Preprint typeset in JHEP style - HYPER VERSION LAPTH-1350/09

Features in the primordial power spectrum? A frequentist analysis

Jan Hamann
Department of Physics and Astronomy, University of Aarhus
8000 Århus C, Denmark
and
LAPTh, Université de Savoie, CNRS
BP 110, 74941 Annecy-le-Vieux Cedex, France
E-mail: hamann@phys.au.dk

Arman Shafieloo
Department of Physics, University of Oxford
1 Keble Road, Oxford, OX1 3NP, UK
E-mail: a.shafieloo1@physics.ox.ac.uk

Tarun Souradeep
IUCAA
Post Bag 4, Ganeshkhind, Pune 411 007, India
E-mail: tarun@iucaa.ernet.in

Abstract: Features in the primordial power spectrum have been suggested as an explanation for glitches in the angular power spectrum of temperature anisotropies measured by the WMAP satellite. However, these glitches might just as well be artifacts of noise or cosmic variance. Using the effective $\Delta\chi^2$ between the best-fit power-law spectrum and a deconvolved primordial spectrum as a measure of "featureness" of the data, we perform a full Monte-Carlo analysis to address the question of how significant the recovered features are. We find that in 26% of the simulated data sets the reconstructed spectrum yields a greater improvement in the likelihood than for the actually observed data. While features cannot be categorically ruled out by this analysis, and the possibility remains that simple theoretical models which predict some of the observed features might stand up to rigorous statistical testing, our results suggest that WMAP data are consistent with the assumption of a featureless power-law primordial spectrum.

Keywords: CMBR theory, cosmological parameters from CMBR, initial conditions and eternal universe.

1. Introduction

The advent of a multitude of cosmological precision data in the past decade has led to the emergence of the so-called concordance, or "vanilla" model of cosmology. A key part of the vanilla model is the assumption that the spectrum of primordial curvature perturbations is smooth and featureless, and can be described by a simple power-law – consistent with their origin from an earlier period of slow-roll inflation [1–4].

Nonetheless, the inflationary mechanism also allows for more complex shapes of the spectrum, caused for instance by non-standard initial conditions [5–8], or unusual dynamics of the inflaton field due to, e.g., a phase transition [9, 10], non-smoothness of the inflaton potential [11–13] or particle production [14, 15] during inflation (see also [16–22]). In any case, any detection of features in the spectrum, i.e., a deviation from the standard power-law behaviour, would yield invaluable clues on the physics of the early Universe. Presently, the most powerful single source of information about the primordial state of perturbations are the observations of the temperature anisotropies of the Cosmic Microwave Background (CMB) by the WMAP satellite [23–25]. Since the first data release, a number of efforts have been undertaken to ascertain the compatibility of the data with the power-law paradigm.

In a top-down approach, specific non-smooth models of the primordial spectrum or the inflaton potential have been fit to the data [26–39]. Alternatively, the bottom-up approach of trying to reconstruct the shape of the primordial spectrum from these data has been employed, involving for example binning of the primordial spectrum [40–42], principal component analysis [43], or a direct reconstruction via deconvolution methods [44–53].

Generally, these results have shown that by introducing suitable features in the primordial spectrum, the fit to the data can be improved by $\Delta\chi^2_{\rm eff} \sim \mathcal{O}(10)$ over the power-law fit. Taken by itself, however, $\Delta\chi^2_{\rm eff}$ does not really help answer the crucial question whether the better fit indicates a real feature in the primordial spectrum or whether it just stems from over-fitting the scatter in the data from noise and cosmic variance. Indeed, attempts at interpreting results of parameter inference in terms of 'ruling out the power-law model in favour of a feature model' are likely to be prone to grossly overestimating the 'evidence' for features. A simplistic goodness-of-fit χ^2/dof analysis would not be very enlightening either, due to the large number of degrees of freedom involved. One might thus be tempted to turn to the formalism of Bayesian model selection [54, 55]. Though aside of the potential technical problems of evaluating Bayes factors for feature models with a large number of free parameters, the interpretation of results from a Bayesian model selection analysis may be problematic [56], owing to the absence of a unique well-motivated choice of priors for these mostly empirical parameterisations.

We believe that for this particular problem the toolbox of frequentist statistics provides a useful implement, and propose to perform an analysis based on the tech-

nique of hypothesis testing. Given the null hypothesis of an underlying power-law spectrum, this involves the evaluation of a suitably chosen statistic on a large number of Monte Carlo simulated mock CMB temperature anisotropy data sets. From the resulting frequency distribution of this statistic one can derive a p-value, i.e., the probability that, given the null hypothesis, the value of the statistic is larger than the one observed. We note that in defining the statistic one has to beware of a posteriori interpretations of the data; a particular feature observed in the real data may be very unlikely (and lead to a low p-value; see e.g., [57, 58]), but the probability of observing some feature may be quite large.

In Section 2 (and in the Appendix) we describe in detail the methods used in our analysis, the results of which are presented in Section 3. We discuss and interpret our findings in Section 4.

2. Method

2.1 Null hypothesis

We assume as our null hypothesis that the temperature angular power spectrum inferred from WMAP can be explained by the present cosmological standard ("vanilla") model. In particular, the underlying primordial spectrum of curvature perturbations is taken to have a smooth power-law form

$$\mathcal{P}_{\mathcal{R}}(k) = A_{\rm S}\ (k/k_*)^{n_{\rm S}-1}\,, \tag{2.1}$$

determined by the two parameters $n_{\rm S}$, the scalar spectral index, and $A_{\rm S}$, the amplitude of fluctuations at the pivot scale $k_* = 0.05$ Mpc^{-1}. The remaining four free parameters of the vanilla model are the baryon and cold dark matter densities $\omega_{\rm b}$ and $\omega_{\rm c}$, the ratio of sound horizon to angular diameter distance at decoupling θ, and the reionisation optical depth τ.

The fiducial spectrum $\mathcal{C}_\ell^{\rm fid}$ used to generate random realisations of WMAP5 data should be chosen as the maximum likelihood spectrum. However, since temperature data alone are not very sensitive to τ, using temperature data alone leads to a value of τ which is at odds with results of a combined temperature/polarisation analysis. We therefore fix τ to the best-fit result for the full WMAP5 likelihood function, and then determine the values of the other five parameters by fitting the thus reduced five-parameter vanilla model to the temperature data only (using the likelihood function described in Appendix A.1).

