PROJECT DESCRIPTION

1 Introduction

Mapping Our Universe: The earliest examples of astronomy included following the nearby planets and charting the "fixed" stars which were projected onto the celestial sphere and organized into constellations. Ultimately this led to a physics-based, low resolution, 3D description of the Galaxy. The situation today in cosmology is somewhat similar. We have large surveys of comparatively nearby galaxies [1–10] and a splendid two dimensional map of the microwave background [11], all of which have led to a "standard cosmological model" of the Universe in which inflation-based, Gaussian, potential fluctuations, with a well-defined spectrum, grew according to deterministic laws to produce contemporary large scale structure in a flat Universe endowed with a cosmological constant. However, the traditional goal of astronomy, to describe the complete disposition of this actual structure, has hitherto been subsumed into statistical investigations designed to elucidate the underlying physics.

We present a staged proposal to combine recent observations with what we have learned about the physics to *make the best map we can of the 3D structure of the Universe within and slightly beyond our horizon.* In addition to satisfying a natural desire to describe our Universe, success in this program will naturally furnish ongoing and planned physics investigations with additional priors which should tighten up their accuracy, and enable us to make scientific predictions about future surveys.

Before proceeding to a detailed description of our proposed program in Section 2, we first provide a brief summary of the context in which it will be carried out. Brief outlines of our plans regarding personnel, schedule and broader impact can be found in Sections 3–5 respectively.

Our Universe's Contents: The last decade has seen remarkable advances in cosmology, spearheaded by increasingly detailed measurements of the cosmic microwave background (CMB) radiation (see e.g. [11, 12]). These accurate measurements have affirmed that a description of a homogeneous, spatially flat general relativistic Universe with relatively few ingredients – photons ($T_\gamma = 2.7$ K), neutrinos (three flavors), baryons ($\Omega_b = 0.05$), dark matter ($\Omega_d = 0.26$) and a cosmological constant ($\Omega_\Lambda = 0.69$) supplemented by (almost) scale-free, adiabatic, Gaussian initial perturbations suffices to describe essentially all that is secure in the observations [12]. There is still room for revision, retraction and major discovery but, right now, we have a good working hypothesis that the Universe is basically this simple (e.g. [13, 14]). There is some tension in the reported measurements, *e.g.* of the Hubble constant, but this is not important for our purpose and we shall simply adopt Planck values. Much effort is being expended to see if a ubiquitous and eternal cosmological constant needs to be replaced by a dynamical dark energy. If this turns out to be true then only simple changes will be needed to what follows.

Our Universe's Evolution: The description of the average expansion of the Universe is relatively uncontroversial. When $t \sim 50$ kyr, the scale factor – the size of a region relative to its contemporary size – was $a \sim 0.0003$ and the Universe became (dark) matter-dominated. When $t \sim 380$ kyr, $a = 0.00093$, the hydrogen plasma quickly formed atoms, decoupling from the radiation and forming the inside-out, CMB photosphere where the majority of CMB photons we observe today were last scattered with a temperature ~ 2900 K. When $t \sim 600$ Myr, $a \sim 0.1$, the first stars formed and the Universe (re)-ionized. This epoch is becoming accessible to observation. Finally when $t \sim 8$ Gyr, $a \sim 0.6$, the cosmological constant began to dominate the matter, and the universal expansion started to accelerate.

Fluctuations in Our Universe: The observed fluctuations are adequately described by a set of spatial Fourier modes expressed in terms of contemporary or comoving coordinates. These modes are longitudinal and the ones that mostly concern us here evolved linearly. It is convenient to describe their amplitude using the (effective) Newtonian potential/gauge Φ from which the density and fluid velocity perturbations can be computed (in the "Sachs-Wolfe" limit at long wavelength [15]). [1] The modes that mostly concern us remain linear and are "adiabatic", so that we only need to know their amplitude and phase at one epoch, e.g. recombination, to predict them for all time.

It has recently been demonstrated, mainly using CMB observations [12], that the adiabatic hypothesis is quite accurate. Furthermore, the amplitudes associated with each mode of the initial potential scale approximately as $k^{-3/2}$ and are drawn from a Gaussian distribution, so that the potential fluctuations associated with each length scale are scale-independent. This behavior is consistent with a remarkable early conjecture by Harrison [16] as elaborated by Zel'dovich [17].

The potential associated with a mode was essentially frozen until it "entered the horizon," that is, until the timescale for its dynamical evolution, became smaller than the expansion timescale. We are mostly concerned with long wavelength modes for which this happened after recombination and $k \lesssim 40\text{Gpc}^{-1}$.[2] For such wavenumbers, the waves evolve at roughly constant Φ until the cosmological constant takes over and the potential falls by roughly 20 percent today. We make this explicit in Fig. 1 where we show the interior of the last scattering surface in comoving coordinate space and a particular wave whose amplitude and phase we are trying to measure.

The Inflation of Our Universe: The flatness of the geometry today, the isotropy of the CMB temperature, and the very existence of fluctuations with wavelengths longer than naively allowed by causality are all consistent with the simplest version of a much more specific and even bolder conjecture by Guth [18], Linde [19] and others (e.g. [20–25]), that the Universe underwent a period of "inflation" at much earlier times. This theory is based on the idea that all the structures in the observed Universe emerged from quantum fluctuations about 10^{-33} seconds after the Big Bang. Inflation, which describes a phase of accelerated cosmic expansion, is the leading theory providing a causal mechanism for generating these fluctuations and stretching them to cosmological scales. The microphysics of inflation makes detailed predictions for the spectra of these fluctuations as observed in the CMB, in particular a slight tilt in the power spectrum, which has been measured (see e.g. [26, 27] and references therein).

[1] There may also be tensor modes which we shall ignore here. If, and when, they are detected, only minor modifications will be needed.

[2] k_0 can be considered as approximately the wavenumber associated with the first acoustic peak and its consequence, Baryon Acoustic Oscillations (BAO).

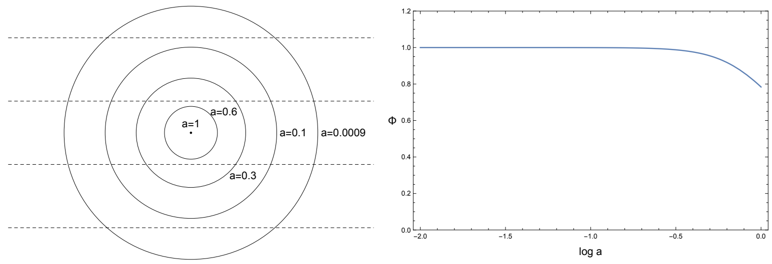


Figure 1: a) The local Universe interior to the CMB photosphere expressed in comoving coordinates. The circles are, in order, $a = 0.6$, ($z \sim 0.7$), schematically the limit of present surveys, $a = 0.3$, ($z \sim 2$) roughly the effective limit of future surveys, the nominal Epoch of Reionization at $a = 0.1$, ($z \sim 9$ and the CMB photosphere at $a = 0.00093$, ($z \sim 1100$) and a distance $x_{\rm CMB} \sim 13.9$ Gpc. The comoving radius of the big bang is 14.2 Gpc. Also shown as dashed lines are the nodes of a single wave mode with $k \sim 0.45$ Gpc^{-1} which contributes significantly to spherical harmonics with $\ell \lesssim 10$. b) Variation of the amplitude of this wave with scale factor a.

