PROJECT DESCRIPTION

1 Introduction

1.1 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 [**?** **?** **?** **?** **?** **?** **?** **?** **?** **?** **?**] and a splendid two dimensional map of the microwave background [**?**] and this has led to a standard 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 that is 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.

This is 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 investigations with additional priors which should tighten up their accuracy.

1.2 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. [**?** **?**]). 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 [**?**]. There is still room for revision, retractions and new discoveries but, right now, we have a good working hypothesis that the universe is basically this simple (e.g. [**?** **?**]). There is some tension in the reported measurements of the Hubble constant (68 km s^{-1} Mpc^{-1}) etc 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 simple changes will be needed to what follows.

1.3 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, $a \sim 0.0009$, 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 came to dominate over the matter, and the universal expansion started to accelerate.

1.4 Fluctuations

The observed fluctuations are conventionally 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 Φ,[1] from which the density and fluid velocity perturbations can be simply computed (in the "Sachs-Wolfe" limit at long wavelength [**?**]). The radiation and neutrinos have to be treated kinetically and a second scalar curvature potential needs to be introduced for accurate calculations. There may also be tensor modes which we shall ignore here. If, and when, they are detected, only minor modifications will be needed. 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 [**?**], that the adiabatic hypothesis is quite accurate. Furthermore, the amplitudes associated with each mode of the initial potential scale as $k^{-3/2}$ and are drawn from a Gaussian distribution, so that the potential fluctuations associated with each length scale are scale-independent.[2] This behavior is consistent with a remarkable early conjecture by Harrison [**?**] as elaborated by Zel'dovich [**?**].

The potential associated with a mode was essentially frozen until it "entered the horizon," that is, until the timescale for its dynamical evolution, a/k, became smaller than the Hubble timescale (which can be thought of as the estimated age of the Universe if the expansion rate were constant). We are mostly concerned with wavelength modes for which this happened before recombination and $k < k_0 \sim 40\mathrm{Gpc}^{-1}$.[3] 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.

1.5 Inflation

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 [**?**], Linde [**?**] and others (e.g. [**? ? ? ? ? ?**]), that the universe underwent a period of so-called "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 their slightly tilted power spectrum (see e.g. [**? ?**] and references therein).

[1] It is simplest to work in the "Newtonian" gauge throughout.

[2] Actually there is a small tilt in their power spectrum in the sense of there being slightly larger amplitudes at longer wavelengths. This requires a small correction to what follows and is addressed in Section 2.2.5.