The resulting fiducial spectrum is defined by the following parameter values: $\omega_{\rm b} = 0.0224$, $\omega_{\rm c} = 0.109$, $\theta = 1.04$, $\tau = 0.089$, $A_{\rm S} = 2.147 \cdot 10^{-9}$ and $n_{\rm S} = 0.963$.

2.1.1 Alternative hypothesis

The null hypothesis is to be tested against an alternative hypothesis. Our alternative hypothesis is that the primordial spectrum is not given by a power-law, but possesses

features of some sort. It now remains to define a suitable test statistic that can be used to assess the validity of the null hypothesis.

2.2 Statistic

By relaxing the assumption of a precise power-law form of the primordial spectrum one can generally achieve a better fit to the data. A commonly used measure of this improvement is the effective delta-chi-squared

$$\Delta\chi^2_{\rm eff} = -2\ln\mathcal{L}^{\rm V}_{\rm max} + 2\ln\mathcal{L}^{X}_{\rm max}, \tag{2.2}$$

where $\mathcal{L}^{\rm V}_{\rm max}$ is the maximum likelihood of a fit to the vanilla model, and $\mathcal{L}^{\rm X}_{\rm max}$ the maximum likelihood of a fit to a model with an alternative shape of the spectrum.

In the present work we wish to allow a very general form of the primordial spectrum, so we focus on a reconstruction of the primordial spectrum from a binned version of the temperature angular power spectrum with the aid of a modified Richardson-Lucy (RL) deconvolution algorithm as performed in Refs. [45, 48] (see Appendix A.2). Given a transfer function $T_\ell(k)$ and an observed or simulated angular power spectrum $\mathcal{C}_\ell$, this algorithm results in an "optimised" primordial spectrum $\mathcal{P}^{\rm RL}_{\mathcal{R}}(k)$, and is a very powerful tool to find potential features in the spectrum. We take $\mathcal{L}^{\rm X}_{\rm max}$ to be the likelihood of the data for the primordial spectrum $\mathcal{P}^{\rm RL}_{\mathcal{R}}(k)$ combined with the best-fit transfer function of the vanilla model.

This method does have two limitations though: as long as the transfer function is held fixed, one is only sensitive to features that are not degenerate with the cosmological parameters that determine the transfer function (e.g., typically sharp features such as spikes or oscillations, but not broad distortions). It would be possible to circumvent this problem by also optimising the transfer function, but such a procedure would become computationally prohibitively expensive in the context of our analysis. Also, since we deconvolve binned angular power spectrum data, we are not sensitive to extremely high-frequency modulations of the primordial spectrum as reported in Ref. [39]. Nevertheless, the $\Delta\chi^2_{\rm eff}$ obtained with our method provides a good measure of the potential deviation from smoothness for realistic primordial spectra.

2.3 Numerical implementation

Applied to real data, the calculation of $\Delta\chi^2_{\rm eff}$ involves a maximisation of the likelihood using the reduced (i.e., $\tau = 0.089$) vanilla model, and the application of the RL-deconvolution algorithm with the best-fit parameter transfer function.

Naturally, the simulated data sets should be treated in exactly the same way as the real data. Schematically, the procedure to be performed for each random realisation of WMAP data should hence read as follows:

(i) Generate a random realisation ${}^{(i)}\mathcal{C}^{\rm sim}_\ell$ of simulated WMAP temperature data.

(ii) Find minimal $\chi^2_{\rm eff,V}$ for a fit of the vanilla model to ${}^{(i)}\mathcal{C}_\ell^{\rm sim}$.

(iii) Calculate the transfer function ${}^{(i)}T_\ell(k)$ for the best-fit parameters found in step (ii).

(iv) Apply the modified Richardson-Lucy deconvolution algorithm to ${}^{(i)}\mathcal{C}_\ell^{\rm sim}$, using the transfer function ${}^{(i)}T_\ell(k)$.

(v) Determine $\chi^2_{\rm eff,RL}$ from the deconvolved primordial spectrum and calculate the improvement $\Delta\chi^2_{\rm eff}$ over the power-law fit.

To keep the task of simulating mock WMAP data at the level of generating angular power spectra ${}^{(i)}\mathcal{C}_\ell^{\rm sim}$ and avoid having to generate random maps, we adopt a slightly simplified form of the WMAP likelihood function (described in Appendix A.1) throughout this paper. For details of simulating mock WMAP data (step (i)) we refer the reader to Appendix A.3.

The minimisation of step (ii) turns out to be the computationally most expensive part of the algorithm. Deterministic, simplex-like minimisation routines like Numerical Recipes' `amoeba` [59] are not very reliable for our purpose since they run the risk of getting stuck in local minima. Random-walk based routines, such as simulated annealing are more suitable here. We choose a combination of the `amebsa` [59] routine to find a good starting point for a subsequent low-temperature run of the Metropolis-Hastings algorithm (based on the `CosmoMC` code [60]), which results in an estimate of the minimal $\chi^2_{\rm eff,V}$ that is accurate to about 0.1. Nonetheless, this algorithm requires several thousand evaluations of CMB angular power spectra and the corresponding likelihoods, which would take a considerable time if one were to use a conventional Boltzmann code, such as `CAMB` [61]. We resort instead to the interpolation code `PICO` [62, 63] for calculating the $\mathcal{C}_\ell$s, which can considerably speed up this task, returning a reliable estimate for the minimal $\chi^2_{\rm eff,V}$ within a few minutes on a regular desktop CPU.

Having found a good estimate of the maximum likelihood point, we then use `CAMB` to calculate the transfer function for the best-fit cosmological parameters in step (iii). Having applied the RL-deconvolution (step (iv)), giving us the optimised spectrum $\mathcal{P}_\mathcal{R}^{\rm RL}(k)$, we can finally compute the improvement $\Delta\chi^2_{\rm eff}$ the optimised spectrum would yield over the best power-law one (step (v)).

