# Bayesian analysis of inflation. III. Slow roll reconstruction using model selection

Jorge Noreña, Christian Wagner, Licia Verde, Hiranya V. Peiris, and Richard Easther

# Bayesian Analysis of Inflation III: Slow Roll Reconstruction Using Model Selection

Jorge Noreña,[1, *] Christian Wagner,[1, †] Licia Verde,[1, 2, ‡] Hiranya V. Peiris,[3, §] and Richard Easther[4, 5, ¶]

[1]*ICC, University of Barcelona (IEEC-UB), Marti i Franques 1, Barcelona 08028, Spain*

[2]*ICREA, Institució Catalana de Recerca i Estudis Avançats*

[3]*Department of Physics and Astronomy, University College London, London WC1E 6BT, U.K.*

[4]*Department of Physics, Yale University, New Haven, CT 06520, U.S.A.*
[5]*Department of Physics, University of Auckland, Private Bag 92019, Auckland, New Zealand*

We implement Slow Roll Reconstruction – an optimal solution to the inverse problem for inflationary cosmology – within MODECODE, a publicly available solver for the inflationary dynamics. We obtain up-to-date constraints on the reconstructed inflationary potential, derived from the WMAP 7-year dataset and South Pole Telescope observations, combined with large scale structure data derived from SDSS Data Release 7. Using MODECODE in conjunction with the MULTINEST sampler, we compute Bayesian evidence for the reconstructed potential at each order in the truncated slow roll hierarchy. We find that the data are well-described by the first two slow roll parameters, $\epsilon$ and $\eta$, and that there is no need to include a nontrivial $\xi$ parameter.

## I. INTRODUCTION

Inflation [1, 2] postulates that the very early universe underwent a period of accelerated expansion. It is typically described via the dynamics of the inflaton, a scalar effective degree of freedom coupled to gravity. Quantum fluctuations of the inflaton and other light fields constitute the most widely studied mechanism for producing primordial density perturbations (for reviews see Refs. [3, 4]). These perturbations induce temperature fluctuations in the cosmic microwave background radiation (CMB) and, as a consequence of their gravitational growth, lead to the formation of the large scale structure of the universe (LSS).

A key corollary of the inflationary hypothesis is that precise measurements of the CMB and the distribution of galaxies in the sky constrain the physical mechanism that drives the accelerated expansion, opening a remarkable window into the physics of the very early universe. In recent years, high-quality observations of the CMB have been provided by the WMAP satellite [47], ACBAR [5], the South Pole Telescope (SPT) [50] and the Atacama Cosmology Telescope (ACT) [6], while the Sloan Digital Sky Survey (SDSS) [7] has probed the density perturbations at low redshifts. Further, data from the Planck satellite [8] will soon be available to the cosmological community, and several major LSS surveys are now underway or are being planned.

Inflation can be realized in a vast number of ways, and these scenarios are typically distinguishable via their predictions for the statistical properties of the primordial perturbations. Some models are purely *ad hoc*, while others satisfy naturalness constraints, or can be derived from candidate theories of fundamental physics. Given the plethora of models that exist, it is fruitful to consider the corresponding inverse problem: given cosmological data, can we deduce the mechanism underlying inflation? This problem is often referred to as *reconstruction* [9–13]. Ideally, reconstruction would recover the effective action of the "inflationary sector" of high energy physics, but carrying out this program for a fully general scenario is not likely to be feasible, even with data from the next generation of astrophysical experiments. Consequently, present-day implementations of reconstruction incorporate a (sometimes implicit) prior restricting attention to a specific category of models – usually a single inflaton with a canonical kinetic term, minimally coupled to Einstein gravity.

This paper builds on the Slow Roll Reconstruction algorithm, proposed by Easther and Peiris [14–18]. Slow Roll Reconstruction does not require the slow roll *approximation*, but rather relies on the truncated Hamilton-Jacobi slow roll *expansion* [13, 19–23]. This is a phenomenological description of inflation, obtained by expanding the Hubble parameter as a power series, with the inflaton field as the independent variable or "clock". We focus on scales which are directly probed by observations, making minimal assumptions regarding the reheating mechanism and the expansion history of the universe during the "primordial dark age" [24].

Given the long history of the inflationary inverse problem, Slow Roll Reconstruction has a number of antecedents. In particular, other approaches to the inverse problem based on the Hamilton-Jacobi equations include Ref. [25] and "Monte Carlo Reconstruction" [26, 27], which was further developed in Refs [28–30]. Separately,

---
*Electronic address: `jorge.norena@icc.ub.edu`
†Electronic address: `cwagner@icc.ub.edu`
‡Electronic address: `liciaverde@icc.ub.edu`
§Electronic address: `h.peiris@ucl.ac.uk`
¶Electronic address: `r.easther@auckland.ac.nz`

Ref. [31] constrained the slow roll parameters by requiring that the duration of inflation was sufficient to solve the classic cosmological problems, while Cline and Hoi reconstructed inflationary models that allow for a significant running of the scalar spectral index within the Hamilton-Jacobi formalism [32]. Leach and collaborators constrained inflation by writing the spectral indices in terms of the slow roll parameters [33–35]. A similar scheme to Slow Roll Reconstruction was discussed in Refs. [36, 37]. Likewise, Ref. [38] used the WMAP5 dataset and SDSS data release 7 LRG data to constrain the Hubble slow roll parameters, while Ref. [39] obtained constraints on the power spectra of curvature and tensor perturbations using priors based on single field slow roll inflationary models.

We implement Slow Roll Reconstruction within MODECODE[1], a publicly available solver for the inflationary background and perturbations [40, 41]. As shown by Liddle [23] the background dynamics corresponding to the truncated slow roll hierarchy can be solved analytically, yielding the corresponding inflationary potential $V(\phi)$. Adding this potential to the menu of models supported within MODECODE yields a simple and robust implementation of Slow Roll Reconstruction. Further, MODECODE performs a full numerical computation of the inflationary perturbation spectrum, making no use of the slow roll approximation when computing the power spectrum. We then estimate the slow roll parameters using the WMAP7 data, and the recent SPT and SDSS DR7 data releases.

The principal advantage of implementing Slow Roll Reconstruction within MODECODE is that, by using the nested sampler MULTINEST [42, 43] with COSMOMC [44], we compute Bayesian evidence at each order in the truncated slow roll hierarchy. This information determines the number of slow roll parameters that are required to account for the data, and thus the optimal order at which to truncate the hierarchy. In addition, we carefully construct the priors for the slow roll parameters to ensure that the computed evidence values are realistic [41]. Finally, we compute two heuristic model selection statistics – the Profile Likelihood ratio and the Akaike Information Criterion – for the truncated slow roll hierarchy, and compare these to the Bayesian inferences.