Qualitatively, the causal mechanism seeding the primordial perturbations is easily understood. During inflation, the Hubble radius, H^{-1}, which can be thought of as the "apparent horizon", was roughly constant. Meanwhile, quantum fluctuations in the matter field(s) and metric are constantly generated with wavelength H^{-1} at most [28]. Once produced, a fluctuation with comoving wavelength λ is stretched with the expansion of space past the Hubble radius, at which point its dynamical timescale, $\sim ac/k$ was larger than the expansion time and it "exited" the horizon and its amplitude froze. Throughout inflation, such fluctuations are continuously created at the physical scale H^{-1}. Therefore, by the end of inflation, perturbations will finally have been produced on a whole spectrum of physical scales.

2 Proposed Research

The long term goal of this proposal is to connect the CMB to local surveys by producing an evolving three-dimensional (or in other words, four-dimensional) map of the Universe that is valid from before 380 kyr to today, and out beyond 14 Gpc. The exercise is not purely cartographic, as it is essential that we use secure physical inferences about the early Universe in making this map.

The first stage of our proposed program is to use 2D CMB observations alone to test internal consistency and to recover as much as we can of the 3D potential, velocity and density fields interior to the last scattering surface. In the second stage, we will augment CMB data with existing 3D measurements from galaxy surveys, gravitational lensing, intermediate Sachs-Wolfe measurements and so on. This should improve the resolution. The third and final stage will be to use our constrained model to make quantitative predictions about the density field as could be probed by future surveys, and will include an investigation of how far it is possible, even in principle, to reconstruct the structure of the Universe including what can be inferred beyond our horizon.

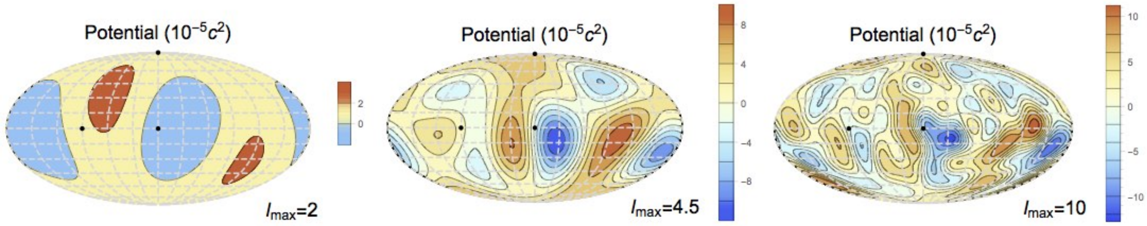


Figure 2: Photospheric potential fluctuations of the CMB for $\ell_{\max} = 2, 4.5, 10$ derived from Planck data shown as Mollweide projections.

2.1 Stage 1. From 2D to 3D: Potential Reconstruction from CMB Data

Temperature Fluctuations: Observations of the CMB are the foundation on which modern quantitative cosmology rests. The conventional way to describe the observations is in terms of spherical harmonics – the generalization of Fourier modes to a sphere – labeled by ℓ and m. It is convenient to use an equivalent vector of real spherical harmonics, $Y_y(\theta,\phi) = \{Y_{0,0}, Y_{1,0}, 2^{1/2}\Re[Y_{1,1}], 2^{1/2}\Im[Y_{1,1}], Y_{2,0}, \ldots, 2^{1/2}\Im[Y_{\ell_{\max},\ell_{\max}}]\}$ of length $(\ell_{\max}+1)^2$ and where θ, ϕ are standard spherical polar coordinates. Note that there are $2\ell+1$ independent, real, basis function in each ℓ-shell. Note also that $\int d\Omega Y_y Y_{y'} = \delta_{yy'}$. It is convenient to treat $\ell_{\max}$ as a continuous variable by adding a fraction between zero and unity of the largest ℓ shell and therefore smoothly change the angular resolution. (The use of a real basis helps identify systematic effects.)

Most investigations have focused on measuring the "power" in the temperature fluctuations (including polarization) associated with a given ℓ, obtained by summing products of the coefficients of the harmonic components over m, and comparing it with the predictions of various cosmological models. This program has been wonderfully productive, and has resulted in the world model just outlined. Furthermore this power spectrum has been successfully reconciled with features of the local Universe, such as galaxy counts. One common assumption is that the particular realization of the Universe that we are observing is drawn from a statistical ensemble of universes. When ℓ is large, we have many independent measurements on the associated angular scale, $\sim \pi/\ell$, and so we can measure an rms value for the harmonic component with a small variance. However, when ℓ is small, we have only a few such measurements and the "cosmic" variance is large. Despite their great value, these statistical measurements inevitably discard information which may be valuable.[3] In the proposed study, only one specific realization of the Universe – the one we inhabit – is considered.

In order to start us on a pilot investigation, Planck team members Wehus & Eriksen (Oslo) have kindly supplied 100 sample Planck temperature fluctuation maps for $0 \le \ell \le 10$ or $1 \le y \le 121$. From this ensemble we are able to compute the mean pho-

[3]This is in the sense that music is far more than a "flicker" power spectrum. To pursue our musical metaphor, different voices and instruments contribute different ranges of frequencies to a musical performance over a total range of roughly ten octaves. We are only listening to the bass range but higher voices and instruments can still contribute to what we hear.

tospheric potential fluctuation map at the time of recombination $\Phi = a_y Y_y$, (including the monopole and dipole components and adopting the summation convention) and the covariance matrix $C_{yy'}$ associated with the harmonic components a_y. We find that this matrix is invertible and can be used directly up to $\ell = 8$. More careful treatment of the data is needed beyond this. The fractional variance in individual harmonics varies between ~ 0.0001 and ~ 0.01. Undoubtedly, there are systematic effects present in this data set which need to be explored, but the accuracy is high enough to proceed without this.