3. Results

3.1 Real data

We commence by applying the RL-deconvolution method to the observed WMAP5 temperature data. The resulting best fit "optimised" primordial spectrum is shown in Fig. 1; its dominant features are qualitatively similar to those found in Ref. [48] for

the 3-year WMAP data: a cutoff-like suppression at the largest scales accompanied by two prominent wiggles around $k \sim 0.002$ Mpc^{-1} and $k \sim 0.07$ Mpc^{-1}. The optimised spectrum gives $\chi^2_{\rm eff,RL} = 1029.45$, which marks an improvement of $\Delta\chi^2 = 24.41$ over the power-law best fit of $\chi^2_{\rm eff,V} = 1053.86$.

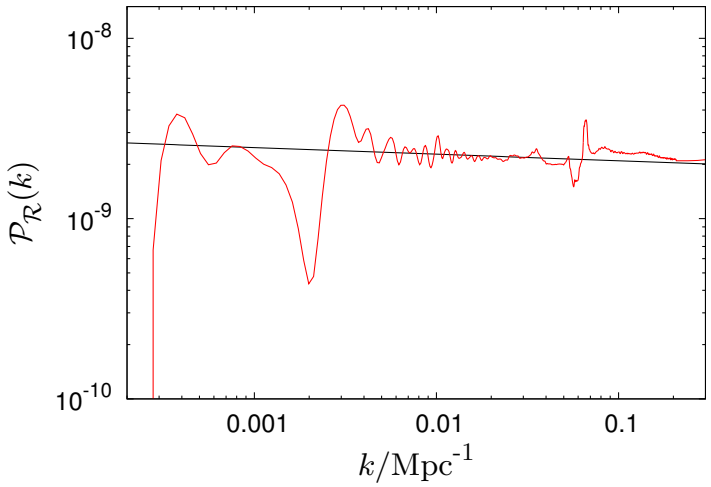

Figure 1: Best-fit power-law (black line), and RL-optimised primordial spectrum (red line).

3.2 Simulated data

We have generated 2000 random realisations of simulated WMAP5 temperature data for multipoles $2 \leq \ell \leq 1000$ and applied the procedure of Section 2.3 to determine the $\Delta\chi^2_{\rm eff}$ for each of them. A histogram of the distribution of $\Delta\chi^2_{\rm eff}$ values is shown in Figure 2. Since the RL-algorithm is applied to binned data whereas $\Delta\chi^2_{\rm eff}$ is defined using un-binned data, for some simulated spectra the resulting $\Delta\chi^2_{\rm eff}$ can be negative, even after wavelet smoothing. In other words, in these cases the RL-optimised spectra yield no improvement over a power-law, and we set $\Delta\chi^2_{\rm eff} = 0$. This has occurred in about 5% of our simulated spectra and is marked by the red part of the lowest-$\Delta\chi^2_{\rm eff}$-bin in the histogram.

The vertical blue line in Figure 2 marks the observed value $\Delta\chi^2_{\rm eff} = 24.41$. Of the 2000 mock data sets, 525 have $\Delta\chi^2_{\rm eff} > 24.41$, corresponding to a p-value of ~ 0.26. To ensure this number is not affected by our simplified likelihood-approximation, we calculated the observed $\Delta\chi^2_{\rm eff}$ using the full WMAP5 likelihood code (version `v3p2`), and found a difference of ~ 1.

Additionally, we considered statistics which are less susceptible to low multipoles where the likelihood approximation is most inaccurate: $\Delta\chi^2_{\rm eff}$ evaluated between $\ell_{\rm min}$ and $\ell_{\rm max} = 1000$ with $\ell_{\rm min} = 10,\ 20$, and 30. The resulting p-values are 0.33, 0.14, and 0.31, respectively. This behaviour can be understood from the features observed

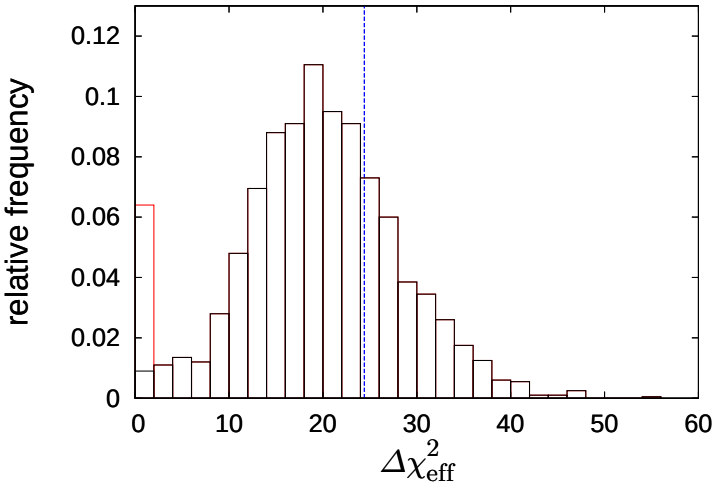

Figure 2: Histogram showing the relative frequency of $\Delta\chi^2_{\rm eff}$ for 2000 simulated realisations of the WMAP5 temperature spectrum. For about 5% of the spectra, the RL-algorithm does not result in a better fit; these cases are displayed in red. The dotted blue line denotes the observed value $\Delta\chi^2_{\rm eff} = 24.41$.

in the real data: removing $\ell < 10$ gets rid of a feature, leading to an increase in the p-value; taking away $10 \leq \ell < 20$ removes a part of the data which does not show features, causing the p-value to drop, only to go up again when one removes the feature at $20 \leq \ell < 30$.

In summary, the improvement in the fit from using an "optimised" primordial spectrum over a power-law spectrum is larger than the observed one in roughly 30% of our simulated data sets which assume an underlying smooth spectrum. We thus conclude that the null hypothesis cannot be rejected at high significance within the limitations of our chosen statistic, and the data are compatible with a smooth power-law primordial spectrum. We illustrate this qualitatively in Fig. 3, which shows that the reconstructed spectrum from the observed data does not have an unusual amount of "featureness" compared to the simulated data.