This is the third in a sequence of papers on Bayesian analysis of inflation, and optimal approaches to constraining inflationary models with astrophysical data. The first [40] focussed on estimating the free parameters in specific inflationary models, while the second [41] addressed the model selection problem in inflation and the computation of Bayesian evidence.

The paper is organized as follows: In Section II we describe Slow Roll Reconstruction and summarize its strengths. Section III contains a detailed description of

our analysis, and summarizes the data used to generate the constraints. We present our results in Section IV, and summarize the implications of our findings in Section V.

## II. SLOW ROLL RECONSTRUCTION

During inflation the background metric is well-described by the flat Friedmann Robertson Walker metric

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + a(t)^2\mathrm{d}x^2\,, \tag{1}$$

where $a(t)$ is the scale factor and, as usual, the Hubble parameter $H \equiv \dot{a}/a$ where dots denote derivatives with respect to time, $t$. Inflation is, by definition, a period of accelerated expansion during which $\ddot{a} > 0$.

Slow Roll Reconstruction [14–18] is based on the inflationary flow hierarchy. This is obtained by rewriting the second Friedmann equation, $\dot{H} = -\dot{\phi}^2/(2M_{\mathrm{pl}}^2)$, as

$$\dot{\phi} = -2M_{\mathrm{pl}}^2 H'(\phi)\,, \tag{2}$$

where $M_{\mathrm{pl}}$ is the reduced Planck mass and primes denote derivatives with respect to $\phi$. We define the slow roll parameters, $^{(n)}\lambda$:

$$^{(0)}\lambda = 2M_{\mathrm{Pl}}^2\left[\frac{H'(\phi)}{H(\phi)}\right]^2 \equiv \epsilon \tag{3}$$

for $n = 0$, and

$$^{(n)}\lambda \equiv (2M_{\mathrm{pl}}^2)^n\frac{(H')^{n-1}}{H^n}\frac{\mathrm{d}^{n+1}H}{\mathrm{d}\phi^{n+1}} \tag{4}$$

for $n \geq 1$. These are related to the Hubble slow roll parameters by noting that $\eta \equiv {}^{(1)}\lambda$ and $\xi \equiv {}^{(2)}\lambda$ [45]. We obtain the Hamilton-Jacobi equations [19–22] by differentiating $^{(n)}\lambda$ with respect to $\phi$, yielding an infinite hierarchy of differential equations where $^{(n+1)}\lambda$ is determined by the two previous terms in the expansion. If $^{(n)}\lambda \equiv 0$ for all $n > N$ at a specific value of $\phi$, the hierarchy truncates, and these terms must be zero for all values of $\phi$. As pointed out by Liddle [23], this ensures that the higher-order derivatives of $H(\phi)$ vanish for all $\phi$, which is equivalent to requiring that $H(\phi)$ is a polynomial of finite order,

$$H(\phi) = H_*\big(1 + A_1\phi + A_2\phi^2 + \cdots + A_N\phi^N\big)\,, \tag{5}$$

where a subscript $*$ denotes a quantity evaluated at $\phi = 0$. Multiplying $H$ by a constant leaves the $^{(n)}\lambda$ unchanged; as we will see below, $H_*$ corresponds to the overall energy scale of the inflationary era. The coefficients $A_1, \ldots, A_N$ in equation (5) are then related to the Hubble slow roll parameters by

$$A_1 = \left(\frac{\epsilon_*}{2M_{\mathrm{pl}}^2}\right)^{1/2}\,, \tag{6}$$

$$A_n = \left(\frac{1}{2M_{\mathrm{pl}}^2}\right)^{n-1}\frac{A_1^{n-1}}{n!}{}^{(n-1)}\lambda_*\,. \tag{7}$$

The system is not modified by the shift $\phi \to \phi + \phi_0$, but this rescaling does implicitly redefine the $A_n$. We remove this ambiguity without loss of generality by assuming that the slow roll parameters are measured at the instant at which $\phi = 0$. Further, the slow roll hierarchy simplifies when $\epsilon \ll 1$, leading to a distinct class of solutions, and we test both "High-$\epsilon$ N-parameter" and "Low-$\epsilon$ N-parameter" models [18] in what follows.

Given an expression for $H(\phi)$ we can always obtain the corresponding potential $V(\phi)$ [19–23] by recalling that $H^2 = (\dot\phi^2/2 + V(\phi))/3M_{\mathrm{pl}}^2$ and replacing $\dot\phi$ with $H'$,

$$V(\phi) = M_{\mathrm{pl}}^2 H_*^2 \left[ 3(1 + A_1\phi + \cdots + A_N\phi^N)^2 \right.$$
$$\left. - 2(A_1 + \cdots + NA_N\phi^{N-1})^2 \right]. \quad (8)$$

This expression is a function of the $A_n$ and thus the slow roll parameters, so Slow Roll Reconstruction can be implemented by adding equation (8) to the "menu" of inflationary models supported by MODECODE [40, 41].

The Fourier components of the scalar and tensor modes $u_k$ and $v_k$ are obtained by solving

$$u_{k,\tau\tau} + \left( k^2 + \frac{z_{,\tau\tau}}{z} \right) u_k = 0\,, \quad (9)$$

$$v_{k,\tau\tau} + \left( k^2 - \frac{a_{,\tau\tau}}{a} \right) v_k = 0\,, \quad (10)$$

where $z \equiv \dot\phi/H$ and the subscript $\tau$ denotes a derivative with respect to conformal time $d\tau \equiv dt/a$. Initial conditions for these modes are set by the Bunch-Davies vacuum when the mode $k$ is well within the horizon, and the amplitude of the power spectrum, defined by

$$\langle \mathcal{R}_{\mathbf{k}} \mathcal{R}_{\mathbf{k}'} \rangle = \frac{2\pi^2}{k^3} \Delta_{\mathcal{R}}^2(k)(2\pi)^3 \delta^{(3)}(\mathbf{k} - \mathbf{k}')\,, \quad (11)$$

is computed from the solution when the mode is much larger than the horizon and frozen in:

$$\Delta_{\mathcal{R}}^2(k) = \frac{k^3}{2\pi^2} \left| \frac{u_k}{z} \right|^2\,, \qquad \Delta_t^2(k) = \frac{4}{\pi^2} \frac{k^3}{M_{\mathrm{pl}}^2} \left| \frac{v_k}{a} \right|^2\,. \quad (12)$$

As usual, $\mathcal{R}$ denotes the curvature perturbations in comoving gauge, and sometimes the notation $P_{\mathcal{R}}(k) = \frac{2\pi^2}{k^3} \Delta_{\mathcal{R}}^2(k)$ is used.

The amplitude of the power spectrum can also be computed in the slow roll approximation. Evaluated at $k_*$ (the scale which exits the horizon as $\phi = 0$), the amplitude is given in terms $H_*$ and the slow roll parameters by:

$$A_{\mathrm{sr}} = \frac{\left[ 1 - (2C + 1)\epsilon_* + C\eta_* \right]^2}{\pi\epsilon_*} \frac{H_*^2}{8\pi}\,, \quad (13)$$

where $A_{\mathrm{sr}}$ denotes $\Delta_{\mathcal{R}}^2(k_*)$ to second order in the slow-roll expansion. Here, $C = -2 + \ln 2 + \gamma$, and $\gamma$ is the Euler-Mascheroni constant [13].