Fourier Modes: It is conventional to Fourier expand the potential Φ today; the expansion at other times is then simply calculable. Although the full spectrum of the Fourier modes we are discussing is continuous in $\mathbf{k}$, the fact that our observations are made over a restricted volume means that we can treat the waves as a discrete Fourier transform of modes associated with a box in comoving space of side L on which periodic boundary conditions are imposed. L is chosen here to have a compromise value of four times the radius of the CMB photosphere, 13.9 Gpc which we adopt as our unit of length.

$$\Phi[\mathbf{x}(r,\theta,\phi)] = \sum_{n=1}^{N/2} [f_n \cos(\mathbf{k}_n \cdot \mathbf{x}) + f_{N+1-n} \sin(\mathbf{k}_n \cdot \mathbf{x})] \tag{1}$$

where the coefficients f_n are also real and $\mathbf{k} = \Delta k\{n_1, n_2, n_3\} = k\{\sin\theta'\cos\phi', \sin\theta'\sin\phi', \cos\theta'\}$, with n_1, n_2, n_3 integers and $\Delta k = 2\pi/L = \pi/2$. We restrict the sum to $(n_1^2+n_2^2+n_3^2)^{1/2} \leq n_{\rm max}$ and only need consider $\mathbf{k}$ over a hemisphere (since the potential must everywhere be real.) We label the coefficients by the index n running from 1 to $N \sim 4\pi n_{\rm max}^3/3$. ($N = 6$ through 4168 for $n_{\rm max} = 1$ through 10.) $\Phi(\mathbf{x})$ can be expanded formally as an infinite sum of Legendre polynomials and approximately as a finite sum:

$$\Phi(\mathbf{x}; \ell_{\rm max}) = \sum_{\ell=0}^{\ell_{\rm max}} (2\ell+1) \sum_{n=1}^{N/2} j_\ell(k_n x) P_\ell(\hat{\mathbf{k}}_n \cdot \hat{\mathbf{x}})[\cos(\ell\pi/2) f_n + \sin(\ell\pi/2) f_{N+1-n}]. \tag{2}$$

Gaussian Prior: Detailed study of the CMB (e.g. [29, 30]) has led to the conclusion that the amplitude of each discrete mode with wave vector $\mathbf{k}$ is well modeled as having been drawn from a Gaussian distribution of variance $\sigma_{\mathbf{n}}^2 = \alpha^2(n_1^2+n_2^2+n_3^2)^{-3+(n_s-1)}$, where n_s is the scalar spectral index. Adopting this model, we can fix the spectral tilt to the Planck best-fit value (this assumption will be relaxed in the next subsection in order to estimate the inflationary hyperparameters) and we can determine the coefficient α approximately and empirically, by computing the spherical harmonic coefficients $a_y = \mathbf{R}_{yn} f_n$ for sets of Gaussian f_ns where the "response matrix" is given by:

$$\mathbf{R}_{yn} = 4\pi Y_y(\theta'\phi') j_\ell(k)[\cos(\pi\ell/2), \sin(\pi\ell/2)] \text{ for } [1 \leq n \leq N/2,\ N/2 < n \leq N]\,, \tag{3}$$

and then adjusting α such that the model-predicted and measured spherical harmonics match. This approach is adequate for our pilot study; an improved determination will be investigated following the evidence analysis of [31], *en route* to a hierarchical modeling of the system where the normalization α is one of a set of hyperparameters governing the statistics of the potential field.

Preliminary Results: The simple question that motivated this investigation, and which did not seem to have a well-known answer, was how much of the 3D potential could be reconstructed interior to the 2D CMB photosphere using CMB observations alone. This is an example of what is sometimes called *holography.* At first sight this might seem hopeless, because if one associates $\ell_{\max}$ with $k\, x_{\rm CMB}$, then we are trying to solve for $O((k_{\max}\, x_{\rm CMB})^3)$ Fourier modes using only $O(\ell_{max}^2)$ spherical harmonics. However, if we confine our attention to the longest wavelength waves,[4] and use all the information that is at our disposal to exploit the high accuracy of the measurements while accepting uncertainty in the result, then it is possible to make some progress.

We approximately characterize the posterior PDF (which under our assumptions is a multi-variate Gaussian distribution) for the coefficients f_n by first finding its peak, minimizing the quantity

$$-2\ln\mathcal{P}\,(f_n|a_y) \approx (a_y - f_n\mathbf{R}_{ny})C^{-1}_{yy'}(a_{y'} - \mathbf{R}_{y'n'}f_{n'}) + \frac{f_n^2}{\sigma_{\mathbf{n}}^2} + \text{const.} \tag{4}$$

with respect to variation of f_n. Note that $\sigma_{\mathbf{n}}$ contains the normalization α.

Here, $C^{-1}_{yy'}$ is the inverse of the covariance matrix. This leads to the linear equations:

$$f_n = \left(\mathbf{R}_{ny}C^{-1}_{yy'}\mathbf{R}_{y'n'} + \frac{\delta_{nn'}}{\sigma_{\mathbf{n}}^2}\right)^{-1}\mathbf{R}_{n'y}C^{-1}_{yy'}a_{y'} \tag{5}$$

The approximate form given here is sufficient for a feasibility check: using mock data, we have found that *it is possible to solve for stable, low order maximum posterior Fourier coefficients, which can recover harmonic coefficients a_y that are consistent with the original data.* We have applied this approximate procedure to the actual Planck data, and our results are exhibited in Fig. 3. There are features of the data which we have yet to understand but this suffices to illustrate what we hope will be possible. It is proposed to refine and improve upon this basic approach to obtain the best map we can based upon the CMB data alone. In particular we intend to set clear, statistical criteria for assessing when it is significant to add additional Fourier components to the map.

Inflationary Origin of Perturbations and Hyperparameter Estimation: As explained above, inflation provides a mechanism for seeding perturbations in the CMB from quantum fluctuations in the early Universe. It is standard, as we shall do here, for most models to assume that the main matter components during inflation are in the form of scalar fields. Moreover, for simplicity, we shall assume that only one scalar field is dynamically relevant, so that we will work within the framework of "single-field" inflation. Apart from these caveats, the framework developed in what follows will remain model-independent.

We want to describe perturbations of the scalar field, which we call φ, about a homogeneous background. It is common to quantify these perturbations with the gauge-invariant

[4] Our approach is quite complementary to that of Yadav and Wandelt [32] who are concerned with wavelengths comparable with the thickness of the recombination surface around the acoustic peak near $\ell \sim 200$.

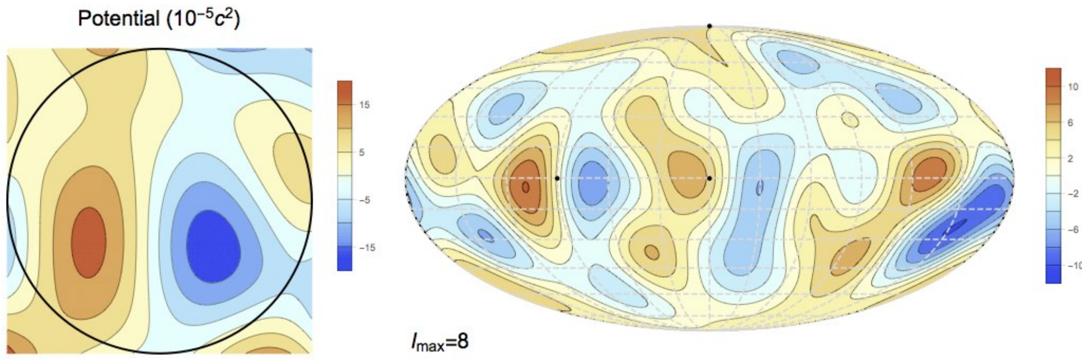


Figure 3: a) 3D gravitational potential at the time of recombination, recovered from 2D fluctuations on the CMB photosphere (represented by the circle), and displayed in the $x_3 = 0$ plane. In this pilot study, 81 spherical harmonics and a Gaussian prior were used to recover 122 Fourier components with $n_{\max} = 3$. The monopole and dipole were not removed. b) The predicted 2D potential fluctuations on the CMB photosphere were then recomputed, exhibiting encouraging consistency with the input data.