4. Conclusions

We have addressed the question whether the 5-year WMAP temperature anisotropy data are compatible with the assumption of an underlying smooth power-law primordial spectrum of curvature perturbations, or whether they show any indication for features or other unaccounted-for systematic effects. Assuming the true underlying primordial spectrum to be of power-law form, we generated 2000 simulated WMAP angular temperature power spectra, and estimated the amount of features by looking at the improvement to the likelihood gained from fitting an "optimised" primordial spectrum, obtained by deconvolving the data, instead of the usual power-

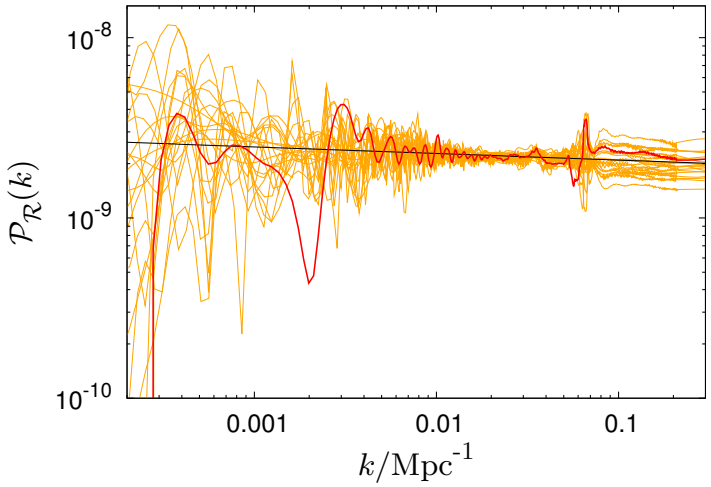

Figure 3: Reconstructed primordial spectra for twenty mock data sets (orange lines). For comparison we show the primordial spectrum reconstructed from WMAP5 data (red line) and our fiducial power-law spectrum (black line). It is evident that the spectrum reconstructed from real data does not have an unusual amount of features. The apparent feature at 0.05 Mpc^{-1} < k < 0.07 Mpc^{-1} is caused by the noise term becoming dominant at the corresponding multipoles in the binned WMAP data. For the last bin, $951 \leq \ell \leq 1000$, the averaged transfer function peaks at $k \sim 0.07$ Mpc^{-1}. Beyond that the reconstructed primordial spectrum is no longer dominated by the peak of one data bin, but rather by the tails of the transfer functions of the last few bins; more structure here does not further improve the fit.

law. We found that 26% of all simulated data sets show a greater improvement than the one observed, which leads us to conclude that the features seen in the WMAP 5-year temperature data are not at all incompatible with the assumption of a smooth primordial spectrum.[1]

However, we emphasise that our analysis does not disprove the existence of features – they are merely not strictly required by present data. Specifically, 'simple' theoretical models that predict similar features could conceivably remain favoured over a power-law. Additionally, since the spectrum reconstruction method used here is not sensitive to features below the binning scale of the WMAP data, we cannot rule out the possibility of extremely high-frequency features. We have demonstrated the feasibility of a full-scale frequentist hypothesis testing analysis in the search for deviations from smooth spectra though, and our method can just as well be applied to other statistics which allow for even more general shapes of the primordial spectrum.

Future analyses and new measurements may well reveal evidence for the existence

[1]While this work was being completed, Ref. [65] appeared, which reaches similar conclusions from cross-validating a spline-reconstruction of the primordial spectrum from various cosmological data sets.

of features, and a search for them will certainly remain a worthwhile endeavour, considering that any detection would profoundly impact our understanding of the physics of inflation and also precise estimation of the cosmological parameters [64].

Besides the imminent improvements to temperature anisotropy data from the Planck satellite [66], better measurements of the E-polarisation of the CMB will greatly enhance sensitivity to primordial features [52, 67] – a dedicated mission like CMBPol [68] may prove particularly helpful here.

Acknowledgments

We thank Pedro Ferreira, Subir Sarkar and Yvonne Wong for interesting discussions and comments. JH thanks the Oxford Physics group for their hospitality during the initial stages of this work. Numerical work was performed on the MUST cluster at LAPP (CNRS & Université de Savoie). JH acknowledges the support of a Feodor Lynen-fellowship of the Alexander von Humboldt foundation. AS acknowledges the support of the European Research and Training Network MRTPNCT-2006 035863-1 (UniverseNet).

A. Appendix: Technical details

A.1 Likelihood function

For reasons of simplicity we adopt here a form of the likelihood function similar to the one used by the WMAP team in their first data release [69]. For an input theoretical angular power spectrum $\mathcal{C}_\ell^{\rm th}$, the likelihood of the data reads

$$\chi^2_{\rm eff} \equiv -2\ln\mathcal{L} = -2\left(\frac{1}{3}\ \ln\mathcal{L}_{\rm Gauss} + \frac{2}{3}\ \ln\mathcal{L}'_{\rm LN}\right), \tag{A.1}$$

with a Gaussian part

$$-2\ln\mathcal{L}_{\rm Gauss} = \sum_{\ell\ell'}(\mathcal{C}_\ell^{\rm th} - \hat{\mathcal{C}}_\ell)Q_{\ell\ell'}(\mathcal{C}_{\ell'}^{\rm th} - \hat{\mathcal{C}}_{\ell'}), \tag{A.2}$$

and a log-normal part

$$-2\ln\mathcal{L}'_{\rm LN} = \sum_{\ell\ell'}(z_\ell^{\rm th} - \hat{z}_\ell)(\mathcal{C}_\ell^{\rm th} + \mathcal{N}_\ell)Q_{\ell\ell'}(\mathcal{C}_{\ell'}^{\rm th} + \mathcal{N}_{\ell'})(z_{\ell'}^{\rm th} - \hat{z}_{\ell'}), \tag{A.3}$$

where $z_\ell^{\rm th} = \ln(\mathcal{C}_\ell^{\rm th} + \mathcal{N}_\ell)$ and $\hat{z}_\ell = \ln(\hat{\mathcal{C}}_\ell + \mathcal{N}_\ell)$, with $\hat{\mathcal{C}}_\ell$ the angular power spectrum estimated from observation. The curvature matrix $Q_{\ell\ell'}$ is given by

$$Q_{\ell\ell'} = D_\ell^{-1}\delta_{\ell\ell'} + \frac{r_{\ell\ell'}}{\sqrt{D_\ell D_{\ell'}}}, \tag{A.4}$$

with

$$D_\ell = 2\,\frac{(\mathcal{C}_\ell^{\rm th} + \mathcal{N}_\ell)^2}{(2\ell+1) f_{\rm sky}^2}. \tag{A.5}$$

The off-diagonal terms induced by the sky cut $r_{\ell\ell'}$, the effective noise power spectrum $\mathcal{N}_\ell$ and the effective sky fraction $f_{\rm sky}$ are supplied by the WMAP team [23] and available for download on the LAMBDA web site[2].