## III. METHOD

Slow Roll Reconstruction treats the individual terms in the slow roll hierarchy as free parameters. By implementing the potential, equation (8), within MODECODE [40, 41] we obtain posterior distributions for $\{\epsilon_*, \eta_*, \xi_*\}$ (or subsets of these variables) and compute Bayesian evidence to determine the optimal truncation-order for this expansion. MODECODE modifies COSMOMC [44], which is used in conjunction with the MULTINEST sampler [42, 43].

### A. Bayesian model comparison using MULTINEST

Consider the probability distribution for a set of parameters $\alpha$, given a model $M$ and the data $D$, denoted by $P(\alpha|D, M)$ [46]. Bayes' theorem yields the *posterior*

$$P(\alpha|D, M) = \frac{P(D|\alpha, M)P(\alpha|M)}{P(D|M)}\,, \quad (14)$$

where $\mathcal{L} \equiv P(D|\alpha, M)$ is the likelihood, $P(\alpha) \equiv P(\alpha|M)$ is the prior, and $E \equiv P(D|M)$ is the Bayesian evidence. Since the probability is normalized to unity we can compute the evidence directly, via

$$E = \int d\alpha^N \mathcal{L}(\theta)P(\alpha)\,, \quad (15)$$

which is an $N$-dimensional integral over a volume defined by the parameter ranges permitted by the prior.

Using Bayes' theorem again, we have

$$\frac{P(M_1|D)}{P(M_2|D)} = \frac{E_1}{E_2} \frac{P(M_1)}{P(M_2)}\,. \quad (16)$$

Given the *a priori* probability ratio $P(M_1)/P(M_2)$, $E_1/E_2$ yields the ratio of probabilities for the two models, in the light of the data. Any *a priori* preference for $M_1$ over $M_2$ is quantified by $P(M_1)/P(M_2)$; this ratio is always set to unity in our analysis. Consequently, the evidence ratio $E_1/E_2$ expresses the relative "betting odds" for two models.

The evidence, equation (15), depends on both the value (or height, given that this is a positive definite integrand) of the likelihood $\mathcal{L}$ and the total volume of the parameter space. The evidence thus encodes the understanding that a model in which $\mathcal{L}$ has a substantial amplitude over a large fraction of the permitted parameter volume is more predictive than one which does not [41].

### B. Implementing Slow Roll Reconstruction

We fix $\phi = 0$ to correspond to the time when the pivot scale $k_*$ leaves the horizon and choose $k_* = 0.05\,\mathrm{Mpc}^{-1}$. We sample the slow roll parameters at the pivot scale,

with priors specified in Section III C. In principle, the initial value of the field $\phi_i$ is determined by the value of $\phi$ at the onset of inflation, or at which the density approaches Planckian values. In practice, the largest physical scale to which data are sensitive is $k_{\min} \sim 10^{-5}$ Mpc$^{-1}$ and we set $\phi_i$ by solving for the value of $\phi$ at which the mode with wavenumber $k = k_{\min}/100$ leaves the horizon. The initial value of $\dot{\phi}_i$ is fixed using the slow roll attractor $\dot{\phi}_i = -2M_{\text{pl}}^2 H'(\phi_i)$. We discard models for which inflation breaks down (*i.e.* $\epsilon > 1$) as the field evolves from $\phi = \phi_i$ to $\phi = 0$, as these scenarios are not consistent with the assumption that the primordial universe is inflating.

The smallest scale we consider is $k_{\max}$ which we take to be $k_{\max} \sim 10^2$ Mpc$^{-1}$, the smallest scale accessible to cosmological observations. The power spectrum is computed from the amplitudes of the solutions to the mode equations when the corresponding scale is far outside the horizon, so for self-consistency we require inflation to last until $k_{\max}$ is a hundred times larger than the horizon. Assuming $H$ to be roughly constant during inflation (which is consistent with limits from data), this puts a lower bound on the number of $e$-folds of inflation occurring after the pivot scale leaves the horizon,

$$N_{\min} \gtrsim \log\left(\frac{100\, k_{\max}}{k_*}\right) \approx 12\,. \tag{17}$$

We discard models with fewer than $N_{\min}$ $e$-folds, measured from the instant $k_*$ leaves the horizon.

As implemented here, Slow Roll Reconstruction only considers the piece of the potential which is directly probed by cosmological data. This amounts to marginalizing over the value of $\phi$ at which inflation ends – physically, this can occur for example because the potential has a sharp "cliff" at some value of $\phi$ for which $\epsilon > 1$ when computed from equation (8). Alternatively, this algorithm can also be implemented by evolving $\phi$ until inflation naturally terminates ($\ddot{a} < 0$) and then imposing a prior based on the total number of $e$-foldings that occur after the pivot scale leaves the horizon [17]. By contrast, in this analysis we assume nothing about the evolution of the universe after $k_{\max}$ has reached its asymptotic value after leaving the horizon.

## C. Models and Priors

The parameters which define the Hubble slow roll model to third order in the expansion defined in Eq. (5) are $H_*$, $\epsilon_*$, $\eta_*$ and $\xi_*$. But for all the models we consider, $H_*$ can, in principle, vary over many orders of magnitude and there is little clear *a priori* theoretical justification for restricting this range. However, given the slow roll parameters, we see from equation (13) that $H_*$ is strongly correlated with the amplitude of the primordial spectrum. Consequently, we treat $\Delta_{\mathcal{R}}^2(k_*)$, the amplitude of the primordial scalar power spectrum, for

which the choice of prior range is more straightforward, as an independent parameter with a logarithmic prior. We allow a range that is generous relative to current constraints on this parameter (e.g. Ref. [47]). However our results do not depend significantly on the specific choice, and setting a narrower range in the prior would implicitly inject information from the data used to compute the posterior. We obtain $H_*$ directly from equation (13), given values for $A_{\text{sr}}$ and the slow roll parameters. This is the one point at which we use the slow roll *approximation*; the precise amplitude of the power spectrum at the pivot $\Delta_{\mathcal{R}}^2(k_*)$ extracted from a numerical solution of the mode equations differs slightly from $A_{\text{sr}}$. In principle we could solve numerically for the exact value of $H_*$ needed to reproduce $\Delta_{\mathcal{R}}^2(k_*)$, but $H_*$ only changes at the level of a few percent if this is done. Note that the slow roll approximation is used only in order to set this prior; for the likelihood computation the numerical solution for $\Delta_{\mathcal{R}}^2(k_*)$ is used.

