In order to properly study the origins of the universe, it is necessary for cosmologists to go back in time, looking at objects and radiation well beyond the Milky Way. One of the oldest and thus most important sources of data is the cosmic microwave background, a near isotropic glow centered in the microwave range of the electromagnetic spectrum that is not associated with any other celestial object. A remnant of the epoch of recombination formed from the photons that were themselves formed from the formation of neutral hydrogen atoms, this cosmic microwave background radiation is polarized into E- and B-modes via Thomson scattering; gravitational waves produce polarized signals such that they contain both components, but the signal due to temperature fluctuation is present within the E-mode only, while the B-mode remains constant. The current models for the universe's long-term inflation support the idea that a primordial gravitational wave background should still be present within the CMB, which itself can be extrapolated from the B-modes of the CMB; as such, by detecting and analyzing those B-modes to greater accuracy, cosmologists are able to draw out information about the rate of expansion of the universe at older and older times in the universe's history.

However, this detection is made difficult, not only because of the much stronger E-mode polarization but because of the foreground radiation that arises from dust and obscures the cosmic microwave background radiation. Cosmic dust acts like a blackbody, radiating light depending on the temperature it is heated to by the interstellar radiation field. This radiation is typically on a scale comparable to or greater than that of the cosmic microwave background radiation, meaning that a typical observation of the CMB is going to have foreground contamination from the Milky Way that has to be accounted for in the predictive model. The level of dust radiation is going to vary relative to the CMB depending on the observance frequency, while the cosmic microwave background is going to show the characteristic curve of a 2.73 K blackbody at all times; with that in mind, current foreground models uses multiple sets of frequency data in order to differentiate the foreground radiation from the desired background radiation.

For this experiment, these multiple sets of data in question come from three different observation experiments that ran between 2001 and 2015 (technically the latter is ongoing, but the accessible data only goes to 2015): the Wilkinson Microwave Anisotropy Probe (WMAP), the *Planck* space observatory operated by the European Space Agency, and the Background Imaging of Cosmic Extragalactic Polarization and Keck Array (BICEP/Keck) experiments. Each of them ran or are running with the goal of measuring the temperature or the polarization of the CMB for both the E- and B-modes, and each ran over different sets of frequencies. The oldest experiment, WMAP, obtained data over five different frequencies ranging from 23 to 94 GHz; however, the only data available to this particular model are the 23 and 33 GHz maps, both of which have relatively high amounts of noise (200 μK/arcmin^2*pixel). *Planck*, which observed the sky for two years directly after WMAP ended, has a much larger range of data to select from, with maps available at 30, 44, 70, 100, 143, 217, and 353 GHz. Finally, the 2015 data release from BICEP/Keck has maps at 95, 150, and 220 GHz, with much lower noise levels when compared to either WMAP or *Planck*. With that in mind, various models were creating using different maps and frequencies as the basis, with the differences in the outputs being a major part of the analysis itself

The actual CMB data that is typically analyzed consists of angular power spectra based on the observed radiation levels at multiple frequencies and the cross spectra between those maps; for example, a comparison between the 95 GHz BICEP-Keck data and the 100 GHz Planck data would have two power spectra, one from each map, and the cross spectrum between the two maps, while a comparison between three separate maps would have three power spectra and three corresponding cross spectra. This actual data is in turn compared to a parametrized model that has estimated values for the magnitude of the CMB radiation and the magnitude of the dust radiation, among other parameters. By comparing the accuracy of the modeled data to the actual data, the range of expectation values for the CMB and dust radiations can be narrowed down into a smaller and smaller range. The issue, then, is determining how likely a specific model is – if the derived parameters of the model match the observed data values but the model is a highly unlikely one, the model itself must be analyzed carefully to make sure it is realistic. This and previous models were based around Bayesian inference, building model parameters *q* from the product of priors of the said parameters and the probability *x* of the data:

$$p(q|x) = \frac{p(x|q)p(q)}{p(x)}$$

The issue with this Bayesian modeling is that the probability values *x* may not be feasible to calculate as the amount of available data increases, due to either the high systematical variance or the increasing computational time. In his August 2017 paper “GLASS: A General Likelihood Approximate Solution Scheme”, Steven Gratton of Cambridge proposes an alternative method for computing these likelihood values using the means and derived cumulants of the parameters rather than any existing priors. While his methodology is proposed for any general mathematical functions involving large sets of data, he notes that it is specifically applicable to CMB analysis and galactic redshift surveys; it is because of this applicability that we decided to further investigate this alternative method and see whether the values it provides for existing CMB data match up with pre-established values derived from other methods.