A.2 RL-algorithm for reconstructing the primordial spectrum

The Richardson-Lucy (RL) algorithm was developed and is widely used in the context of image reconstruction in astronomy [70, 71]. However, the method has also been successfully used in cosmology, to deproject the 3-D correlation function and power spectrum from the measured 2-D angular correlation and 2-D power spectrum [72, 73].

The angular power spectrum, $\mathcal{C}_\ell$, is a convolution of the initial power spectrum $\mathcal{P}_\mathcal{R}(k)$ generated in the early universe, with a radiative transfer kernel $T_\ell(k)$ (the transfer function), that is determined by the values of the cosmological parameters. In our application, we solve the inverse problem of determining the primordial power spectrum, $\mathcal{P}_\mathcal{R}(k)$, from the measured angular power spectrum, $\mathcal{C}_\ell$, using the relation

$$\mathcal{C}_\ell \;= \int {\rm d}k\; T_\ell(k) \mathcal{P}_\mathcal{R}(k) \simeq \sum_i T_\ell(k_i)\, \mathcal{P}_\mathcal{R}(k_i). \tag{A.6}$$

In the above equation, the *target* measured angular power spectrum, $\mathcal{C}_\ell \equiv \mathcal{C}_\ell^{\rm obs}$, is the data given by observations, and the radiative transport kernel,

$$T_\ell(k_i) = \frac{\Delta k_i}{k_i} \left|\Delta_{T\ell}(k_i, \eta_0)\right|^2, \tag{A.7}$$

encodes the response of the present multipoles of the CMB perturbed photon distribution function $\Delta_{T\ell}(k_i, \eta_0)$ to unit of power per logarithm interval of wavenumber, k, in the primordial perturbation spectrum. The kernel $T_\ell(k)$ is completely fixed by the cosmological parameters of the *base* cosmological model. The kernel $T_\ell(k)$ also includes the effect of geometrical projection from the three-dimensional wavenumber, k, to the harmonic multipole, l on the two-dimensional sphere.

Obtaining $\mathcal{P}_\mathcal{R}(k)$ from the measured $\mathcal{C}_\ell$ for a given $T_\ell(k)$ is clearly a deconvolution problem. An important feature of the problem is that $\mathcal{C}_\ell^{\rm obs}$, $T_\ell(k)$ and $\mathcal{P}_\mathcal{R}(k)$ are all positive definite. We employ an improved RL method to solve the inverse problem for $\mathcal{P}_\mathcal{R}(k)$ in Eq. (A.6). The advantage of the RL method is that positivity of the recovered $\mathcal{P}_\mathcal{R}(k)$ is automatically ensured, given positive definite $T_\ell(k)$ and $\mathcal{C}_\ell$s. The RL method, readily derived from elementary probability theory of distributions [70], is an iterative method that can be neatly encoded into a simple recurrence relation. The power spectrum ${\mathcal{P}_\mathcal{R}}^{(i+1)}(k)$ recovered after iteration $(i+1)$ is given by

[2] `http://lambda.gsfc.nasa.gov/`

$$\mathcal{P_R}^{(i+1)}(k) - \mathcal{P_R}^{(i)}(k) \;=\; \mathcal{P_R}^{(i)}(k) \sum_\ell \tilde{T}_\ell(k)\, \zeta_k \, \frac{\tilde{\mathcal{C}}_\ell^{\rm obs} - \mathcal{C}_\ell^{(i)}}{\mathcal{C}_\ell^{(i)}} \;\; \tanh^2 \left[\frac{(\tilde{\mathcal{C}}_\ell^{\rm obs} - \mathcal{C}_\ell^{(i)})^2}{\tilde{\sigma}_l^2} \right], \tag{A.8}$$

where $\tilde{T}_\ell(k)$ is the normalised kernel (on the ℓ space for all k wavenumbers), $\tilde{\mathcal{C}}_\ell^{\rm obs}$ is the normalised measured data (target) and $\mathcal{C}_\ell^{(i)}$ is the angular power spectrum at the $i^{\rm th}$ iteration obtained from $\mathcal{C}_\ell^{(i)} = \sum \tilde{T}_\ell(k)\, \mathcal{P_R}^{(i)}(k)$ using the recovered power spectrum $\mathcal{P_R}^{(i)}(k)$. It is important to keep in mind that, due to the formulation in terms of conditional probability distributions, the RL method requires the kernel $T_\ell(k)$, data $\mathcal{C}_\ell$, and the target vector $\mathcal{P_R}(k)$, all to be normalised at the beginning,

$$\sum_\ell \tilde{\mathcal{C}}_\ell \;=\; 1; \quad \sum_k \tilde{\mathcal{P}_\mathcal{R}}^{(1)}(k) \;=\; 1; \quad \sum_\ell \tilde{T}_\ell(k) \;=\; 1, \tag{A.9}$$

where $\tilde{\mathcal{P}_\mathcal{R}}^{(1)}(k)$ is the normalised initial guess model of the primordial spectrum and the normalisation factor ζ_k is defined by $\zeta_k = \sum_\ell T_\ell(k)$.

Note that Eq.(A.8) represents a modified form of the Richardson-Lucy algorithm specifically tailored for our purpose in this problem [45, 47, 48].

Following Refs. [45, 47, 48], we apply the RL-algorithm to a *binned* version of the (real and simulated) WMAP5 data. The resulting raw deconvolved spectrum has spurious oscillations and features on scales smaller than the bin size, arising largely due to the k space sampling and binning in ℓ space, which adversely impact the (unbinned) full likelihood of the data given the reconstructed spectrum. As shown in Ref. [47], a subsequent smoothing of the raw reconstructed spectrum with a discrete wavelet transform can lead to a significant improvement in the likelihood.

A.2.1 Discrete Wavelet Transform

Wavelet transforms provide a powerful tool for the analysis of transient and non-stationary data and is particularly useful in picking out characteristic variations at different resolutions or scales. This linear transform separates a data set in the form of low-pass or average coefficients, which reflect the average behaviour of the data, and wavelet or high-pass coefficients at different levels, which capture the variations at corresponding scales. As compared to Fourier or window Fourier transform, wavelets allow optimal "time-frequency" localisation in the real as well as in the Fourier domain. The vocabulary of DWT stems from applications in one dimensional time-stream signal trains, but has found wide application in signal in other domains and dimensions. Specifically in our case, the signal being transformed is the primordial power spectrum, $\mathcal{P_R}(k)$, a one dimensional function of wavenumber, k.