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

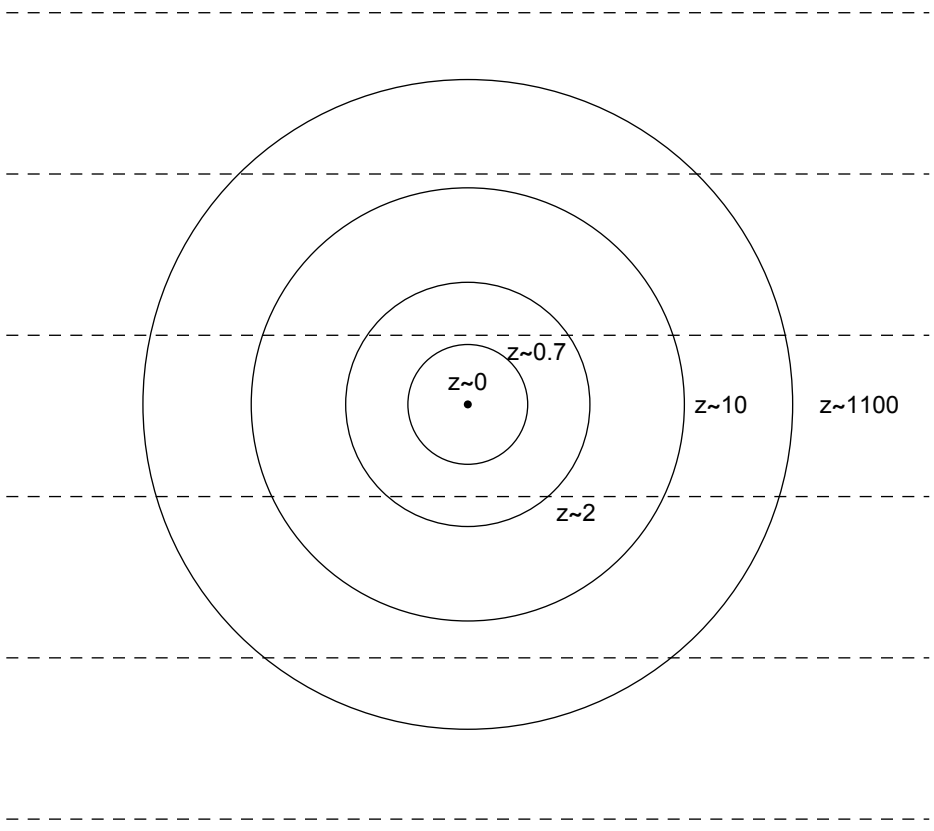

Figure 1: The local universe interior to the CMB photosphere expressed in comoving coordinates. The circles are, in order, redshift $z \sim 0.7$, schematically the limit of present surveys, $z \sim 2$ roughly the effective limit of future surveys, the nominal Epoch of Reionization at $z \sim 10$ and the CMB photosphere at $z \sim 1100$ and a distance $x_{\rm CMB} \sim 14$ Gpc. Also shown as dashed lines are the nodes of a single wave mode with $k \sim 0.4$ Gpc^{-1} which contributes significantly to spherical harmonics up to $\ell \sim 6$. The universe that is directly observable by us (for example using long wavelength gravitational waves) extends only slightly beyond the CMB photosphere. WILL REDRAW AND SHRINK.

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", remains (quasi) constant. Meanwhile, quantum fluctuations in the matter field(s) and metric are constantly generated with wavelength H^{-1} at most [**?**]. 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, a/k, becomes larger than the Hubble time: it "exits the horizon" and its amplitude freezes. 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.

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

2 Proposed Research

2.1 Long-term Goal

The long term goal which this proposal addresses, is to connect the CMB to local surveys, and to produce an evolving three-dimensional (or in other words, a 4D) 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 the secure physical inferences that have been drawn about the early universe in making this map.

The first stage of our proposed program uses 2D Cosmic Microwave Background (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 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 stage involves estimating the improvement that should come on a decade timescale from future surveys such as LSST and Epoch of Reionization investigations, "Stage IV" CMB observations, SKA etc. The fourth and final stage is an investigation of how far it is possible to reconstruct the structure of the universe including what can be inferred beyond our horizon.

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

2.2.1 Temperature Fluctuations

Most quantitative cosmology derives from CMB observations. 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, $Y_y(\theta,\phi) = \{Y_{0,0}, Y_{1,0}, 2^{1/2}\Re[Y_{1,1}], 2^{1/2}\Im[Y_{1,1}], Y_{2,0}, \ldots, 2^{1/2}\Im[Y_{\ell_{\rm max},\ell_{\rm max}}]\}$ of length $(\ell_{\rm max}+1)^2$ 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 $\ell_{\rm max}$ as a continuous variable by adding a fraction between zero and unity of the largest ℓ shell and therefore smoothly change the angular resolution.

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.[4] In the proposed study, only one specific realization of the

[4]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

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 us with 100 sample Planck temperature fluctuation maps for $0 \leq \ell \leq 10$ or $1 \leq y \leq 121$. From this ensemble we are able to compute the mean photospheric 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.

2.2.2 Fourier Modes

It is conventional to Fourier expand the potential Φ at a given time, usually the present; 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 cosmic photosphere, 13.9 Gpc[5] 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 real and $\mathbf{k} = \Delta k\{n_1, n_2, n_3\}$, 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:

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

2.2.3 Gaussian Prior

Detailed study of the CMB (e.g. [**?** **?**]) 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 $\sigma_{\mathbf{n}}^2 = \alpha(n_1^2+n_2^2+n_3^2)^{-3/2}$. We fix α 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 comparing with the measured a_ys. This approach is adequate for our pilot study; an improved determination will be investigated following the evidence analysis of [**?**], *en route* to a hierarchical modeling of the system where α is one of a set of hyperparameters governing the statistics of the potential field (Section 2.2.5).

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.

[5]The comoving radius of the big bang is 14.2 Gpc.