curvature perturbation, ζ.[5] In the co-moving gauge, the Fourier modes of ζ as they exit the horizon are

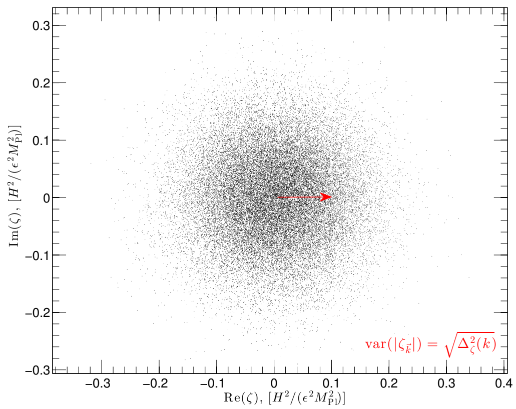
$$\zeta_{\mathbf{k}} = \sqrt{\frac{\pi}{2}} e^{i\theta} e^{i\frac{\pi}{2}\left(\nu+\frac{1}{2}\right)} (-\tau)^{\nu} \mathcal{H}_{\nu}^{(1)}(-k\tau)\,. \tag{6}$$

Here, ν is a slowly varying function which encapsulates the specific dynamics of the inflationary model, $\mathcal{H}_{\nu}^{(1)}$ is a Hankel function of the first kind, τ is the conformal time, defined by $d\tau = a(t)dt$, and θ is a random variable endowing each mode with a random phase. A prediction of inflation is that θ should have a uniform probability distribution between 0 and 2π. This is a fundamental prediction of inflation which stems from the assumption that the fluctuations are quantum mechanical in origin, the very assumption the idea of inflation is based upon.

However, the actual distribution of the phases for the modes of ζ (or Φ) in the CMB has yet to be measured. This is because, so far, most analyses of CMB data have strived to measure the power spectrum of fluctuations, Δ_{ζ}^2. Because the power spectrum is proportional to $|\zeta_{\mathbf{k}}|^2$, when such a measurement is made, the phase information of the modes is lost. With a reconstruction of the interior of the CMB sphere that outputs three-dimensional Fourier modes, this phase information will become readily available. Therefore, the new analysis of the CMB data we are proposing will make the measurement of the distribution of the phases possible. This will constitute a direct test of the quantum origin of the CMB fluctuations.

One on the main uncertainties pertaining to inflation remains the shape of the potential for the scalar field φ, which determines the specific dynamics of inflation, e.g. its energy scale and the precise shape of the power spectrum, bi-spectrum, etc. it produces. Using

[5] At late times (i.e. after the time of BBN) and inside of the horizon, when we use the Newtonian gauge to describe the Newtonian potential Φ, one can think of ζ and Φ interchangably. The main advantage of using ζ during inflation is that, in single-field inflation, the amplitude of its Fourier modes remain constant from horizon exit to horizon re-entry, significantly simplifying their evaluation [13].



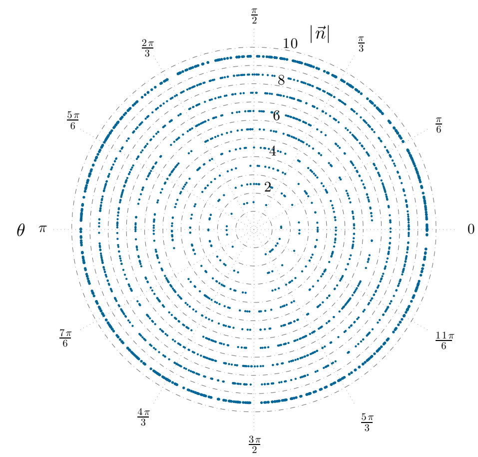


Figure 4: *Left panel*: Distribution of the modes of $\zeta_{\mathbf{k}}$ at fixed $|\mathbf{k}|$. *Right panel*: Representation of the norm of the first few k-modes ($|\mathbf{n}|$ from 2 to 10), versus their random spatial angle θ. If the individual modes were clustering around an angle, this would demonstrate that θ is actually not a uniformly distributed random variable, representing a serious challenge for inflation. Regardless of the distribution of θ, observations of Δ_ζ^2 would not be affected.

our low-k 3D reconstruction of the potential Φ, it will also be possible to reconstruct the inflationary potential over a limited range of φ. This is achieved by expanding the potential locally around a fixed $\varphi = \varphi_*$ in a model independent way in terms of the "slow-roll" parameters. We can then define the inflationary vector parameter

$$\vec{\eta} = (H_*^2/\epsilon_*,\, \epsilon_*,\, \eta_*,\, (\eta_2)_*,\, (\eta_3)_*)\,, \tag{7}$$

which defines uniquely a shape for the potential around φ_*. Here, the star quantities refer to their values when $\varphi = \varphi_*$. Using Bayes theorem, we then infer the probability of a given vector $\vec{\eta}$ given the data a_y, via the potential map f_n (which we marginalize over), $\mathcal{P}(\vec{\eta}|a_y)$,

$$\mathcal{P}(\vec{\eta}|a_y) = \frac{1}{(2\pi)^{p/2}}\left[\mathrm{Det}(C_{nn'})\mathrm{Det}(C_{yy'})\mathrm{Det}\,\mathbf{A}_{nn'}\right]^{-1/2} e^{\left\{\frac{1}{2}\mathbf{B}_n\mathbf{A}^{-1}_{nn'}\mathbf{B}_{n'}-\frac{1}{2}a^{\mathrm{T}}_{y'}C^{-1}_{y'y}a_y\right\}}\,\frac{\mathcal{P}(\vec{\eta})}{\mathcal{P}(a_y)}\,, \tag{8}$$

where $\mathcal{P}(\vec{\eta})$ is the prior on $\vec{\eta}$, $C^{-1}_{n'n}$ is a diagonal matrix with diagonal elements given by $1/|\zeta_{\mathbf{k}}|^2$, generalizing $1/\sigma^2_{\mathbf{n}}$ from our Gaussian prior (note that this relaxes the assumption of fixed spectral index n_s and amplitude α) and the $\mathbf{A}$ matrix and $\mathbf{B}$ vector are given by:

$$\mathbf{A}_{nn'} = \mathbf{R}_{ny'}C^{-1}_{y'y}\mathbf{R}_{yn'} + C^{-1}_{nn'}\,, \qquad \mathbf{B}_n = \mathbf{R}_{ny'}C^{-1}_{y'y}a_y\,. \tag{9}$$

Generalizing this formalism to inflationary parameters that produce primordial non-Gaussianities is straightforward, however, a key assumption in the above formalism is the assumption of a Gaussian prior for the f_n. To exploit the full information present in data samples containing an increasingly larger range of ℓ modes, we will further extend our numerical analysis and develop a pipeline capable of handling non-Gaussian priors for the f_n and solve for the $\vec{\eta}$ vector using MCMC sampling.