The specific set of models we consider is:

- "$\epsilon_*$ only": $\eta_* = \xi_* = 0$, $\epsilon_*$ and $A_{\text{sr}}$ are varied.

- "$\eta_*$ only": $\epsilon_* = 10^{-6}$, $\xi_* = 0$, $\eta_*$ and $A_{\text{sr}}$ are varied.

- "$\epsilon_*$, $\eta_*$": $\xi_* = 0$, $\epsilon_*$, $\eta_*$ and $A_{\text{sr}}$ are varied.

- "$\epsilon_*$, $\eta_*$, $\xi_*$": all parameters varied.

The "$\eta_*$ only" case is the "Low-$\epsilon$, 1-Parameter" scenario described in Ref. [18]: When $\epsilon$ is much smaller than 1, it effectively decouples from the slow roll hierarchy, meaning that the inflationary dynamics are then independent of the exact value of $\epsilon$ and are determined by the other slow roll parameters, $\eta$ in this case. We implement this by fixing $\epsilon_*$ to a small, finite value and have checked that the results we obtain are insensitive to the specific choice at which we fix $\epsilon_*$.

We consider both uniform and logarithmic priors on $\epsilon_*$, and impose uniform priors on $\eta_*$ and $\xi_*$. Since the tensor-to-scalar ratio is strongly correlated with $\epsilon$, a uniform prior for this parameter biases the analysis toward models where inflation occurs at high energy scales, which are likely to produce a significant background of tensor perturbations [48]. Likewise, $A_{\text{sr}}$ is drawn from a logarithmic prior. The prior ranges for all model parameters are listed in Table I, while the priors for other free cosmological parameters are specified in Table II.

The ranges of the slow roll parameters allowed by these priors are substantially larger than those used in most previous analyses [14, 15, 17]; since we are computing the evidence we only want to stipulate that inflation is taking place, without requiring that the slow roll parameters are small, in order to compute a self-consistent value for the evidence. Requiring that inflation lasts $N_{\min}$ $e$-folds after the pivot scale exits the horizon imposes an additional prior on the slow roll parameters. The resulting joint prior on any pair of slow roll parameters is no longer "rectangular" – Figure 1 illustrates the effective priors after this constraint is imposed.

| Parameter | Lower limit | Upper limit |
|-----------|-------------|-------------|
| $\log_{10} \epsilon_*$ | $-10$ | $0$ |
| $\epsilon_*$ | $0$ | $1$ |
| $\eta_*$ | $-1$ | $1$ |
| $\xi_*$ | $-1$ | $1$ |
| $\log(10^{10} A_{\mathrm{sr}})$ | $2.7$ | $4$ |

TABLE I: Prior ranges for the model parameters. Runs are performed with both log and uniform priors on $\epsilon_*$; the corresponding ranges are given in the first two rows.

| Parameter | Prior |
|-----------|-------|
| Baryon fraction | $0.015 < \Omega_b h^2 < 0.035$ |
| Dark matter | $0.05 < \Omega_{\mathrm{dm}} h^2 < 0.2$ |
| Reionization | $0.01 < \tau < 0.25$ |
| Projected acoustic scale | $0.8 < \theta < 1.2$ |
| Sunyaev-Zel'dovich Amplitude | $0 < A_{\mathrm{SZ}} < 2$ |

TABLE II: Prior ranges for other free parameters. All priors are uniform. The universe is assumed to be flat ($\Omega_k = 0$), so that the fractional energy density of dark energy is not an independent parameter. The effective number of relativistic species is set to 3.046. For the SPT dataset, we also include two nuisance parameters for the power from Poisson and clustered point sources (see [50] for details).

If $N_{\min}$ is increased, the allowed region of parameter space becomes smaller, as we discuss in Section IV B. A small volume of parameter space for which inflation lasts longer than $N_{\min}$ e-folds is excluded by our prior ranges on the slow roll parameters – however, this region is small and we have checked that modifying this choice does not impact our results.

### D. Data

We use the WMAP 7 year likelihood (WMAP7) [49] with data on the damping tail of the CMB temperature power spectrum from the South Pole Telescope [50]. We follow Ref. [50] in computing the SPT likelihood, setting $l_{\max} = 3000$, and marginalizing over foreground contributions from unresolved point sources and Sunyaev-Zel'dovich (SZ) clusters. These data are presented in Figure 2, which also shows the physical wavenumbers which contribute to the corresponding angular scales. In addition, we use the power spectrum of luminous red galaxies (LRG) derived from the SDSS data release 7 [51]. We adopt the public LRG likelihood code released by the SDSS collaboration, marginalizing over the amplitude and using only the information from the shape of the power spectrum. This likelihood automatically marginalizes over a possible scale dependence of the small-scale bias. The LRG power spectrum is shown in Figure 3.

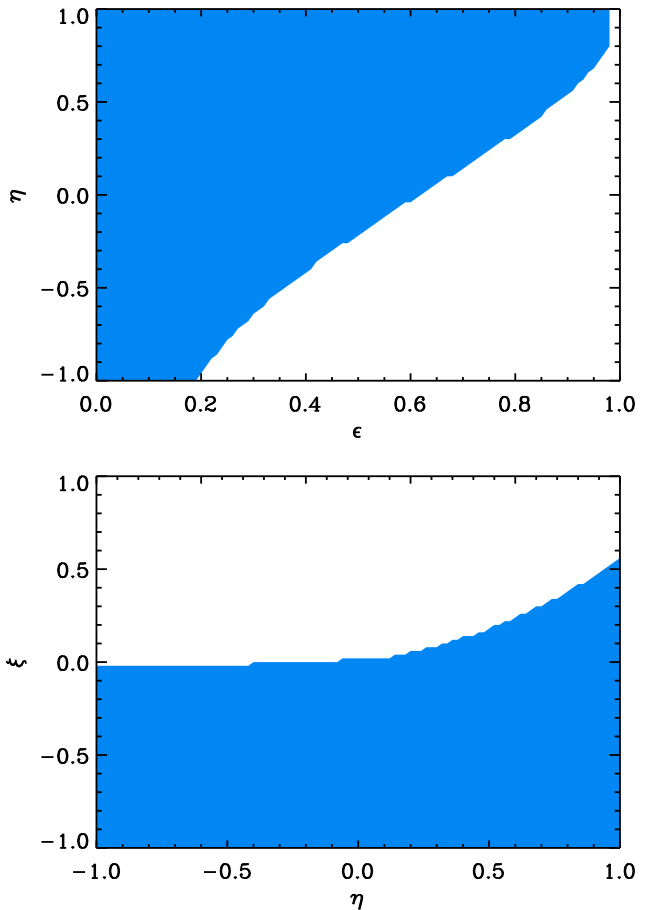In what follows we use the data combinations WMAP,



FIG. 1: The shaded region shows the projections of the prior volume after imposing the condition that the model achieves at least $N_{\min} = 12$ e-folds of inflation after the pivot scale leaves the horizon.