Wavelets are an orthonormal basis of small waves, with their variations primarily concentrated in a finite region, which make them ideal for analysing localised transient signals. Wavelets can be continuous or discrete. In the latter case, the

basis elements are strictly finite in size, enabling them to achieve localisation, while disentangling characteristic variations at different frequencies [74]. For more details about DWT and its theoretical basis, see [47].

In this paper, we use the prescription of Ref. [48], and smooth the raw deconvolved spectrum with the following method:

1. Increase the number of k-values of the spectrum to a power of 2, by padding evenly on both sides

2. Perform discrete wavelet transform

3. Keep the 2^N wavelet coefficients corresponding to lowest frequency features

4. Perform inverse discrete wavelet transform and calculate the likelihood for the smoothed spectrum

5. Repeat steps 3.-4. for $N \in \{0, 11\}$ and keep the spectrum that provides the best likelihood.

A.3 Generating random realisations of WMAP data

Given an observed CMB temperature anisotropy $\Delta T(\hat{\mathbf{n}})/\bar{T}$, after performing the usual expansion into spherical harmonics

$$\Delta T(\hat{\mathbf{n}})/\bar{T} = \sum_{\ell m} a_{\ell m}^{\rm obs} Y_{\ell m}(\hat{\mathbf{n}}), \tag{A.10}$$

we can define an observed angular power spectrum

$$\mathfrak{C}_\ell^{\rm obs} \equiv \mathcal{C}_\ell^{\rm obs} + \mathcal{N}_\ell^{\rm obs} = \frac{1}{2\ell+1} \sum_{m=-\ell}^{\ell} \left| a_{\ell m}^{\rm obs} \right|^2, \tag{A.11}$$

which can be split up into an original signal $\mathcal{C}_\ell^{\rm obs}$ and a contribution from experimental noise $\mathcal{N}_\ell^{\rm obs}$. Theory, on the other hand, can only predict the average of this quantity over an ensemble of independent observations, $\mathfrak{C}_\ell^{\rm th} = \langle \mathfrak{C}_\ell^{\rm obs} \rangle$.

In the following, we shall describe how to generate independent realisations of simulated mock data (which replace the $\hat{\mathcal{C}}_\ell$ in the associated mock likelihood function) from a theoretical input spectrum $\mathfrak{C}_\ell^{\rm fid}$. Since the $\hat{\mathcal{C}}_\ell$ are actually defined as the difference between the observed spectrum and the expected noise power spectrum, $\hat{\mathcal{C}}_\ell = \mathfrak{C}_\ell^{\rm obs} - \mathcal{N}_\ell^{\rm th}$, the corresponding simulated quantity is $\mathfrak{C}_\ell^{\rm sim} - \mathcal{N}_\ell^{\rm th}$.

For the sake of clarity we commence with the case of an ideal (i.e., noise-free, full-sky) observation.

A.3.1 No experimental noise, full sky-coverage, no correlations

Even in the case of an ideal measurement of the CMB temperature anisotropies ($\mathcal{N}_\ell = 0$), the observed angular power spectrum would still be subject to cosmic variance. Under the assumption of Gaussian fluctuations, the coefficients $a_{\ell m}$ can be considered Gaussian random variables for the purpose of simulating data sets $\mathcal{C}_\ell^{\rm sim}$. Isotropy dictates $\langle a_{\ell m} a^*_{\ell' m'} \rangle = \mathcal{C}_\ell^{\rm fid} \delta_{\ell\ell'} \delta_{mm'}$, so different multipoles will be independent of each other. Hence, for each ℓ the simulated $\mathcal{C}_\ell^{\rm sim}$ can be constructed by drawing random numbers x_ℓ from a $\chi^2_{2\ell+1}$-distribution:

$$\mathfrak{C}_\ell^{\rm sim} = \mathcal{C}_\ell^{\rm sim} = \mathcal{C}_\ell^{\rm fid} \frac{x_\ell}{\langle x_\ell \rangle} = \mathcal{C}_\ell^{\rm fid} + \mathcal{C}_\ell^{\rm fid} \left(\frac{x_\ell}{2\ell+1} - 1 \right). \tag{A.12}$$

A.3.2 Adding experimental noise

Realistic observations are of course subject to instrumental noise.

Assuming the noise to be isotropic and Gaussian, its effect on the distribution the $a_{\ell m}$ are drawn from can be considered as a convolution with a Gaussian of width $\sqrt{\mathcal{N}_\ell^{\rm th}}$, and we have

$$\mathfrak{C}_\ell^{\rm sim} - \mathcal{N}_\ell^{\rm th} = \mathcal{C}_\ell^{\rm sim} + \mathcal{N}_\ell^{\rm sim} - \mathcal{N}_\ell^{\rm th} = \mathcal{C}_\ell^{\rm fid} + \left(\mathcal{C}_\ell^{\rm fid} + \mathcal{N}_\ell^{\rm th} \right) \left(\frac{x_\ell}{2\ell+1} - 1 \right). \tag{A.13}$$

A.3.3 Partial sky coverage and correlations

Even for nominally full-sky observations, some parts of the sky cannot be used to construct spectra and need to be cut out, since the signal is obscured by point sources and galactic foregrounds. This has two consequences on the recovered $\hat{\mathcal{C}}_\ell$. First, using only part of the sky leads to a loss of information, resulting in a larger scatter of $\hat{\mathcal{C}}_\ell$ around the true underlying spectrum $\mathcal{C}_\ell^{\rm fid}$. Second, since the sky cut breaks isotropy, the (pseudo)-$\hat{\mathcal{C}}_\ell$ constructed from an incomplete map will no longer be uncorrelated [75].