Detailed Approach Incorporating Planck Polarization Data: Up to this point, we have only considered the CMB temperature data. We anticipate making significant gains in map fidelity when we incorporate the Planck polarization data as well after it is made public. While these will be more sensitive to systematic effects (such as galactic dust and synchrotron emission), the additional signal to noise alone should allow us to make maps of higher spatial resolution. The prediction of E and B mode CMB polarization at low ℓ from a 3D model potential is more involved compared to the temperature field. Initial simple Monte Carlo ray tracing experiments suggest that this may be an instructive way to capture the dependence of the polarized radiation field on the underlying potential – a challenge will be to capture this in a form that preserves our ability to use only the fast linear inversions described above.

Then, we will carry out an extensive systematic error analysis, assessing sources of contamination from the various foregrounds (using the Planck products as templates) and quantifying the robustness of our results to them. We expect all our inferences to increase in precision as a result of including the polarization information; whether we can reach the commensurate degree of accuracy will depend on both the computational and modeling problems outlined here.

Tree Representation and Non-Parametric Investigation of Gaussianity: There is another way to think about this problem. Let us build up the resolution of the potential on the CMB photosphere by increasing ℓ continuously from zero. Saddle points – designated S – accompanied by extrema – either maxima, designated H, or minima, designated, L – will be created. They will be accompanied by fresh separatrices – the contours that pass through the saddles. These come in two types - "lemniscates", like an infinity symbol, and designated X, and "limaçons", with the shape of a pinched annulus, and designated K [33]. New separatrices may be created between existing separatrices or out of a contour encircling a L or a H. Occasionally, the inverse process – annihilation – will occur. If we designate the total number of maxima, minima, saddles, lemniscates and limaçons by N_H, N_L, N_S, N_X, N_K, respectively, then clearly $N_S = N_X + N_K = N_H + N_L - 2$.[6]

The nesting of these contours defines a specific topology which suffices to describe all the equipotentials. It is convenient to represent it using a "tree" containing "forks" (corresponding to separatrices and labeled K or S), and branches terminating on "leaves" (corresponding to extrema and labeled H or L) [35]. There is only one "path" connecting any two leaves. Although our investigation has only begun, it is already clear that the statistics and structure of the tree, satisfy many rules if the underlying fluctuations are truly drawn from a random distribution, for example $N_K \sim N_X$ as $\ell \to \infty$. We propose to explore this novel approach to testing Gaussianity. We can actually extend this 2D approach on the photosphere to a 3D approach[7] exploring the nesting of the equipotential surfaces within the photosphere. This leads to similar type of tree. We propose to explore this as well with similar goals in mind.

[6] It is also instructive to calculate the Hessian matrix for the potential on the sphere and divide the sphere into "H-zones" (where both eigenvalues are negative), "L-zones" (where they are both positive) and "S-zones" (where they have opposite signs). These "curvature maps" [34] are complementary to the tree representation.

[7] Another generalization that we plan to explore is describe the topological arrangement of the polarization patterns [36].

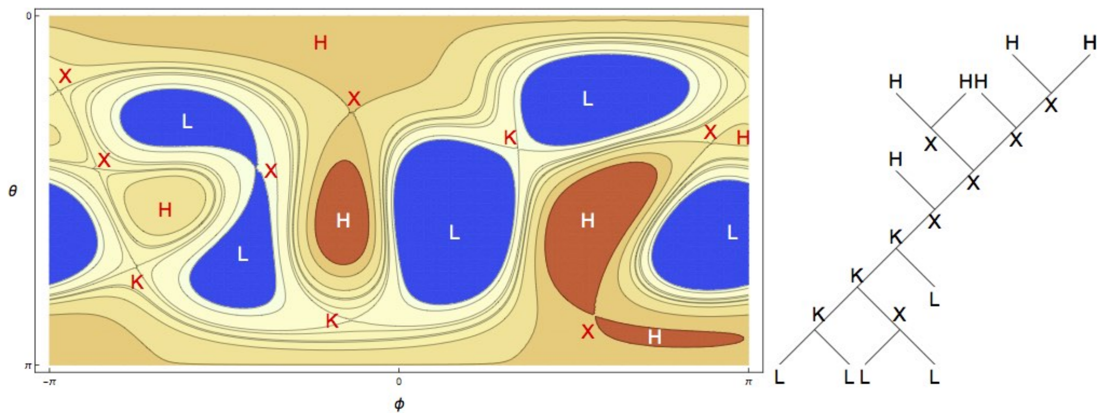


Figure 5: a) Nesting of separatrices for Planck data, with $\ell_{\max} = 4.5$. The extrema, H, L and saddles, K, X are identified. b) Equivalent tree describing the same data.

2.2 Stage 2. From 3D to 4D: Incorporating Existing "Local" Volumetric Data

At the end of Stage 1, we will have in hand the posterior PDF for the fourier mode applitudes of a 3D model of the potential on large scales throughout the observable Universe (and slightly beyond) at the epoch of last scattering. If we now assume a Friedman-Robertson-Walker expansion model, with suitably-parameterized expansion rate and a particular choice of those expansion parameters, we can evolve the fourier modes of any sample potential map we draw from this PDF forward in cosmic time in order to make posterior predictions about the "local" (i.e, $z << 1000$) Universe. The likelihood of these predictions given various observations in the local Universe will allow us to downweight some possible models, and hence further reduce the uncertainty in the map. As we discuss below, the process of combining CMB data with local surveys will not always be as straightforward as this simple picture of prediction and evaluation: Stage 2 of this program will involve investigation of how to carry out the joint inferences robustly and efficiently.

CMB Lensing and ISW Measurements: An important way to add 3D information on the potential throughout a significant fraction of the Universe's volume is to include the lensing of the CMB [12]. A uniform CMB is unchanged by gravitational lensing. However, if there is a gradient in the background temperature, intervening structure will appear as extra power on the scale of intervening large scale structure.[8] The consequences are largest on much smaller scales than those in which we are primarily interested. However there are still integral effects with $\ell \sim 30 - 100$ which are relevant. Furthermore the intense interest in the claim that inflationary B-modes have been detected [12] has focused much observational and analytical effort on this region of the spectrum. It is proposed to see if the addition of these measurements will improve the specification of the 3D body modes.

Note that in the case of CMB lensing, it is the same 3D potential model that will be predicting both the lensing effect on the CMB temperature map (due to structures

[8]More subtle manifestations including those involving polarization are possible [37], but this is the main effect.

at $z < 1000$), and also the intrinsic structure in this "background" temperature map (at $z \sim 1000$) itself. The consequence will be a non-linear system, which will need to be treated with care in the inference. The posterior PDF for the Fourier modes will no longer have a Gaussian form, but the weakness of the lensing effect may leave it to be close enough to Gaussian for a simple Gaussian approximation to give sufficient accuracy. This will be the starting point for our research in this area, which may develop into an exploration of better approximations to the posterior PDF for the Fourier modes which retain as much of the the computational efficiency as the linear model as possible.

Similar remarks apply to the Integrated Sachs Wolfe effect which is caused by variation in the potential over time, attributable to the cosmological constant (or a "dark energy" component) at late times. It is proposed to see if such measurements can also contribute to the specification of structure on the largest scales.