WMAP+SPT and WMAP+SPT+LRG – the dependence of the evidence ratios on the different data combinations reveals the information content of each dataset in this model selection problem.

### IV. RESULTS

#### A. Parameter estimation

Figure 4 presents the posterior constraints from WMAP+SPT+LRG on $\epsilon_*$ and $\eta_*$ for the "$\epsilon_*$, $\eta_*$" and "$\epsilon_*$, $\eta_*$, $\xi_*$" models, while Figure 5 show the constraints on $\epsilon_*$ and $\xi_*$ for both logarithmic and uniform priors on $\epsilon_*$. Appendix A tabulates the constraints on each of the slow roll parameters for all combinations of model, $\epsilon$ prior and dataset.

Observe that the posteriors all peak in regions of parameter space where the slow roll parameters are small. Consequently we can conclude that, within the context of this model, slow roll inflation is preferred by the data.
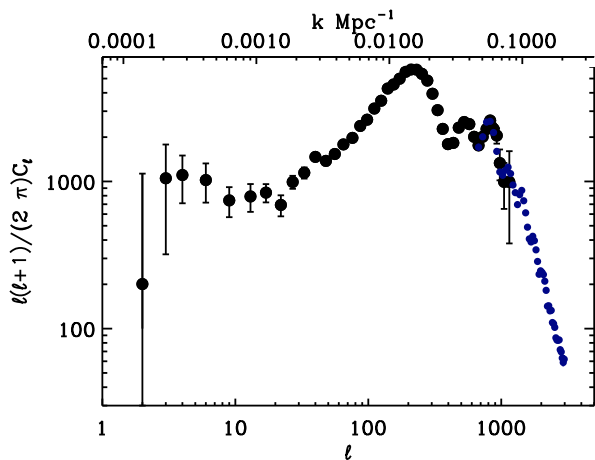
FIG. 2: The CMB temperature angular power spectrum as measured by WMAP 7 year data (black dots) and the South Pole Telescope (blue dots). The error bars on the SPT data are smaller than the dots.
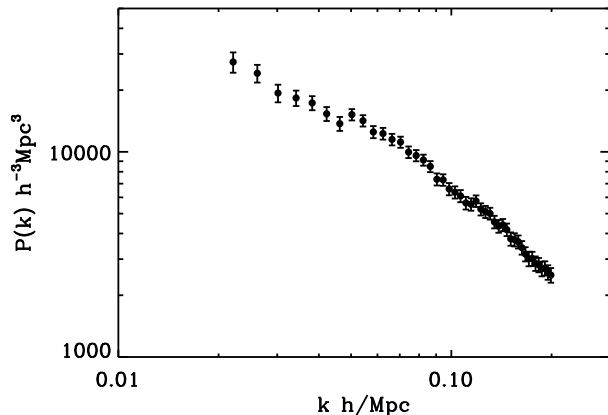


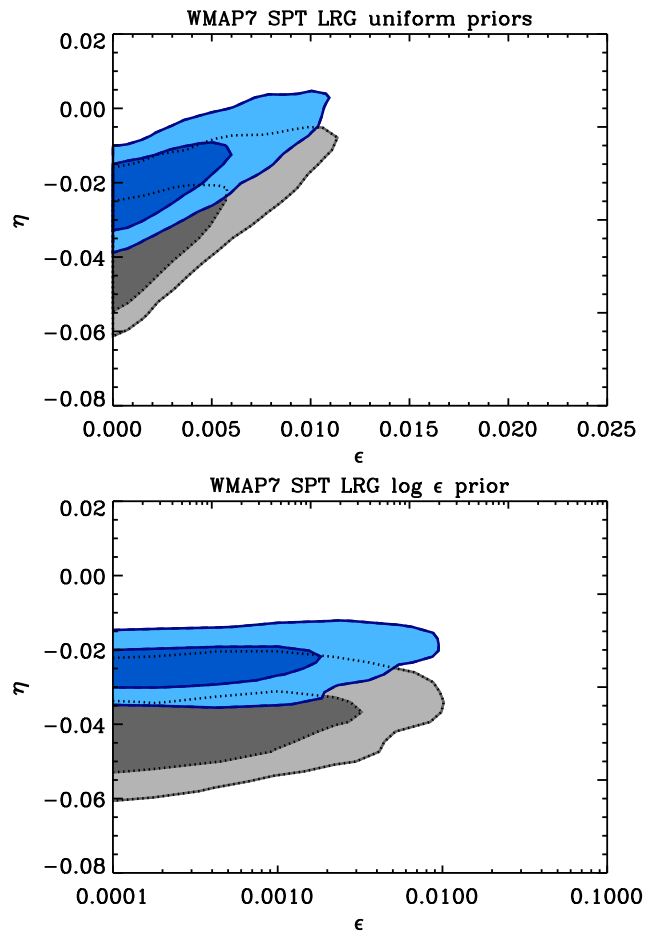FIG. 3: LRG power spectrum from the SDSS data release 7.



FIG. 4: Joint constraints on $\epsilon_*$ and $\eta_*$ using data from WMAP7, SPT and SDSS LRGs, marginalizing over all other parameters. The blue (upper) contours correspond to the "$\epsilon_*$, $\eta_*$" model while the grey (lower) contours correspond to the "$\epsilon_*$, $\eta_*$, $\xi_*$" model. Contours correspond to 68 and 95% joint confidence levels; blue denotes the "$\epsilon_*$, $\eta_*$" model and grey denotes the '$\epsilon_*$, $\eta_*$, $\xi_*$" model. Top: uniform prior on $\epsilon_*$. Bottom: log prior on $\epsilon_*$.

Constraints on $\eta_*$ differ when the uniform prior on $\epsilon_*$ is replaced with a logarithmic prior. We interpret this as the data currently not being informative enough to overcome the priors; see section IV B for further discussion.

For comparison, we show the same information in Figure 6 using only WMAP7 data. We see that the addition of SPT and LRG data lead to better constraints on $\eta_*$. Note that one cannot directly compare the constraint $\epsilon_* < 0.016$ (95% CL) to the constraint on the tensor-to-scalar ratio $r < 0.36$ (95% CL) reported by the WMAP7 team [49]. We find $r < 0.27$ (95% CL), which is close to the value one would obtain by simply using the slow roll approximation $r \approx 16\epsilon_*$. This $\sim 30\%$ difference can be explained by the different choice for the pivot scale (Ref. [49] uses $k_p = 0.002$ Mpc$^{-1}$) and the difference in the models used. The Hubble slow roll model "$\epsilon_*$, $\eta_*$" is not just a simple reparametrization of the phenomenological model parametrized by the spectral index $n_s$ and the

tensor-to-scalar ratio $r$, which was applied in Ref. [49].