Since the curvature matrix (A.4) appearing in the likelihood function scales like $f_{\rm sky}^2$, the standard deviation of $\mathcal{C}_\ell^{\rm sim} - \mathcal{C}_\ell^{\rm fid}$ should scale like $f_{\rm sky}^{-1}$. This increase in scatter around the fiducial model can be modelled by drawing the random variables x_ℓ from a $\chi^2_{f_{\rm sky}^2(2\ell+1)}$-distribution. Since it is numerically much easier to generate χ^2 variates with an integer number of degrees of freedom we round $f_{\rm sky}^2(2\ell+1)$ to the nearest integer and eventually rescale the resulting $y_\ell \equiv x_\ell / \langle x_\ell \rangle - 1$ accordingly, to recover the correct standard deviation.

The so-generated y_ℓ are uncorrelated random variables. For the simulated spectrum to have the same correlation properties as the real WMAP data, we take instead the vector

$$(\tilde{y}_\ell) \equiv {\rm Chol}(K_{\ell\ell'})(y_{\ell'}), \tag{A.14}$$

where the lower triangular matrix $\mathrm{Chol}(K_{\ell\ell'})$ is the Cholesky decomposition of the correlation matrix $K_{\ell\ell'} = \delta_{\ell\ell'} + r_{\ell\ell'}$. Finally, the simulated spectra are given by

$$\mathfrak{C}_\ell^{\rm sim} - \mathcal{N}_\ell^{\rm th} = \mathcal{C}_\ell^{\rm fid} + \left(\mathcal{C}_\ell^{\rm fid} + \mathcal{N}_\ell^{\rm th}\right) \tilde{y}_\ell. \tag{A.15}$$

The above procedure can rarely result in negative values of $\mathcal{C}_\ell^{\rm sim} + \mathcal{N}_\ell^{\rm sim}$ at low multipoles, due to the rescaling of the y_ℓ or due to the transformation of Eq. (A.15). Such data is unphysical and not compatible with the log-normal piece of the likelihood function and must therefore be discarded. As a result, the simulated data can be slightly biased; however, the bias is negligible and remains smaller than $\mathcal{O}(1\%)$ of cosmic variance for all ℓ.

References

[1] A. A. Starobinsky, Phys. Lett. B **117** (1982) 175.

[2] A. H. Guth and S. Y. Pi, Phys. Rev. Lett. **49** (1982) 1110.

[3] J. M. Bardeen, P. J. Steinhardt and M. S. Turner, Phys. Rev. D **28** (1983) 679.

[4] V. F. Mukhanov, H. A. Feldman and R. H. Brandenberger, Phys. Rept. **215** (1992) 203.

[5] J. Martin, A. Riazuelo and M. Sakellariadou, Phys. Rev. D **61** (2000) 083518 [arXiv:astro-ph/9904167].

[6] J. Martin and R. H. Brandenberger, Phys. Rev. D **63** (2001) 123501 [arXiv:hep-th/0005209].

[7] U. H. Danielsson, Phys. Rev. D **66** (2002) 023511 [arXiv:hep-th/0203198].

[8] C. R. Contaldi, M. Peloso, L. Kofman and A. Linde, JCAP **0307** (2003) 002 [arXiv:astro-ph/0303636].

[9] J. A. Adams, G. G. Ross and S. Sarkar, Nucl. Phys. B **503** (1997) 405 [arXiv:hep-ph/9704286].

[10] M. Joy, V. Sahni and A. A. Starobinsky, Phys. Rev. D **77** (2008) 023514 [arXiv:0711.1585 [astro-ph]].

[11] A. A. Starobinsky, JETP Lett. **55** (1992) 489 [Pisma Zh. Eksp. Teor. Fiz. **55** (1992) 477].

[12] J. A. Adams, B. Cresswell and R. Easther, Phys. Rev. D **64** (2001) 123514 [arXiv:astro-ph/0102236].

[13] R. K. Jain, P. Chingangbam, J. O. Gong, L. Sriramkumar and T. Souradeep, JCAP **0901** (2009) 009 [arXiv:0809.3915 [astro-ph]].

[14] D. J. H. Chung, E. W. Kolb, A. Riotto and I. I. Tkachev, Phys. Rev. D **62** (2000) 043508 [arXiv:hep-ph/9910437].

[15] Ø. Elgarøy, S. Hannestad and T. Haugbølle, JCAP **0309** (2003) 008 [arXiv:astro-ph/0306229].

[16] L. A. Kofman and A. D. Linde, Nucl. Phys. B **282** (1987) 555.

[17] L. A. Kofman and D. Y. Pogosyan, Phys. Lett. B **214**, 508 (1988).

[18] D. S. Salopek, J. R. Bond and J. M. Bardeen, Phys. Rev. D **40** (1989) 1753.

[19] D. Polarski and A. A. Starobinsky, Nucl. Phys. B **385** (1992) 623.

[20] N. Kaloper and M. Kaplinghat, Phys. Rev. D **68** (2003) 123522 [arXiv:hep-th/0307016].

[21] J. M. Cline and L. Hoi, JCAP **0606** (2006) 007 [arXiv:astro-ph/0603403].

[22] R. K. Jain, P. Chingangbam, L. Sriramkumar and T. Souradeep, arXiv:0904.2518 [astro-ph.CO].

[23] M. R. Nolta *et al.* [WMAP Collaboration], Astrophys. J. Suppl. **180** (2009) 296 [arXiv:0803.0593 [astro-ph]].

[24] J. Dunkley *et al.* [WMAP Collaboration], Astrophys. J. Suppl. **180** (2009) 306 [arXiv:0803.0586 [astro-ph]].

[25] E. Komatsu *et al.* [WMAP Collaboration], Astrophys. J. Suppl. **180** (2009) 330 [arXiv:0803.0547 [astro-ph]].

[26] J. Barriga, E. Gaztanaga, M. G. Santos and S. Sarkar, Mon. Not. Roy. Astron. Soc. **324** (2001) 977 [arXiv:astro-ph/0011398].

[27] H. V. Peiris *et al.* [WMAP Collaboration], Astrophys. J. Suppl. **148** (2003) 213 [arXiv:astro-ph/0302225].

[28] J. Martin and C. Ringeval, Phys. Rev. D **69** (2004) 083515 [arXiv:astro-ph/0310382].

[29] J. Martin and C. Ringeval, Phys. Rev. D **69** (2004) 127303 [arXiv:astro-ph/0402609].

[30] J. Martin and C. Ringeval, JCAP **0501** (2005) 007 [arXiv:hep-ph/0405249].