Galaxy Surveys and the "Local" Universe: Most of the use of galaxy surveys to date has been for drawing statistical inferences relating the growth of structure to the CMB emphasizing shorter length scales, notably those associated with BAO and the largest voids ~ 0.1 Gpc. However, these same surveys can also be used to augment the long wavelength CMB data and improve the accuracy and resolution of the resulting 3D potential map. A good example is the SDSS/BOSS program,[9] which covered nearly a third of the sky with over a million redshifts and photometry on galaxies out to $z \sim 0.7$.[10] For our purposes this translates to a comoving volume $\sim 50\text{Gpc}^3$, about 0.005 of the total. Surveys of much rarer quasars and the brightest star forming galaxies which extend to $z \sim 6$ provide much greater volumes over which the potential on Gpc scales can be estimated, albeit with inferior precision.

It is helpful at this point to consider a volume limited-survey of objects out to some radius r. Suppose we have a set of objects, (L^* galaxies, quasars, bright, star-forming galaxies ...) with space density n, and we want to measure the amplitude of a given Fourier component with wave vector k of the relative density perturbation associated with this potential $\delta \sim -2k^2\Phi/3a^2H^2$. Now, the accuracy with which the amplitude of a single relative density perturbation Fourier mode can be measured is comparable to the precision with which the fractional density perturbation can be measured in a single region of size equal to the associated length scale. This is $\sim k^{3/2}n^{-1/2}$, and must exceed δ. This suggests that the density of such objects must exceed $\sim H_0^4/c^2\Phi k_m ax$ if local surveys can possible connect with the CMB. A slightly more careful calculation indicates that making such a connection with existing survey and CMB data from stage 1 is just possible, and so it is worth exploring this further. If this is achievable, then although the data increment will be small, its value will be much greater because it can act as a *phase reference* for anchoring the imperfectly specified modes measured by CMB observations.

Further Constraints on Inflation: The newly developed effective field theory of large scale structure [38] has made great promises in terms of pushing constraints on the primordial Universe far beyond the scope of the CMB. This formalism relies on a separation of scales between the linear physics on large, cosmological scales in the infra-red (IR)

[9]http://www.sdss.org

[10]21 cm redshift surveys provide an important complement to optical surveys but the survey volumes to date are comparatively modest.

and the non-linear physics on small scales in the ultra-violet (UV) (under ~ 10 Mpc). These small scales are strongly coupled as a result of gravitational collapse. The UV modes are then integrated out in a way that yields a classical loop expansion correcting the dynamics of the linear IR theory. By pushing the theoretical control of modes with k closer to the non-linear regime, a far greater amount of information becomes available. The application of this method to current surveys will certainly allow the improvement of our 3D map in accuracy and resolution, which will in turn be translated into better constraints on the inflaton potential, primordial non-Gaussianities, etc. Quantifying the astrophysical systematic errors will be an important part of this phase of the investigation: the connection between observed galaxy summaries such as their number density and the underlying potential will need to be quantified and explored. Both galaxy bias and halo bias will need to be incorporated into the treatment.

2.3 Stage 3. Posterior Predictions and Ultimate Limits

The final stage of our proposed research program will be predictive in character. A fully constrained 4D model of the large scale gravitational potential in the Universe, with well characterized posterior PDF, will enable us to make predictions, with uncertainties, about the density field probed by a number of different upcoming sky surveys. Beyond the fundamental scientific importance of making, and so enabling the testing of, such predictions, our map will provide a new tool for the teams analyzing these future surveys. Systematic error control on these large scales is as yet uncharted territory: our 4D potential maps will enable the effective regularization of the analysis of new data on the largest angular scales, and thus potentially provide a higher contrast view of any anomalies present.

Over the next decade, a suite of all-sky (or at least, wide field) cosmological galaxy surveys are planned, including those to be carried out with LSST[11]), DESI[12], SphereX[13], Euclid[14], WFIRST-AFTA[15], and CHIME[16]. We will be able to make predictions about the large scale density of galaxies and clusters of galaxies, as well as the large scale tomographic weak lensing density fluctuations, that will be probed by these surveys.

There is a large effort underway to probe the Epoch of Reionization, (EoR) $6 \lesssim z \lesssim 30$ through hydrogen line measurements, with facilities such as the planned Square Kilometer Array[17] and its precursors. This is an exciting area of discovery, as the relevant physics depends upon many factors, notably the formation of the first stars and galaxies, that are very hard to anticipate. We will be able to make predictions about the large scale structure present in the volume at $6 \lesssim z \lesssim 30$ probed by these surveys. Again, such a prediction should assist in the interpretation of the survey data, and increase the fidelity of the new measurements made there.

It is also of interest to consider the limitations to what could be learned in principle

[11] http://www.lsst.org/lsst/
[12] http://desi.lbl.gov
[13] http://spherex.caltech.edu/
[14] http://www.euclid-ec.org
[15] http://wfirst.gsfc.nasa.gov
[16] http://chime.phas.ubc.ca
[17] https://www.skatelescope.org

about the idiosyncratic structure of our Universe with *any* conceivable observing facility. The many galaxy and EoR surveys referred to above, combined with CMB lensing and ISW measurement, should ultimately be able to give a quite detailed description of 4D potential. Exactly how detailed the map can be made is an interesting question to ask, and in doing so anticipate being led to new applications of our approach. Any explicit (not just statistical) linkage between large scale structure at recombination and today must strengthen investigations into basic physics questions including the properties of dynamical dark energy if it is present. Implicit in our approach is the opportunity to make statements about structure somewhat outside our horizon, predicated on our adopted inflationary model on these large scales. This raises interesting issues of theoretical principle which we intend to try to clarify.

3 Personnel Plans

This work will be primarily a collaboration between the PI and the collaborator Dr. Phil Marshall, who is currently a Staff Scientist at SLAC and one of the LSST Dark Energy Science analysis working group conveners. It is proposed to support a postdoctoral fellow, Laurence Perreault Levasseur to work on this project half time for two years. As we are conducting this research in open view, we expect to attract additional collaborators and several local colleagues have already expressed interest in it.

4 Responsibilities and Schedule

Blandford will start by taking the lead experimenting with the simulation of potential maps and then CMB temperature and polarization maps and finally using the publicly released Planck temperature and polarization data to do the best job we can on this data alone. Meanwhile, Marshall will lead the consideration of existing local surveys, including those listed above, and work on addressing the underlying Bayesian inference problems posed by combining them with CMB data. Perreault Levasseur is already working on the connection to inflation and intends to participate fully in developing the new statistical approaches as well as lead the Stage 4 discussion of the limits of the approach. Our goal, if funded, is to complete the baseline investigation described in this proposal by summer, 2018.