### B. Bayesian evidence

As discussed in Section III A the evidence is not an absolute quantity: the ratio of the evidence values for two models expresses the relative "betting odds" that these models are responsible for the observed state of the universe. In Figure 7 we plot $\Delta \log E(\epsilon_*; \epsilon_*, \eta_*) \equiv \log[E(\epsilon_*)/E(\epsilon_*, \eta_*)]$ for uniform and log priors on $\epsilon_*$. There appears to be "strong" evidence for the "$\epsilon_*$ only" model compared to the "$\epsilon_*$, $\eta_*$" model when using uniform priors, while the log priors give "inconclusive" results. A similar result is seen in Figure 8, which shows $\Delta \log E(\eta_*; \epsilon_*, \eta_*)$. Again the uniform priors on $\epsilon_*$ gives "strong" evidence for "$\eta_*$ only" scenario compared to the
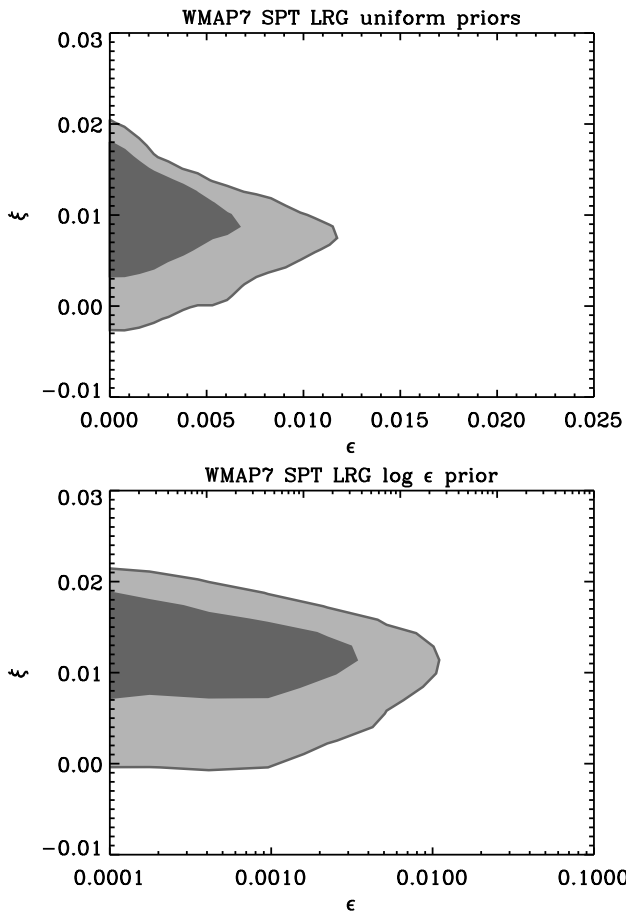
FIG. 5: Joint constraints on $\epsilon_*$ and $\xi_*$ using data from WMAP7, SPT and SDSS LRGs, marginalizing over all other parameters. Contours correspond to 68 and 95% joint confidence levels. Top: uniform prior on $\epsilon_*$. Bottom: log prior on $\epsilon_*$.
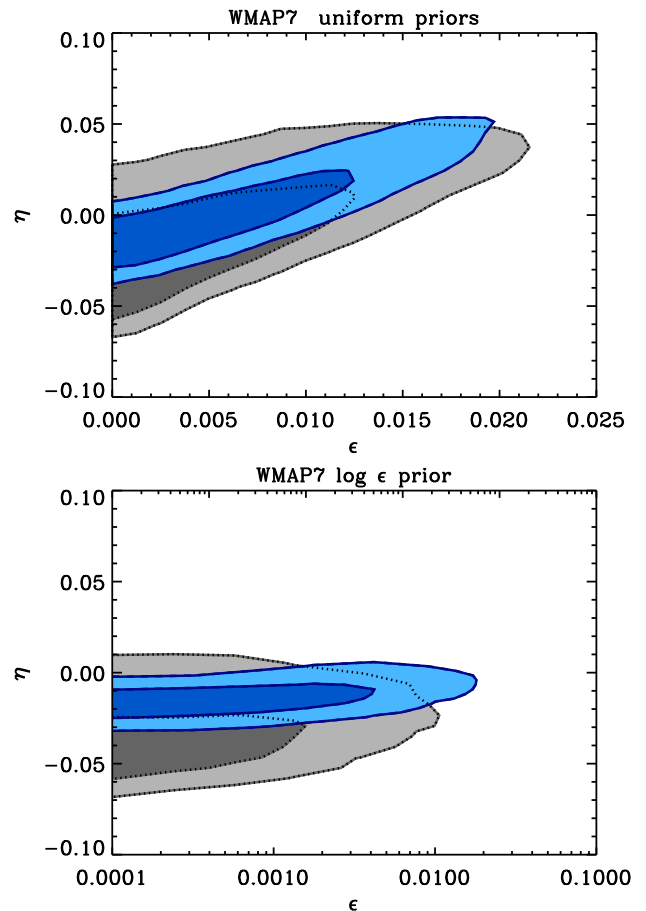
FIG. 6: Joint constraints on $\epsilon_*$ and $\eta_*$ using data from WMAP7 only, after marginalizing over all other parameters. The contours are in the same format as in Figure 4. Top: uniform prior on $\epsilon_*$. Bottom: log prior on $\epsilon_*$.

"$\epsilon_*$, $\eta_*$" model, while the evidence ratio is "inconclusive" with the log prior.

In the cases with uniform $\epsilon_*$ priors, the evidence clearly tells us that a single parameter (either "$\epsilon_*$ only" or "$\eta_*$ only") is preferred by the data over the two-parameter "$\epsilon_*$, $\eta_*$" case; however, the current data are not strong enough to tell us *which* of the single parameter models to pick.

When using the log $\epsilon_*$ priors, the parameter space is heavily weighted towards small $\epsilon_*$, approaching the "Low-$\epsilon$" scenario, $\epsilon \ll 1$, where this parameter effectively decouples from the Hubble slow roll hierarchy. Hence, in the scenarios with log priors, even though at face value we are comparing a two parameter model with a single parameter model, we are effectively comparing two single-parameter models. In the absence of a detection of tensor modes or a much stronger upper limit on tensors than is currently available, $\epsilon$ becomes a nuisance parameter which does not have a significant effect on the likelihood.

We can gain further insight into these results by map-

ping the slow roll parameter into the usual spectral index $n_s$ and tensor-to-scalar ratio $r$ which, to first order in slow roll,

$$n_s - 1 \approx 2\eta_* - 4\epsilon_* , \tag{18}$$

$$r \approx 16\epsilon_* . \tag{19}$$

Varying $\epsilon_*$ changes both $n_s$ and $r$, while varying $\eta_*$ changes only $n_s - 1$. Data favor models with $n_s - 1 \approx -0.05$. With a log prior the "$\epsilon_*$ only" model gives greater weight to the $r \approx 0$, $n_s \approx 1$ region of parameter space, which is disfavored by data, compared to the uniform prior. Further, with a log prior the likelihood changes very slowly over most of the range spanned by log $\epsilon_*$. In this case, we would expect the evidence computed for the "$\epsilon_*$, $\eta_*$" model to approach the "$\eta_*$ only" model, as seen in Figure 7 and 8. This result also mirrors earlier discussions of the prior dependence of estimates of the tensor-to-scalar ratio [48].