Figure 3: Preliminary recovery of the potential Φ at the time of inflation from the CMB data alone. Fourier coefficients up to $n_{\rm max}$ =??? are used, and the posterior mode located. A higher resolution map should be possible using a more sophisticated inversion process and including polarization data.

2.2.4 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.*[6] At first sight this might seem hopeless, because if one associates ℓ with $k\,x_{\rm CMB}$, then we are trying to solve for $O((k_{\rm max}\,x_{\rm CMB})^3)$ Fourier modes using only $O(\ell^2_{max})$ spherical harmonics. However, if we confine our attention to the longest wavelength waves,[7] 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 multivariate Gaussian distribution) for the coefficients f_n by first finding its peak, minimizing the quantity

$$-2\ln\mathcal{P} = (a_y - \mathbf{R}_{yn}f_n)^{\rm T}C^{-1}_{yy'}(a_{y'} - \mathbf{R}_{y'n'}f'_n) + \frac{f_n^2}{\sigma^2_{\mathbf{n}}} \tag{4}$$

with respect to variation of f_n.

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

$$f_n = \left(\frac{\delta_{nn'}}{\sigma^2_{(n)}} + \mathbf{R}^{\rm T}_{yn}C^{-1}_{yy'}\mathbf{R}_{y'n'}\right)^{-1}\mathbf{R}^{\rm T}_{yn'}C^{-1}_{yy'}a_{y'} \tag{5}$$

We have tested this procedure using mock data and found that it is generally possible to solve for stable, low order maximum posterior Fourier coefficients which can then be used to recover the original harmonic coefficients a_y. We have also followed this procedure on the actual Planck data, and our results are exhibited in Fig. 3. 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.

2.2.5 Inflationary Origin of Perturbations and Hyperparameter Estimation

As explained in Section 1.5, 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. Appart from these caveats, the framework developed in what follows will remain model independent.

[6] This is not the original meaning of the word.

[7] In so doing, are approach is quite complementary to that of Yadav and Wandelt (2005) who are concerned with wavelengths comparable with the thickness of the recombination surface around the acoustic peak near $\ell \sim 200$.

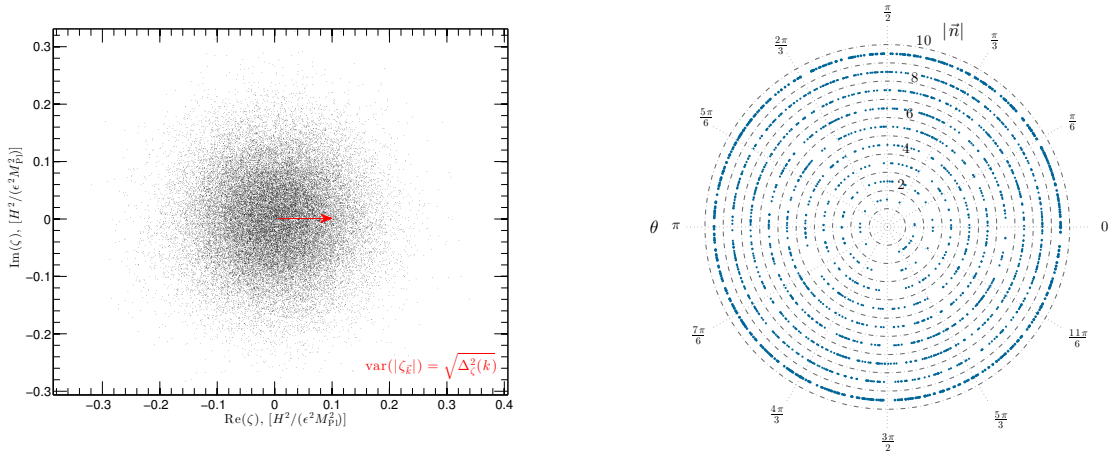

Figure 4: *Left panel*: Distribution of the modes of $\zeta_{\mathbf{k}}$ at fixed $|\mathbf{k}|$. *Right panel*: Representation of the norm of the first few k-modes ($|\mathbf{n}|$ from 2 to 10), versus their random spatial angle θ. If the individual modes were clustering around an angle, this would 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^2_ζ would not be affected.