5 Broader Impact

The research program that we propose has a broad and popular interest analogous to the images of say Comet 67P.[18] As the map is essentially 3D, we will explore the use of 3D printing as well as sophisticated 2D movie representations to exhibit the results. This project also necessarily brings together many disparate research communities both

[18] If WMAP produced the "baby pictures of the Universe", then perhaps the goal here is to produce the corresponding adult movies.

astronomical and statistical. As a consequence, we are developing the statistical machinery for combining the various cosmological datasets in the open, via the GitHub web service at `http://github.com/rogerblandford/Music`, to enable and encourage broad participation.[19] If our approach is fruitful, we believe that it may be of value to other investigations. It will certainly help disseminate our 3D models and 2D posterior predictive distributions, in the interest of those working in other big dataset visualizations: interactive presentation of these products via IPython notebooks is ongoing, and can be supplemented by screenshared video recording to provide a narrative for a wider audience.

Blandford continues to give many public talks on black holes, high energy astrophysics and cosmology. He has also spent much of the past year completing a 1600+ pp text book, Modern Classical Physics, coauthored with Kip Thorne, which will appear in the spring, published by Princeton University Press. The 28th and final chapter of this book comprises a relatively new approach to presenting the results of modern cosmology to a graduate physics audience. The ideas discussed above arose out of the writing of this chapter.

[19] Program collaborator Marshall has worked in this way on other projects that lend themselves to this approach, most notably a recent Annual Reviews article on Ideas for Citizen Science in Astronomy [39], and the Space Warps citizen science project [40].

References

[1] N. Kaiser, H. Aussel, B. E. Burke, H. Boesgaard, K. Chambers, M. R. Chun, J. N. Heasley, K.-W. Hodapp, B. Hunt, R. Jedicke, D. Jewitt, R. Kudritzki, G. A. Luppino, M. Maberry, E. Magnier, D. G. Monet, P. M. Onaka, A. J. Pickles, P. H. H. Rhoads, T. Simon, A. Szalay, I. Szapudi, D. J. Tholen, J. L. Tonry, M. Waterson, and J. Wick, *Pan-STARRS: A Large Synoptic Survey Telescope Array*, in *Survey and Other Telescope Technologies and Discoveries* (J. A. Tyson and S. Wolff, eds.), vol. 4836 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, pp. 154–164, Dec., 2002.

[2] Ž. Ivezić, K. Menou, G. R. Knapp, M. A. Strauss, R. H. Lupton, D. E. Vanden Berk, G. T. Richards, C. Tremonti, M. A. Weinstein, S. Anderson, N. A. Bahcall, R. H. Becker, M. Bernardi, M. Blanton, D. Eisenstein, X. Fan, D. Finkbeiner, K. Finlator, J. Frieman, J. E. Gunn, P. B. Hall, R. S. J. Kim, A. Kinkhabwala, V. K. Narayanan, C. M. Rockosi, D. Schlegel, D. P. Schneider, I. Strateva, M. SubbaRao, A. R. Thakar, W. Voges, R. L. White, B. Yanny, J. Brinkmann, M. Doi, M. Fukugita, G. S. Hennessy, J. A. Munn, R. C. Nichol, and D. G. York, *Optical and Radio Properties of Extragalactic Sources Observed by the FIRST Survey and the Sloan Digital Sky Survey*, *Ap. J.* **124** (Nov., 2002) 2364–2400, [astro-ph/0202408].

[3] M. Davis, S. M. Faber, J. Newman, A. C. Phillips, R. S. Ellis, C. C. Steidel, C. Conselice, A. L. Coil, D. P. Finkbeiner, D. C. Koo, P. Guhathakurta, B. Weiner, R. Schiavon, C. Willmer, N. Kaiser, G. A. Luppino, G. Wirth, A. Connolly, P. Eisenhardt, M. Cooper, and B. Gerke, *Science Objectives and Early Results of the DEEP2 Redshift Survey*, in *Discoveries and Research Prospects from 6- to 10-Meter-Class Telescopes II* (P. Guhathakurta, ed.), vol. 4834 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, pp. 161–172, Feb., 2003. astro-ph/0209419.

[4] M. Giavalisco, H. C. Ferguson, A. M. Koekemoer, M. Dickinson, D. M. Alexander, F. E. Bauer, J. Bergeron, C. Biagetti, W. N. Brandt, S. Casertano, C. Cesarsky, E. Chatzichristou, C. Conselice, S. Cristiani, L. Da Costa, T. Dahlen, D. de Mello, P. Eisenhardt, T. Erben, S. M. Fall, C. Fassnacht, R. Fosbury, A. Fruchter, J. P. Gardner, N. Grogin, R. N. Hook, A. E. Hornschemeier, R. Idzi, S. Jogee, C. Kretchmer, V. Laidler, K. S. Lee, M. Livio, R. Lucas, P. Madau, B. Mobasher, L. A. Moustakas, M. Nonino, P. Padovani, C. Papovich, Y. Park, S. Ravindranath, A. Renzini, M. Richardson, A. Riess, P. Rosati, M. Schirmer, E. Schreier, R. S. Somerville, H. Spinrad, D. Stern, M. Stiavelli, L. Strolger, C. M. Urry, B. Vandame, R. Williams, and C. Wolf, *The Great Observatories Origins Deep Survey: Initial Results from Optical and Near-Infrared Imaging*, *Ap. J. Lett.* **600** (Jan., 2004) L93–L98, [astro-ph/0309105].

[5] The Dark Energy Survey Collaboration, *The Dark Energy Survey*, *ArXiv Astrophysics e-prints* (Oct., 2005) [astro-ph/0510346].

[6] S. M. Faber, C. N. A. Willmer, C. Wolf, D. C. Koo, B. J. Weiner, J. A. Newman, M. Im, A. L. Coil, C. Conroy, M. C. Cooper, M. Davis, D. P. Finkbeiner, B. F. Gerke, K. Gebhardt, E. J. Groth, P. Guhathakurta, J. Harker, N. Kaiser, S. Kassin, M. Kleinheinrich, N. P. Konidaris, R. G. Kron, L. Lin, G. Luppino, D. S. Madgwick, K. Meisenheimer, K. G. Noeske, A. C. Phillips, V. L. Sarajedini, R. P. Schiavon, L. Simard, A. S. Szalay, N. P. Vogt, and R. Yan, *Galaxy Luminosity Functions to z˜1 from DEEP2 and COMBO-17: Implications for Red Galaxy Formation*, *Ap. J.* **665** (Aug., 2007) 265–294, [astro-ph/0506044].

[7] N. Scoville, H. Aussel, M. Brusa, P. Capak, C. M. Carollo, M. Elvis, M. Giavalisco, L. Guzzo, G. Hasinger, C. Impey, J.-P. Kneib, O. LeFevre, S. J. Lilly, B. Mobasher, A. Renzini, R. M. Rich, D. B. Sanders, E. Schinnerer, D. Schminovich, P. Shopbell, Y. Taniguchi, and N. D. Tyson, *The Cosmic Evolution Survey (COSMOS): Overview*, *Ap. J. Supp.* **172** (Sep., 2007) 1–8, [astro-ph/0612305].

[8] N. Kaiser, W. Burgett, K. Chambers, L. Denneau, J. Heasley, R. Jedicke, E. Magnier, J. Morgan, P. Onaka, and J. Tonry, *The Pan-STARRS wide-field optical/NIR imaging survey*, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, vol. 7733 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, p. 0, July, 2010.