Considered together, Figs. 7 and 8 for a uniform $\epsilon_*$ prior indicate that the evidence "decisively" prefers a single-parameter model (to explain $n_s < 1$) over a two-
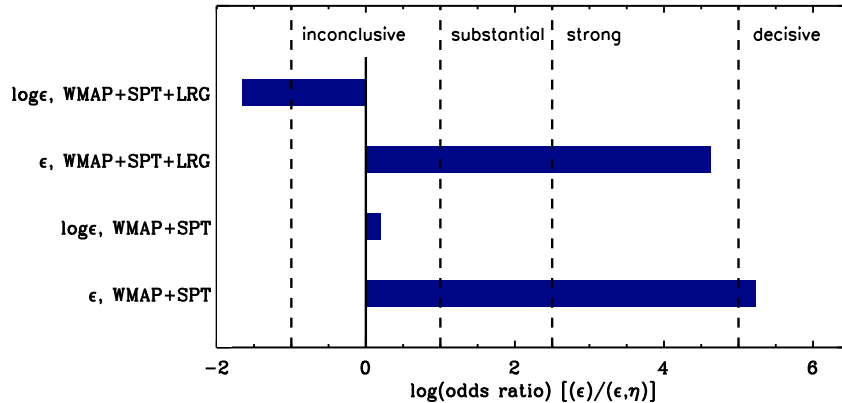
FIG. 7: $\Delta \log E(\epsilon_*; \epsilon_*, \eta_*)$ for log and uniform priors on $\epsilon_*$ and different data sets, as indicated on the vertical axis. The estimates of $\Delta \log E(\epsilon_*; \epsilon_*, \eta_*)$ have an uncertainty of $\sim 0.3$ at the numerical settings used in this work. When taking the ratios, the same priors were used in the numerator and denominator.
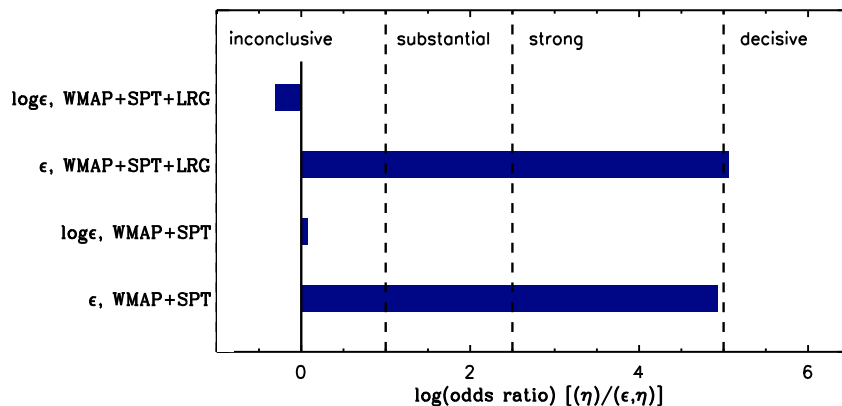


FIG. 8: $\Delta \log E(\eta_*; \epsilon_*, \eta_*)$ for log and uniform priors on $\epsilon_*$ and different data sets, as indicated on the vertical axis. The estimates of $\Delta \log E(\eta_*; \epsilon_*, \eta_*)$ have an uncertainly of $\sim 0.3$ at the numerical settings used in this work. When taking the ratios, the same priors were used in the numerator and denominator.

parameter model, but without a tensor detection (or a strong upper limit) it cannot discriminate whether that parameter should be $\epsilon$ or $\eta$.

In Figure 9 we plot $\Delta \log E(\epsilon_*, \eta_*; \epsilon_*, \eta_*, \xi_*)$ for different choices of $\epsilon$ priors and data. According to the usual scale [52] for comparing models, there is "strong" evidence in favour of the "$\epsilon_*$, $\eta_*$" model compared to the "$\epsilon_*$, $\eta_*$, $\xi_*$" model, indicating that $\xi_*$ is not needed to explain the data, and this result is largely independent of the prior on $\epsilon_*$. However, adding SPT and LRG data *decreases* the significance of this conclusion. If the inflationary phase was well described by only $\epsilon_*$ and $\eta_*$, adding more data should increase the evidence ratio, in the absence of systematic effects in the data. However, even if the true underlying model had only $\epsilon_*$ and $\eta_*$, a small systematic mismatch between different data sets (e.g. normalization issues) can lead to a spurious preference for a non-zero $\xi_*$.

To determine the impact of $N_{\min}$ on our results we also performed runs with $N_{\min} = 24$. Figure 10 shows the constraints on $\eta_*$ and $\xi_*$ for $N_{\min} = 12$ (grey contours) and $N_{\min} = 24$ (blue contours). As expected, requiring fewer $e$-folds typically allows for a broader range in the slow roll parameters, especially for $\xi_*$ [14]. Note also that the region allowed with $N_{\min} = 12$ but excluded with $N_{\min} = 24$ has a reasonably high likelihood. Consequently, excluding this region lowers the computed evidence for the "$\epsilon_*$, $\eta_*$, $\xi_*$" model, further disfavoring the presence of $\xi_*$ in the parameter set. Figure 10 also shows that the posterior distribution for $\xi_*$ is truncated by the $e$-folds prior with $N_{\min} = 24$, while it more closely resembles an ellipse with $N_{\min} = 12$. For our $N_{\min} \approx 24$ runs, the log of the evidence ratios increased by $\sim 1$, since in this case the $N_{\min}$ prior truncates a high-likelihood region with somewhat large values of $\xi_*$.
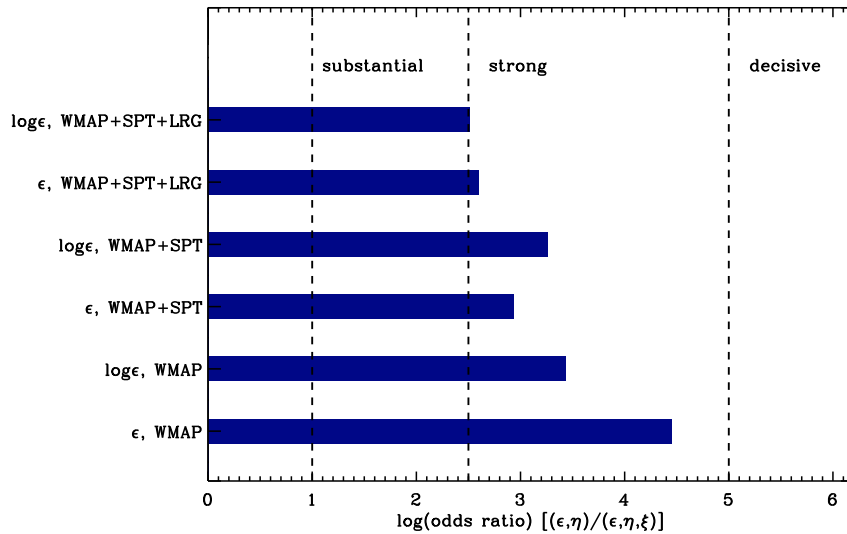
FIG. 9: $\Delta \log E(\epsilon_*, \eta_*; \epsilon_*, \eta_*, \xi_*)$ for log and uniform priors on $\epsilon_*$ and different data sets as indicated in the vertical axis. The values of $\Delta \log E(\epsilon_*, \eta_*; \epsilon_*, \eta_*, \xi_*)$ have an uncertainty of $\sim 0.3$ at the numerical settings used in this work.
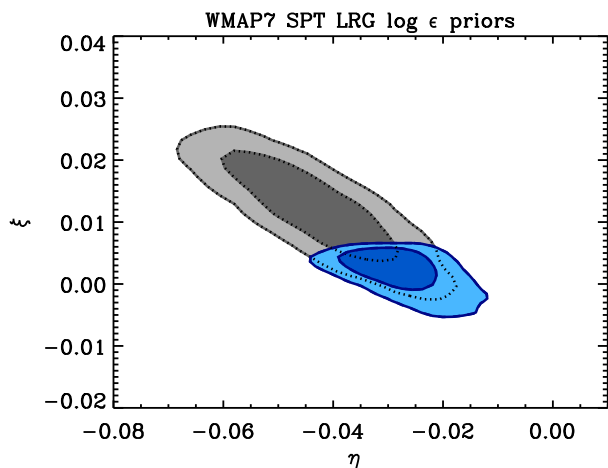