The underlying principle behind the calculation methodology as presented by Gratton is to determine the broadest possible sampling distribution by maximizing the entropy of the system while still keeping consistent the assumed parameters defined for the model and the known statistics calculated before creating the model. Around some statistic *x* as a function of a defined parameter q, with some known prior probability p_0, this entropy is expressed as:

$$H(p) = -\int p(x) \ln \frac{p(x)}{p_0(x)} \, dx$$

which in turn yields the probability distribution over x:

$$p(x) = \frac{p_0(x) \exp(-\lambda_x x - \lambda_{xx} x^2)}{\int p_0(x) \exp(-\lambda_x x - \lambda_{xx} x^2) \, dx}$$

where λ_x and λ_{xx} are Lagrange multipliers that impose the necessary constraints on the statistic *x*. These constraints are what allow the likelihood model to conform to known parameters, and solving for these multipliers numerically as functions of the parameter *q* (denoted as $\lambda_x(q)$ and $\lambda_{xx}(q)$) gives an approximate probability distribution for *q* for any given or modeled value of *x*; immediately taking that likelihood p(x) and multiplying by some prior p(q) outputs the posterior p(q|x), which is the main value that is being sought after in the first place. Gratton proposes moving further from that distribution $p(x)$; denoting the denominator of the original distribution *Z*, he introduces the action distribution *S* equivalent to the negative log of the probability distribution:

$$S(x,q) = -\log p(x) = -\log p_0(x) + \lambda_x(q)x + \lambda_{xx}(q)x^2 + \log Z(\lambda(q))$$

This action distribution can theoretically be extended to multiple dimensions (either multiple statistics x^i, $i = 1,\dots,n$ or multiple parameters q^a, $a = 1,\dots,m$) with the Lagrange multipliers becoming indexed matrix functions of the parameter and the action distribution expanding to multiple indices:

$$S(x,q) = -\log p_0(x) + \lambda_i x^i + \lambda_{(ij)} x^i x^j + \lambda_{(ijk)} x^i x^j x^k + \cdots + \log Z(\lambda(q^a))$$

However, solving for the multidimensional Lagrange multipliers requires multidimensional numerical integrals that become rapidly difficult to the point of being unfeasible with every new parameter. The strength of this model comes from the fact that it is possible to obtain it without explicitly solving for those Lagrange multipliers for certain classes of models.

The final result of this vector manipulation yields an equation for the derivative of the action distribution in which the Lagrange multipliers do not appear:

$$\frac{\partial S}{\partial q^a} = -(X - \langle X \rangle)^T \langle\langle XX^T \rangle\rangle^{-1} \frac{\partial \langle X \rangle}{\partial q^a}$$

where X and X^T are simply the vector for the mean values of the three power spectra obtained from comparing two different maps, and $\langle\langle XX^T \rangle\rangle$ is the covariance matrix for that vector X. This model requires only the desired moments of whatever statistic(s) *x* is being analyzed, the derivatives of those moments with respect to the parameter(s) *q,* and the second cumulants from the theoretical model of *x.* Integrating that value over two separate points in *q*-space yields a value for the action distribution over that range, which in turn yields a value for the probability distribution over that same range. Gratton goes on to apply this theoretical model to three cases that are specifically relevant to analyzing CMB data: an auto power spectrum with 2l + 1 map modes, a correlated power spectra example, and a cross-spectrum example with noise involved, all of which were only mapped over one parameter. With confidence that the model worked in one dimension, Gratton proposed that the next useful step would be understanding how taking a multidimensional integral affects the error in the action distribution and whether the potential path-dependence of the integration will affect the final action and likelihood distributions. Analyzing that path-dependence for multidimensional integrals was the crux of this particular project as a check for both the theory as a whole and as a way to see whether this model could potentially be used as a less-computational intensive version of the current Hamimeche-Lewis model.