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 curvature perturbation, ζ.[8] In the co-moving gauge, the Fourier modes of ζ as they exit the horizon are

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

Here, ν is a slowly varying function which encapsulates the specific dynamics of the inflationary model, $\mathcal{H}^{(1)}_{\nu}$ 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 experiments have strived to measure the power spectrum of fluctuations, Δ^2_ζ. 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 the reconstruction method proposed in this proposal, such measurement will become possible.

One on the main uncertainty 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 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 so-called 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

[8] 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 interchangeably. 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 [**?**].

$\vec{\eta}$ given the data a_y, via the potential map f_n (which we marginalize over), $\mathcal{P}(\vec{\eta}|a_y)$,

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

where $\mathcal{P}(\vec{\eta})$ is the prior on $\vec{\eta}$, $C^{-1}_{n'n}$ is a diagonal matrix with diagonal elements given by $1/|\zeta_{\mathbf{k}}|^2$ or, equivalently, $1/\sigma^2_{\mathbf{n}}$ as defined in Section 2.2.3,[9] and the $\mathbf{A}$ matrix and $\mathbf{B}$ vector are given by:

$$\mathbf{A}_{nn'} = \mathbf{R}^{\mathrm{T}}_{y'n}C^{-1}_{y'y}\mathbf{R}_{yn'} + C^{-1}_{nn'}\,, \qquad \mathbf{B}^{\mathrm{T}}_n = a^{\mathrm{T}}_{y'}C^{-1}_{y'y}\mathbf{R}_{yn}\,. \quad (9)$$

2.2.6 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. 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 a viable 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.

PJM: ROGER TO ADD FEW SENTENCES DETAIL ON PHYSICS BEING PROBED, AND RAY-TRACING EXPTS?

Then, we will carry out an extensive systematic error analysis, introducing various 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.

PJM: SUFFICIENT ON SYSTEMATICS? WHAT OTHER GAINS/PITFALLS DO WE EXPECT?

2.2.7 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 cosmic 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 (cf. Blandford and Narayan 1986). 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 $N_S = N_X + N_K = N_H + N_L - 2$.[10]

[9] $C^{-1}_{n'n}$ therefore corresponds to the covariance matrix of the Gaussian prior from Section 2.2.3 for fixed $\vec{\eta}$ values

[10] 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" (cf. Nye, J. F. 1978 Proc. Roy. Soc. A361 21) are complementary to the tree representation.

Figure 5: a) Nesting of separatrices on the cosmic photosphere for the Planck data, with ℓ =???, on the cosmic photosphere punctured at the absolute minimum potential. The extrema, H, L and saddles, K, X are identified. Also shown are the L (yellow), S (green) and H (blue) zones. b) Equivalent tree describing the same data as in a). The ordinates of the forks and leaves corresponds to the actual potential values.

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) (e.g West, D. B. Introduction to Graph Theory1996 London: Prentice-Hall). 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.[11] The equipotential contours become surfaces and we can describe their nesting by a very similar type of tree. We propose to explore this as well with similar goals in mind.

2.3 Stage 2. From 3D to 4D: Incorporating Existing Volumetric Data

2.3.1 CMB Lensing and ISW Measurements

An important way to add 3D information is to include CMB lensing (Ade et al. 2014b). 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.[12] 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 (Ade et. 2014c) 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.

Similar remarks apply to the Integrated Sachs Wolfe effect which is caused by changes in the potential due to the cosmological constant at late times. It is proposed to see if such measurements can also contribute to the specification of large scale structure although here the challenge seems even greater.

2.3.2 Galaxy Surveys and the Local Universe

Most of the use of surveys 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 map. A good example is the SDSS/BOSS program http://www.sdss.org which covered nearly a third

[11] Another generalization that we plan to explore is describe the topological arrangement of the polarization patterns (e.g. Scheuer, Hannay and Hargrave 1977).

[12] More subtle manifestations including those involving polarization are possible (Hu & Okamoto 2002), but this is the main effect.

of the sky with over a million redshifts and photometry on galaxies out to $z \sim 0.7$.[13] For our purposes this translates to a comoving volume $\sim 50\mathrm{Gpc}^3$, 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 but 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 precision with which the amplitude of a single relative density perturbation Fourier mode can be measured is comparable with 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.