[31] P. Hunt and S. Sarkar, Phys. Rev. D **70** (2004) 103518 [arXiv:astro-ph/0408138].

[32] P. Hunt and S. Sarkar, Phys. Rev. D **76** (2007) 123504 [arXiv:0706.2443 [astro-ph]].

[33] M. Joy, A. Shafieloo, V. Sahni and A. A. Starobinsky, JCAP **0906** (2009) 028 [arXiv:0807.3334 [astro-ph]].

[34] M. Bridges, A. N. Lasenby and M. P. Hobson, Mon. Not. Roy. Astron. Soc. **369** (2006) 1123 [arXiv:astro-ph/0511573].

[35] D. N. Spergel *et al.* [WMAP Collaboration], Astrophys. J. Suppl. **170** (2007) 377 [arXiv:astro-ph/0603449].

[36] L. Covi, J. Hamann, A. Melchiorri, A. Slosar and I. Sorbera, Phys. Rev. D **74** (2006) 083509 [arXiv:astro-ph/0606452].

[37] J. Hamann, L. Covi, A. Melchiorri and A. Slosar, Phys. Rev. D **76** (2007) 023503 [arXiv:astro-ph/0701380].

[38] N. Barnaby and Z. Huang, arXiv:0909.0751 [astro-ph.CO].

[39] K. Ichiki, R. Nagata and J. Yokoyama, arXiv:0911.5108 [astro-ph.CO].

[40] S. L. Bridle, A. M. Lewis, J. Weller and G. Efstathiou, Mon. Not. Roy. Astron. Soc. **342** (2003) L72 [arXiv:astro-ph/0302306].

[41] S. Hannestad, JCAP **0404** (2004) 002 [arXiv:astro-ph/0311491].

[42] P. Mukherjee and Y. Wang, Astrophys. J. **599** (2003) 1 [arXiv:astro-ph/0303211].

[43] S. M. Leach, Mon. Not. Roy. Astron. Soc. **372** (2006) 646 [arXiv:astro-ph/0506390].

[44] N. Kogo, M. Matsumiya, M. Sasaki and J. Yokoyama, Astrophys. J. **607** (2004) 32 [arXiv:astro-ph/0309662].

[45] A. Shafieloo and T. Souradeep, Phys. Rev. D **70** (2004) 043523 [arXiv:astro-ph/0312174].

[46] D. Tocchini-Valentini, M. Douspis and J. Silk, Mon. Not. Roy. Astron. Soc. **359** (2005) 31 [arXiv:astro-ph/0402583].

[47] A. Shafieloo, T. Souradeep, P. Manimaran, P. K. Panigrahi and R. Rangarajan, Phys. Rev. D **75** (2007) 123502 [arXiv:astro-ph/0611352].

[48] A. Shafieloo and T. Souradeep, Phys. Rev. D **78** (2008) 023511 [arXiv:0709.1944 [astro-ph]].

[49] R. Nagata and J. Yokoyama, Phys. Rev. D **78** (2008) 123002 [arXiv:0809.4537 [astro-ph]].

[50] R. Nagata and J. Yokoyama, Phys. Rev. D **79** (2009) 043010 [arXiv:0812.4585 [astro-ph]].

[51] K. Ichiki and R. Nagata, Phys. Rev. D **80** (2009) 083002.

[52] G. Nicholson and C. R. Contaldi, arXiv:0903.1106 [astro-ph.CO].

[53] G. Nicholson, C. R. Contaldi and P. Paykari, arXiv:0909.5092 [astro-ph.CO].

[54] R. Trotta, Mon. Not. Roy. Astron. Soc. **378** (2007) 72 [arXiv:astro-ph/0504022].

[55] A. R. Liddle, arXiv:0903.4210 [hep-th].

[56] G. Efstathiou, arXiv:0802.3185 [astro-ph].

[57] D. N. Spergel *et al.* [WMAP Collaboration], Astrophys. J. Suppl. **148** (2003) 175 [arXiv:astro-ph/0302209].

[58] A. Lewis, arXiv:astro-ph/0310186.

[59] W. H. Press, S. A. Teukolsky, W. T. Vetterling and B. P. Flannery, "Numerical Recipes in Fortran," Cambridge University Press, 1992

[60] A. Lewis and S. Bridle, Phys. Rev. D **66** (2002) 103511 [astro-ph/0205436].

[61] A. Lewis, A. Challinor and A. Lasenby, Astrophys. J. **538** (2000) 473 [arXiv:astro-ph/9911177].

[62] W. A. Fendt and B. D. Wandelt, Astrophys. J. **654** (2006) 2 [arXiv:astro-ph/0606709].

[63] W. A. Fendt and B. D. Wandelt, arXiv:0712.0194 [astro-ph].

[64] A. Shafieloo and T. Souradeep, arXive:0901.0716 [astro-ph]

[65] H. V. Peiris and L. Verde, arXiv:0912.0268 [astro-ph.CO].

[66] [Planck Collaboration], arXiv:astro-ph/0604069.

[67] M. J. Mortonson, C. Dvorkin, H. V. Peiris and W. Hu, Phys. Rev. D **79** (2009) 103519 [arXiv:0903.4920 [astro-ph.CO]].

[68] D. Baumann *et al.* [CMBPol Study Team Collaboration], AIP Conf. Proc. **1141** (2009) 10 [arXiv:0811.3919 [astro-ph]].

[69] L. Verde *et al.* [WMAP Collaboration], Astrophys. J. Suppl. **148** (2003) 195 [arXiv:astro-ph/0302218].

[70] L. B. Lucy, Astron. J. **79** (1974) 745.

[71] B. H. Richardson, J. Opt. Soc. Am., **62**, 55 (1972).

[72] C. M. Baugh and G. Efstathiou, Mon.Not.Roy.Astron.Soc. **265**, 145 (1993).

[73] C. M. Baugh and G. Efstathiou, Mon.Not.Roy.Astron.Soc. **267**, 323 (1994).

[74] I. Daubechies, Ten Lectures on Wavelets, SIAM, Philadelphia, PA, 1992. Vol.64, CBMS-NSF Conference Series in Applied Mathematices.

[75] B. D. Wandelt, E. Hivon and K. M. Gorski, Phys. Rev. D **64** (2001) 083003 [arXiv:astro-ph/0008111].