Before moving directly to the multidimensional analysis, multiple steps were taken to build up an understanding of the model and how to relate the math to a general program that could take in and model data from various CMB sources and at varying wavelengths. First, a series of toy bandpowers was produced over the different frequencies as detailed above, with model "maps" consisting of a vector of random numbers distributed over a Gaussian curve, mean of zero. For each of the observing frequencies, a model map would be created for the CMB, the dust contamination (with amplitude dependent on the frequency), and random observational noise. The power spectra could then be obtained for each frequency as well as cross spectra for separate frequencies by taking the dot product of two corresponding maps and dividing by the number of modes; for this experiment that number of modes was kept to 40, roughly corresponding to the number of modes present in the BICEP/Keck data. After working with those toy models, the model as proposed in Gratton's article was extended to multiple maps at once while keeping the other foreground parameters unaccounted for, followed by doing the reverse and activating the foreground parameters over a single map. Once that had been established, the next step was running the model with multiple simulated maps and parameters activated at once and comparing the derived probability distributions with those actually found from the Hamimeche-Lewis model to see whether this new Gratton model held up at higher levels.

In the current state, the model is able to compare two separate bandpowers and provide both the log likelihood and the true likelihood models of the dust value and the CMB power based on those bandpowers and the predicted value, which is assigned as a constant within the model itself. In the case where the two bandpowers are close together (i.e. 143 and 150 GHz), the dust and the background radiation are hard to disentangle, and as such the likelihood is more likely to have a strong negative correlation between the two parameters; the model itself has to take into account the possibility that there is high dust contamination and low background radiation or low dust contamination and high background radiation. With that in mind, the model tends to have a strong negative slope, with values centered on the expected values as pre-determined within the model itself. In the case where the bandpowers are very far apart, the radiation values of the dust are going to be very different when compared to the background radiation, and as such are going to be much easier to disentangle from each other. As such, there is going to be a much weaker correlation between the two parameters, and the potential error in the dust is going to be much smaller. The model tends to have a very small slope and be much more tightly centered on the expected parameter values. As it is now, this model holds up far better for the second case rather than the first, as shown in the figures below (note that the axes are reversed on the GLASS approximation and the Hamimeche-Lewis approximation):

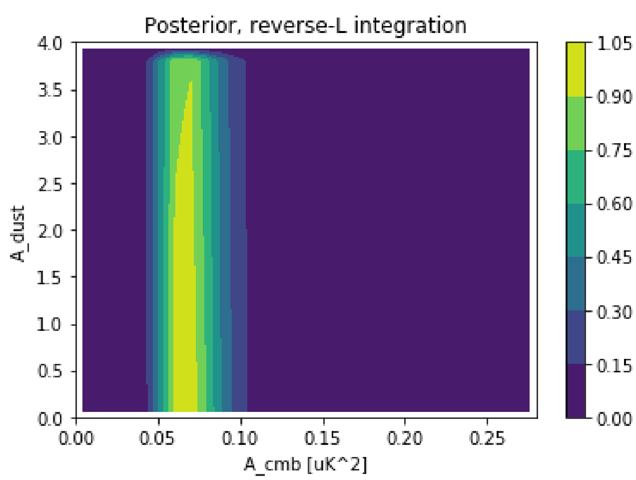

Figure 1a: *The GLASS approximation model for the 143 and 150 GHz maps, centered at A_{cmb} = 0.07 and A_{dust} = 1*

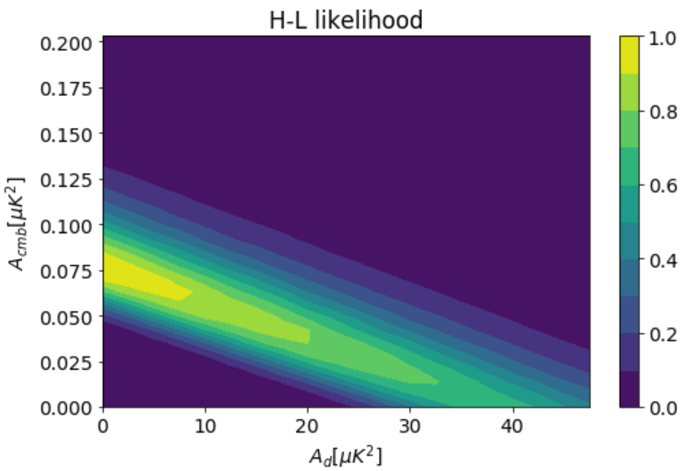

Figure 1b: *The Hamimeche-Lewis approximation model for the 143 and 150 GHz maps, centered at A_{cmb} = 0.07 and A_{dust} = 1*

The GLASS model here is overly constrained, not showing the expected correlation between the two parameters that has been established from the Hamimeche-Lewis model. However, the model is much more properly constrained for the second case:

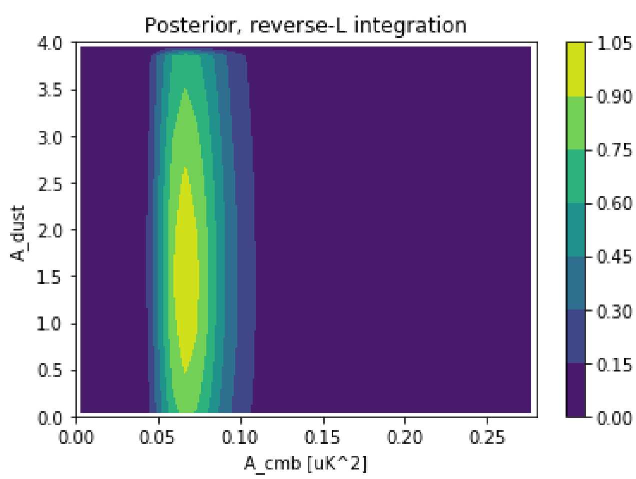

Figure 2a: *The GLASS approximation model for the 150 and 343 GHz maps, centered at A_{cmb} = 0.07 and A_{dust} = 1*

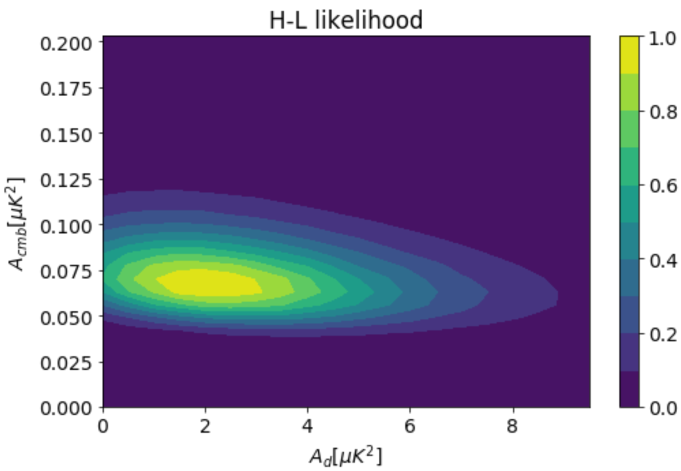

Figure 2b: *The Hamimeche-Lewis approximation model for the 150 and 343 GHz maps, centered at* $A_{cmb} = 0.07$ *and* $A_{dust} = 1$

Both the shape of the model and the overall range for potential dust values are much closer to the expected values as derived from the Hamimeche-Lewis likelihood, supporting the idea that this model at least works for the further apart bandpowers.

As of now, this model still has plenty of room to be improved; it was due to the time constraints of the semester that we weren't able to explore the mathematics behind the modeling itself more. There are two major steps to be pursued next; firstly, the current model is based off of a single method of integrating the log likelihood over the full range of values. However, because this model is a two variable system, there are multiple potential paths of integration that the model could take, and it is possible that the output could be more or less constrained based on which path of integration is taken. Developing a model that can compare the results from different integration paths will show whether the GLASS model truly is universal or whether there are certain cases where it does and doesn't work. Secondly, and more importantly for long term utility, is applying this model to real-world data. Currently the values of the dust and background parameters that the model is using are assigned within the code itself and based on what is assumed to be the case; by using real-world values for each parameter the model itself will be able to provide a more accurate picture and hopefully begin to be used on a much wider scale.

Works Cited

Bennett, C. e. (2013). Nine-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Final Maps and Results. *Astrophysical Joural Supplement Series*.

Gratton, S. (2017). GLASS: A General Likelihood Approximate Solution Scheme. *arXiv* .

Hamimeche, S., & Lewis, A. (2012). Likelihood Analysis of CMB Temperature and Polarization Power Spectra. *Physical Review*.

Kamionkowski, M., & Kovetz, E. (n.d.). The Quest for B Modes from Inflationary Gravitational Waves. *Annual Review of Astronomy and Astrophysics*, 2015.

Keck Array, BICEP2 Collaborations. (2015). BICEP2 / Keck Array VI: Improved Constraints On Cosmology and Foregrounds When Adding 95 GHz Data From Keck Array. *Physics Review Letters*.

Planck Collaboration. (2015, August 10). Planck 2015 results. I. Overview of products and scientific results.

Planck Collaboration. (2016). Planck intermediate results. XXX. The angular power spectrum of polarized dust emission at intermediate and high Galactic latitudes. *Astronomy & Astrophysics*.