2.3.3 Constraints on Inflation

The newly developed effective field theory of large scale structure [**?**] has made great promises in terms of pushing constraints on the primordial Universe far beyond what the CMB is theoretically capable of. This formalism relies on a separation of scales between the linear physics on large, cosmological scales in the infra-red (IR) 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 future application of this method to current and upcoming 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.

2.4 Stage 3. Future Surveys

Many facilities that are being planned or built will provide data relevant to our approach. In this phase of our program we will use simulated data for a selection of these future surveys to make forecasts of how our maps should improve, and look for new opportunities to exploit them.

2.4.1 Ground-based CMB Telescopes

While there are exciting proposals for future space-based CMB measurements, most attention is currently focused on the next two generations – Stages 3 and 4 – of ground-based CMB telescopes, which have been proposed to deliver results in very roughly five and ten years respectively. While these are mostly focused on probing the physics of inflation and cosmological neutrinos [**?** **?**], they will also improve the measurement of CMB temperature and E-mode

[13] 21 cm redshift surveys provide an important complement to optical surveys but the survey volumes to date are comparatively modest.

polarization on the scales $\ell \lesssim 200$ in which we are primarily interested, with signal to noise $\sim 3 \times 10^{-4}$, roughly ten times higher than Planck.

2.4.2 Survey Telescopes

Construction has begun on the Large Synoptic Survey Telescope (LSST[14]), which will commence a decade-long survey in 2022. It will survey half the sky from Chile (with consequently very strong overlap with the ground-based CMB observations) in six optical and near infrared bands, detecting about ten billion galaxies out to $z \sim 2$ for L^* galaxies and $z \sim 6$ for bright, star-forming galaxies and quasars [**?**]. Its primary cosmological goal is to perform a multi-probe joint analysis (involving weak lensing, galaxy clustering, type Ia supernovae, time delay lenses and cluster number counts) to see if dark energy (and not just a cosmological constant) is responsible for cosmic acceleration [**?**]. However, it will, in practise, contribute to many more important cosmological tests. LSST will only provide photometric redshifts, but these will be well-calibrated by large spectroscopic surveys including that of DESI[15], which is projected to measure ~ 20 million redshifts to $z \sim 1$ starting in 2018. A complementary source of new local data will be the Euclid space mission[16], which is scheduled for a 2020 launch and which will carry out weak lensing, baryon acoustic oscillation and redshift space distortion measurements using 1.5 billion galaxies and 50 million redshifts over more than a third of the sky [**?**]. The proposed WFIRST-AFTA http://wfirst.gsfc.nasa.gov also has an impressive program in observational cosmology [**?**], that will build on the LSST and Euclid surveys [**?**]. Meanwhile, at radio wavelengths, the Canadian High Intensity Mapping Experiment (CHIME[17], see e.g. [**?**]) will measure BAO out to $z \sim 2.5$ over half the sky.

As with the CMB, weak lensing and galaxy clustering observations can provide "tomographic" distance information and, in principle, should lead to a better map of the long wavelength potential perturbations. Systematic error control on these large scales is as yet uncharted territory: our 4D potential maps may have a role to play in this, effectively regularizing the local analysis on the largest angular scales. At the same time, local constraints on the 3D potential will pin down the model consdierably, improving our constraints on the inflation model.

2.4.3 Epoch of Reionization

There is a large effort underway to probe the Epoch of Reionization, (EoR) $6 \lesssim z \lesssim 30$ through hydrogen line measurements. This is an exciting area of discovery, as the relevant physics depends upon many factors, notably first star formation and galaxy assembly that are very hard to anticipate.[18] The experiments will probe an ideal range of comoving radius $\sim 8 - 12$ Gpc, interpolating between the CMB photosphere and local surveys, for either contributing to or expanding upon our incorporating our 3D potential map.

On a longer time scale there are ambitious plans to construct an international Square Kilometer Array (SKA[19]). The long term goals include measuring the redshifts of a billion galaxies, performing weak lensing surveys and carrying out more sensitive surveys of the epoch of reionization (see [**?**] and references therein). It is likely that the SKA capabilities and schedule will

[14] http://www.lsst.org/lsst/
[15] http://desi.lbl.gov
[16] http://www.euclid-ec.org
[17] http://chime.phas.ubc.ca
[18] JWST, http://www.jwst.nasa.gov scheduled for launch in 2018, will also help indirectly in understanding the universe during this epoch but seems unlikely to provide quantitative measurements of very large scale structure.
[19] https://www.skatelescope.org

become better-defined over the lifetime of this proposed research program.