[9] C. Blake, E. A. Kazin, F. Beutler, T. M. Davis, D. Parkinson, S. Brough, M. Colless, C. Contreras, W. Couch, S. Croom, D. Croton, M. J. Drinkwater, K. Forster, D. Gilbank, M. Gladders, K. Glazebrook, B. Jelliffe, R. J. Jurek, I.-H. Li, B. Madore, D. C. Martin, K. Pimbblet, G. B. Poole, M. Pracy, R. Sharp, E. Wisnioski, D. Woods, T. K. Wyder, and H. K. C. Yee, *The WiggleZ Dark Energy Survey: mapping the distance-redshift relation with baryon acoustic oscillations*, *MNRAS* **418** (Dec., 2011) 1707–1724, [arXiv:1108.2635].

[10] S. Alam, F. D. Albareti, C. Allende Prieto, F. Anders, S. F. Anderson, T. Anderton, B. H. Andrews, E. Armengaud, É. Aubourg, S. Bailey, and et al., *The Eleventh and Twelfth Data Releases of the Sloan Digital Sky Survey: Final Data from SDSS-III*, *Ap. J. Sup. Series* **219** (July, 2015) 12, [arXiv:1501.0096].

[11] **Planck** Collaboration, R. Adam et al., *Planck 2015 results. IX. Diffuse component separation: CMB maps*, arXiv:1502.0595.

[12] **Planck** Collaboration, P. A. R. Ade et al., *Planck 2015 results. XIII. Cosmological parameters*, arXiv:1502.0158.

[13] S. Weinberg, *Cosmology*. Cosmology. OUP Oxford, 2008.

[14] P. Schneider, *Extragalactic astronomy and cosmology: An introduction*. SpringerLink : Bücher. Springer Berlin Heidelberg, 2014.

[15] R. K. Sachs and A. M. Wolfe, *Perturbations of a Cosmological Model and Angular Variations of the Microwave Background*, *Ap. J* **147** (Jan., 1967) 73.

[16] E. R. Harrison, *Fluctuations at the threshold of classical cosmology*, *Phys. Rev.* **D 1** (May, 1970) 2726–2730.

[17] Y. B. Zel'dovich, *A hypothesis, unifying the structure and the entropy of the universe*, *MNRAS* **160** (Mar., 1972) 1–3.

[18] A. H. Guth, *Inflationary universe: A possible solution to the horizon and flatness problems*, *Phys. Rev.* **D 23** (Jan., 1981) 347–356.

[19] A. D. Linde, *A new inflationary universe scenario: A possible solution of the horizon, flatness, homogeneity, isotropy and primordial monopole problems*, *Phys. Lett.* **B 108** (Feb., 1982) 389–393.

[20] V. F. Mukhanov and G. Chibisov, *Quantum fluctuation and nonsingular Universe. (In Russian)*, *JETP Lett.* **33** (1981) 532–535.

[21] K. Sato, *First-order phase transition of a vacuum and the expansion of the Universe*, *MNRAS* **195** (May, 1981) 467–479.

[22] S. Hawking, *The development of irregularities in a single bubble inflationary universe*, *Phys. Lett.* **B 115** (1982), no. 4 295 – 297.

[23] A. A. Starobinsky, *Dynamics of phase transition in the new inflationary Universe scenario and generation of perturbations*, *Phys. Lett.* **B 117** (1982) 175–178.

[24] A. Albrecht and P. J. Steinhardt, *Cosmology for grand unified theories with radiatively induced symmetry breaking*, *Phys. Rev. Lett.* **48** (Apr., 1982) 1220–1223.

[25] J. M. Bardeen, P. J. Steinhardt, and M. S. Turner, *Spontaneous Creation of Almost Scale-Free Density Perturbations in an Inflationary Universe*, *Phys. Rev.* **D 28** (1983) 679.

[26] K. A. Malik and D. Wands, *Cosmological perturbations*, *Phys. Rept.* **475** (2009) 1–51, [`arXiv:0809.4944`].

[27] C. Gordon, D. Wands, B. A. Bassett, and R. Maartens, *Adiabatic and entropy perturbations from inflation*, *Phys. Rev.* **D 63** (2001) 023506, [`astro-ph/0009131`].

[28] N. Birrell and P. Davies, *Quantum fields in curved space.* Cambridge monographs on mathematical physics. Cambridge University Press, 1984.

[29] **Planck** Collaboration, N. Aghanim et al., *Planck 2015 results. XI. CMB power spectra, likelihoods, and robustness of parameters*, *Submitted to: Astron. Astrophys.* (2015) [`arXiv:1507.0270`].

[30] **Planck** Collaboration, P. A. R. Ade et al., *Planck 2015 results. XVII. Constraints on primordial non-Gaussianity*, `arXiv:1502.0159`.

[31] S. H. Suyu, P. J. Marshall, M. P. Hobson, and R. D. Blandford, *A bayesian analysis of regularised source inversions in gravitational lensing*, *MNRAS* **371** (2006) 983–998, [astro-ph/0601493].

[32] A. P. Yadav and B. D. Wandelt, *CMB tomography: Reconstruction of adiabatic primordial scalar potential using temperature and polarization maps*, *Phys. Rev. D* **71** (June, 2005) 123004, [astro-ph/0505386].

[33] R. Blandford and R. Narayan, *Fermat's principle, caustics, and the classification of gravitational lens images*, *ApJ* **310** (Nov., 1986) 568–582.

[34] J. F. Nye, *Optical Caustics in the Near Field from Liquid Drops*, *Proceedings of the Royal Society of London Series A* **361** (May, 1978) 21–41.

[35] D. B. West et al., *Introduction to graph theory*, vol. 2. Prentice hall Upper Saddle River, 2001.

[36] P. A. G. Scheuer, J. H. Hannay, and P. J. Hargrave, *A note on the interpretation of polarization maps*, *MNRAS* **180** (July, 1977) 163–168.

[37] W. Hu and T. Okamoto, *Mass Reconstruction with Cosmic Microwave Background Polarization*, *ApJ* **574** (Aug., 2002) 566–574, [astro-ph/0111606].

[38] J. J. M. Carrasco, M. P. Hertzberg, and L. Senatore, *The effective field theory of cosmological large scale structures*, *JHEP* **9** (Sep., 2012) 82, [arXiv:1206.2926].

[39] P. J. Marshall, C. J. Lintott, and L. N. Fletcher, *Ideas for Citizen Science in Astronomy*, *Ann. Rev. Astron. Astrophys.* **53** (Aug., 2015) 247–278, [arXiv:1409.4291].

[40] P. J. Marshall, A. Verma, A. More, C. P. Davis, S. More, A. Kapadia, M. Parrish, C. Snyder, J. Wilcox, E. Baeten, C. Macmillan, C. Cornen, M. Baumer, E. Simpson, C. J. Lintott, D. Miller, E. Paget, R. Simpson, A. M. Smith, R. Küng, P. Saha, T. E. Collett, and M. Tecza, *Space Warps: I. Crowd-sourcing the Discovery of Gravitational Lenses*, *ArXiv e-prints* (Apr., 2015) [arXiv:1504.0614].