of the CMB. The students learned a great deal and are now contributing directly to analysis as well as instrumentation.

4 Workshop Feedback Survey

At the end of the school, the participants were asked for anonymous feedback on the school and the topics covered. The career demographics of those participants who responded to this survey are summarised in Figure 6.

The questionnaire was designed particularly to determine if the participants were satisfied with the pace of the school (this is shown in Figure 7), and to determine which larger projects they attempted (shown in Figure 8).

5 Lessons learned and Future Plans

Feedback from the respondents was generally very positive, and the school will continue in coming years at various ACT institutions. Based on the lecturers' observations and the feedback from the participants, some refinements will be made. For example, the group projects should be outlined and explained in more detail. Some participants felt that the workbooks

Figure 5: The participants of the inaugural CMB analysis summer school held at the University of Michigan.

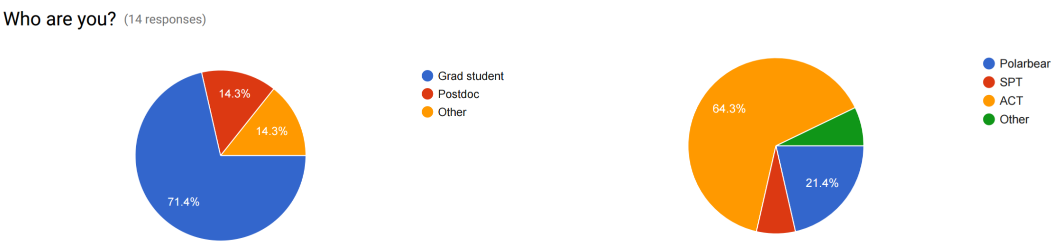

Figure 6: The respondents to the feedback survey.

should be made more separate (i.e. there was an issue with redundancy of variables and making sure that code didn't overwrite previous work). Suggestions for future schools include a more detailed look at the analysis pipelines, and how the beams are measured, how time ordered data becomes maps (this was given as a group exercise but should be revised).

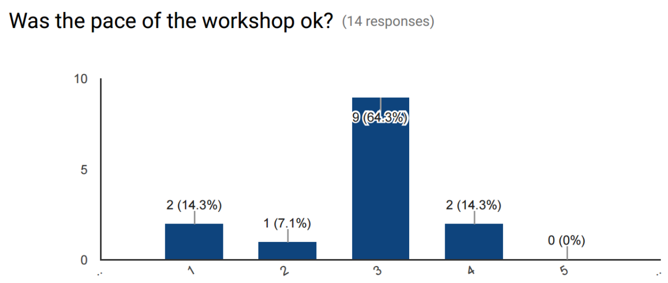

Figure 7: The responses to the question about the pace of the school. There were four workbook topics covered on the first day (with the morning of the second day given as a summary and introduction to the fifth workbook). The rest of day two was given to working on the group projects. The y-axis is the number of respondents who supplied a given answer. The scale ranged from 1 ("It was too slow. I could do everything with ease.") to 5 ("The pace was too fast, I struggled to keep up."). The form is attached at the end of this document.

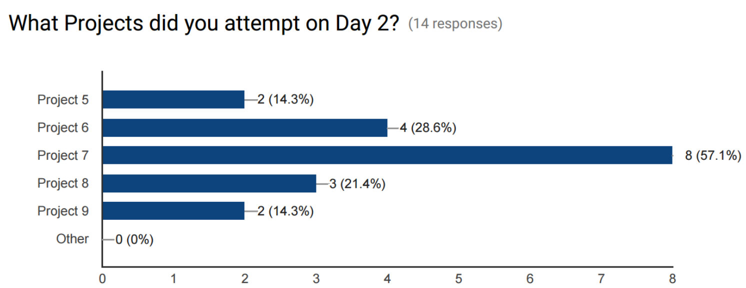

Figure 8: The breakdown of the projects that groups attempted, including the stacking analysis project five, which was carried over from the previous day. A large number of groups did the polarization exercise, with some also attempting the application to real data and the MCMC analysis. Most groups were able to tackle more than one project (albeit not to completion). The feedback form is attached to the end of this report.

As explained earlier, one of the main goals for future schools is to improve access to CMB tools and techniques for those who are not explicitly part of a CMB collaboration, and to improve the diversity (both racial diversity and gender balance) of the CMB community.

We seek to expand the school along the model of this first one, and perhaps host multiple schools in tandem. Given the open source nature of the course, this is possible but would require coordination with different groups. One model would be for the instructors to travel 'with the workshop' to institutions that may be more geographically remote or without financial ability to send students to the school. In order to provide access to the students, there ought to be no more than 15 participants per instructor; and ideally two instructors should be present. This allows one instructor to present the material with one roaming for questions.

6 Summary

The CMB Analysis Summer School provides participants with hands-on experience in CMB analysis, from the simulation of maps to the parameter estimation techniques used to constrain cosmological models. This end-to-end curriculum is developed over a two day workshop with interactive workbooks and discussion. The CMB Analysis Summer School will not only enrich the current generation of researchers, but provides an avenue to teach these analysis skills more broadly, and to recruit researchers from further afield into the field of CMB cosmology.

In setting up the 'flipped classroom', we designed the workshop to be scalable and flexible. While in-person participation is preferred, the material is designed to facilitate remote participation. This model is transferable to other groups and subject areas: we hope to broaden the content from CMB-focused to a broader cosmology summer school that will improve skills of our students and postdoctoral researchers more broadly.

CMB Analysis Summer School

Jeff and I are pleased that we could have you at the first CMB Analysis Summer school held at the University of Michigan, and would value your feedback! This won't take long, and will help us improve future schools.

*Required

1. **Who are you?**
Mark only one oval.

- ( ) Grad student
- ( ) Postdoc
- ( ) Other: ______________

2. **Where are you from?** *
Mark only one oval.

- ( ) Polarbear
- ( ) SPT
- ( ) ACT
- ( ) Other: ______________

3. **How prepared did you feel for the workshop?**
Mark only one oval.

	1	2	3	4	5	
Totally prepared - the initial workbook helped	( )	( )	( )	( )	( )	Unprepared - I was lost throughout

4. **How were the logistical arrangements?**
Mark only one oval.

	1	2	3	4	5	
Great, no issues - communication was great.	( )	( )	( )	( )	( )	I was confused about some issues.

5. **Was the pace of the workshop ok?**
Mark only one oval.

	1	2	3	4	5	
It was too slow - I could do everything with ease.	( )	( )	( )	( )	( )	The pace was too fast, I struggled to keep up.

6. **Do you have suggestions for the individual Parts 1-4 on Day 1? Be as specific as you like.**

7. **What Projects did you attempt on Day 2?**

Tick all that apply.

- [ ] Project 5
- [ ] Project 6
- [ ] Project 7
- [ ] Project 8
- [ ] Project 9
- [ ] Other:

8. **Do you have suggestions for the more self-guided "Projects" Parts 5-9?**

9. **What would you have liked to see that wasn't on offer here?**

10. **If you are willing, please drop your code in the link below: https://www.dropbox.com/sh/o4sqjst7s2wx9ot/AACrlsoX5DssPNh00RwGnGxJa?dl=0**

Tick all that apply.

- [ ] Yes, I added it
- [ ] Nope, didn't have any code to share.