With a 4D potential model constrained both at $z = 1100$ by the CMB and at $z = 0.5$ by the Dark Energy surveys of the previous subsection, we will be able to make predictions about the large scale structure present in the volume at $6 \lesssim z \lesssim 30$ probed by EoR surveys. Such a prediction should assist in the interpretation of the survey data, and increase the fidelity of the measurements made there.

2.5 Stage 4. Limits

It is of interest to consider the limitations to what could be learned in principle about the idiosyncratic structure of our universe with *any* conceivable observing facility (cf. Leclerc et al. 2014). Galaxy and quasar survey combined with EoR studies an CMB lensing should be able to give a detailed description of 4D potential though the resolution during the dark ages will be poor relative to that earlier and later epochs. Any explicit (not just statistical) linkage between large scale structure at recombination and today must strengthen investigations into basic physics questions including the presence of dynamical dark energy instead of a cosmological constant. Implicit in this approach is the opportunity to make statements about structure somewhat outside our horizon predicated on our adoption of the Gaussian prior 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 co-Investigator Dr. Phil Marshall who is currently a Project Scientist at SLAC and chair of the LSST Dark Energy Science Collaboration Council. 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. Marshall will also handle the inclusion of future surveys under stage 3, 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

PHIL AND LAURENCE TO ADD MORE ON PUBLIC OUTREACH

The research program that we propose has an broad and popular interest analogous to the images of say Comet 67P.[20] As the map is essentially 3D, we will explore the use of 3D printing as well a sophisticated 2D movie representations to exhibit the results. This project also necessarily brings together many disparate research communities both astronomical and statistical. As a consequence, we intend to develop the statistical machinery for combining the various cosmological datasets on an open website, to enable and encourage broad participation.[21] If our approach is fruitful, we believe that it may be of value to other investigations and we hope that this device will help disseminate our 3D models and 2D posterior predictive distributions which among those working in other big dataset visualizations.

Marshall has a long and ongoing interest and record in public outreach including various initiatives in citizen science, public talks and large outreach events. Blandford continues to give many public talks on black holes, high energy astrophysics and cosmology. He has also spent much of the past year computing 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.

6 Results from AST08-07458. PI: R. Blandford

THIS HAS TO BE PART OF THE 15 PAGES AND I WILL SHORTEN IT

This proposal was to carry out research on gravitational lensing. Some principal results will be summarized here.

6.1 Strong Lensing

6.1.1 Time delays

Blandford and Hilbert, lately joined by Hezaveh, as members of a large collaboration led by former student Suyu and also including KIPAC member Marshall, refined the measurement of the Hubble constant and other cosmological parameters using the gravitational lens B1608+656. They successfully tested and implemented new formalisms to include the perturbative effects of intervening galaxies and large scale structure upon the time delays. They have also used the Millennium simulation to devise a procedure for estimating the statistical distribution of the mean convergence along the line of sight, given a catalog of observations of nearby galaxies. This improves the accuracy with which lensing determinations of the Hubble constant can be achieved with imminent observations. Recent research has been extended to RXJ1131-1231, where the effect of extrinsic convergence should be lower than for B1608+656. This has already furnished a seven percent distance solution. More generally, the convergence distributions are also important for galaxy counts or halo models based upon photometry. There has been recent tension between Planck-based determinations of the Hubble constant and lensing on most other approaches. Some progress has been made on clarifying the practical strengths and limitations of lens approaches.

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

[21] 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 (Marshall et al. 2015).

6.1.2 Galactic structure and Initial Mass Function

Barnabe continued to refine the CAULDRON code that he developed for his thesis that combines stellar dynamics with gravitational lensing. He then applied it to observations made under the Sloan Lens ACS Survey (SLACS). The main galaxies of study were early type and it was possible to study their IMF, especially their low mass cut offs and to deduce that there had been little evolution since z 0.35. Barnabe has also been a leader of the Sloan WFC Edge on Late type Lens Survey (SWELLS). This is a major project to model edge on spirals that act as gravitational lenses. The goal is to obtain a better understanding of the dynamical structure of spiral galaxies and, especially, to study the stellar distribution and to infer the initial mass functions associated with the different components of the galaxy the disk, bulge and the halo. Over several publications it has been concluded that there is no universal IMF and that

6.2 ALMA

Hezaveh and collaborators have demonstrated that ALMA observations of strong gravitational lenses can be used to provide a gravitational measurement of the power spectrum of dark matter subhalos. Observations will be taken shortly which have the potential to validate the particulate nature of dark matter. Hezaveh has proposed that high redshift galactic nuclei that are strongly lensed may have their gas kinematics well enough resolved to furnish reliable black holes masses. This technique will soon be practical with ALMA and eventually with GSMTs. A paper has been published.

6.3 Cosmic Shear

6.3.1 Weak lensing surveys

Hilbert and others used shear correlations from the Deep Lens Survey to derive constraints on the cosmic mean matter density and the amplitude of matter fluctuations. In particular, Hilbert contributed cosmic shear data covariance matrices. The first full analysis has been published and a tomographic extension is in preparation. Hilbert and colleagues developed a new method for detecting shear peaks in weak lensing surveys and studying their abundance and spatial correlation using simulations, They showed how to constrain more effectively cosmological parameters from weak lensing surveys when this analysis was combined with complementary measurements. This approach is currently being applied ot Looking to the future, Hilbert has simulated the cosmological constraints one could obtain from an LSST or a Euclid weak lensing survey by combining shear correlations, aperture mass statistics, and shear peak counts. He demonstrated that competitive constraints on cosmological parameters, including non-Gaussianity would be possible. Such observations might thus provide valuable constraints on models of inflation and the physics of the very early Universe. Schrabback, who has been designated as the deputy lead for Euclid weak lensing shape measurements, has studied the calibration and the influence of color gradients in galaxies on the accuracy

6.3.2 Improving photometric redshifts

Schrabback and others have re-analyzed the CFHT Legacy Survey as part of the CFHT Lens Survey. They have paid particular attention to performing more accurate photometry and used the results to improve the determination of photometric redshifts. The results of this exercise are mixed with the most positive outcome being an improved PSF procedure which may lead ultimately to improved redshift determinations.

6.3.3 Galaxy shapes, intrinsic alignment, and weak lensing

Hilbert and colleagues used simulations by Springel of cosmic structure to make a more direct assessment of the impact of intrinsic alignment on cosmic shear surveys. Two papers have been published and a third is in preparation. Hilbert, Schrabback and colleagues then investigated the possible contamination of weak lensing measurements by intrinsic alignments of galaxy shapes. They quantified the shape distribution of various galaxy samples in the COSMOS survey and compared the results to predictions from N-body simulations finding that simple models of galaxy shapes in the literature fail to reproduce the observed shape distribution. Currently, they are considering a broader set of semi-analytic models models to understand this discrepancy. Schrabback and colleagues have made a study of the influence of color gradients on the shapes of faint galaxies used for gravitational lensing investigations of dark energy. They find that these effects can be significant but are correctable and give prescriptions for keeping the errors that they may engender below statistical errors.

6.4 Galaxy-Galaxy Lensing

6.4.1 Light-matter correlation

The CFHT Legacy Survey-Wide is the most powerful data set for weak lensing measurements currently existing. It comprises 170 square degrees of sky imaged in five bands. Schrabback is a core member of the team which is conducting a thorough analysis of the complete dataset, which is now corrected for systematic bias and includes all the available photometric redshift information. The primary goals of the survey are to explore the relationship between luminous and dark matter and to place significant constraints on cosmological parameters. He has also completed an analysis of the flattening of halos. Hilbert et al. used N-body simulations of structure formation along with semi-analytic galaxy-formation models and ray-tracing to show that higher-order galaxy-galaxy lensing correlations can be used to provide new information about galaxy formation. An analysis of galaxy-galaxy lensing in the CFHTLS yielded a number of unexpected results for the relation between matter and galaxies which may be a consequence of systematic error. A paper on lensing 2- and 3-point correlations has been published and another paper is in preparation.

6.4.2 Evolution of the ratio of stellar to dark mass in galaxies

Schrabback and colleagues have used data from the COSMOS field to perform a new investigation of how the light and mass from galaxies varies with redshift out to z=1. They are able to measure the downsizing at modest redshift and discuss some mechanisms involving AGN feedback and disk instabilities which may be responsible.

6.5 Clusters

6.5.1 Weak lensing observations of rich clusters of galaxies

Schrabback has also been devising new techniques for galaxy shape measurements in weak lensing studies of rich clusters of galaxies. The goal is to increase the number density of background galaxies that can be included. These techniques have been successfully applied mainly to HST-ACS observations. In particular they are being used by von der Linden in the analysis of MACS 0717.5+3745, which also includes wide-field imaging from the ground. The pipeline now includes the merging system MACSJ0417.5-1154 as well as several clusters from the HST Multi-Cycle

Treasury Program CLASH, and discovered by the South Pole Telescope. Studies of the cosmological evolution of clusters are limited by the small number of high redshift clusters suitable for comboined strong and weak lensing analysis. In order to increase the sample, Schrabback is combining the ROSAT X-ray Survey and the Sloan Digital Sky Survey samples. This involves optical imaging with the WHT and LBT telescopes, and radio (Sunyaev-Zel'dovich) observations with the CARMA (SZA) Array. Chandra and HST observations have also been proposed. Von der Linden has led a large team including Blandford on an ambitious project to carry out a systematic weak lensing analysis of a sample of 52 rich clusters. The project involved developing new accurate photometric procedures and new algorithms for measuring cluster masses In ongoing work, this sample is being used to propose new cluster scaling relationships and to furnish new constraints on cosmological parameters .

6.5.2 Detecting tidal stripping of halos using weak lensing

A new approach to demonstrating the stripping of dark matter from the halos around distant group using weak gravitational lensing has been proposed. Its implementation was heavily simulated using the Millennium Simulation.

6.5.3 Chandra observations of the largest quasar lens

Blandford and Schrabback have participated in a investigation led by former postdoc Oguri to analyze Chandra observations of the largest separation, triply-imaged quasar, SDSS J 1029+2623. The X-ray observations from the associated z=0.58 rich cluster exhibit a subpeak suggestive of a merger and a mass consistent with that obtained from gravitational lensing. The consistency of the magnifications suggested that microlensing is not important in this system. However, their ratios point to the presence of dark matter substructure. A multi-telescope study of this massive X-ray cluster has been performed, finding additional multiple images, measuring a time delay and performing a weak lensing analysis. The lens model is able to account for the salient features of these observations, deriving a magnification of a background quasar of 30. There is a serious discrepancy with the X-ray cluster mass estimate which is attributed to shock heating in a merger.

6.6 Miscellaneous

6.6.1 Cosmic strings

Morganson, Marshall, Schrabback and Blandford continue to explore ways to improve the lensing limit on the incidence of cosmic strings.

6.6.2 Hubble workshop

Suyu, Blandford, Freedman and Treu, organized a highly successful workshop on the measurement of the Hubble constant. Many competing approaches were contrasted and compared.

6.6.3 Angular Correlation Function

Blandford is collaborating with Morganson and Schrabback on the measurement of cosmic shear using asymmetry in the angular correlation function on arc second angular scales following earlier work by Morganson and Blandford. The theory has been developed and an attempt is underway to seek this effect in existing data.

6.6.4 Microlensing studies of nomad planets

Blandford, Barnabe and Marshall have collaborated with former postdoc Strigari on an investigation of the use of microlensing observations, from ground and space, to enhance our statistical understanding of the frequency of planetary companions to main sequence stars in mass - semi major axis - eccentricity space in the light of the results from the Kepler space mission. They have gone on to assess the prospects of detecting interstellar planets and dwarf planets using microlensing. They argued that there could be as many as a hundred thousand of these nomads per star with masses larger than that of Pluto. In collaboration with Porter, they considered γ-ray limits on the incidence of nomads and began a study the probabilities of their carrying microbial life between stars (Panspermia). This last has led to a discussion of several experiments designed to quantify the viability of microorganisms under interstellar and re-entry conditions.