FIG. 10: Constraints on $\eta_*$ and $\xi_*$. These constraints were obtained by using data from WMAP7, SPT and SDSS LRGs after marginalizing over all other parameters with a log prior on $\epsilon_*$. Contours correspond to 68 and 95% joint confidence levels. The blue contours correspond to discarding models which don't achieve $N_{\min} = 24$ e-folds from the moment at which the pivot scale exits the horizon, the gray contours correspond to $N_{\min} = 12$ e-folds.

## C.   The Profile Likelihood Ratio.

Given an $n$-parameter model, the $n$-dimensional likelihood encodes the information content of the data. Often, in parameter estimation problems, we are primarily interested in the confidence intervals for a single parameter. Using Bayes theorem, we can promote the likelihood to a probability density function (the posterior) by multiplying by the prior probability density, and then integrat-

ing – or "marginalizing" the posterior over the remaining $n-1$ parameters. Confidence intervals obtained from the marginalized posterior thus depend on the prior.

Comparing the "$\epsilon_*$ only", "$\eta_*$ only" and "$\epsilon_*$, $\eta_*$" models in Section IV B we see that these results are strongly prior dependent. However, these priors are phenomenologically-motivated, insofar as they are not derived from fundamental physical considerations, so unlike a set of physically-derived model priors, we cannot be confident that these priors are appropriate. Thus, at least with current data, the answers to the model selection problems for this specific subset of models cannot be considered definitive.

Consequently, it is worth investigating statistics which rely only on the likelihood and are thus prior-independent, even if these do not provide a fully consistent model selection criterion, and we now discuss the Profile Likelihood Ratio (PLR) (see e.g. Refs. [53–55]) and turn to the Akaike Information Criterion (AIC) [56] in the following subsection. The PLR is obtained by taking the maximum value of the likelihood for fixed values of the interesting parameter; it is the ratio of the conditional to the unconditional maximum likelihood. This is a straightforward generalization of the delta chi-square ($\Delta\chi^2$) for a multidimensional likelihood in the case that we only need constraints on a single parameter: under certain regularity conditions the distribution of $-2\ln$ PLR converges to a chi-square distribution [53].

By construction, the PLR statistic is prior-independent and has an interpretation similar to that of the $\Delta\chi^2$ where an effective chi-square is identified with $-2\ln\mathcal{L}$. This analogy is exact when the likelihood is Gaussian. However, one must keep in mind that the confidence intervals may not have strict frequentist coverage, especially if the likelihood is far

from Gaussian. With this caveat in mind, best fit values and their confidence intervals can be obtained from the $\Delta\chi^2$, or the PLR. If the best fit parameter differs from zero at $> n$–$\sigma$ one reports an $n$–$\sigma$ "evidence" for that quantity. This "evidence" should not be confused with the Bayesian evidence (also called evidence for short). Being solely a description of the likelihood, the PLR is not a self-consistent model-comparison statistic, but allows us to investigate whether preferences for extra parameters in posterior confidence intervals are driven by the data or by the prior.

The use of the PLR in cosmology is relatively recent [54, 55]. It can be simply computed from a MULTI-NEST output using the following algorithm. Assume we have $m$ uninteresting parameters (which are marginalized over in the standard Bayesian approach) and one parameter, $\beta$, on which we want to report constraints. For each value of $\beta$, we find the maximum likelihood value $\mathcal{L}_\beta$, for all values of the other $m$ parameters. We then compute $\ln(\mathcal{L}_\beta/\mathcal{L}_{\max})$, where $\mathcal{L}_{\max}$ is the overall maximum likelihood, which we identify with the maximum likelihood found by MULTINEST during its exploration of the full-dimensional prior. We then use the pseudo-chisquare defined as $\ln(\mathcal{L}_\beta/\mathcal{L}_{\max}) = 1/2\chi^2$, so that $\Delta\ln(\mathcal{L}_\beta/\mathcal{L}_{\max}) = 0.5$ and 2 correspond to the 68.3% and 95.4% confidence regions respectively. This quantity is, at least in principle, independent of priors.

The PLR is useful in testing nested models: cases where a more complex model has one extra parameter ($\beta$), compared to a simpler model in which $\beta$ is fixed. For instance, in the three-parameter model we can identify $\beta$ with $\xi_*$; in the simpler model $\xi_* \equiv 0$.

Since MULTINEST uses the uncertainty on the Bayesian Evidence integral as a criterion for convergence, the computed PLR may be noisy. The MULTINEST runs terminate when the evidence integral converges – and this convergence criterion does not guarantee that the PLR has also converged. This can be a problem in practice, and the resulting numerical noise in the computed PLR is visible in some of the plots we present.

In Fig. 11 we show the PLR-derived constraints on the $\xi_*$ parameter. From this plot we conclude that including $\xi_*$ as an extra parameter improves the likelihood so that $\xi_* = 0$ is excluded, at about the 2-$\sigma$ limit when the LRG data are included. This should be compared with Fig. 5 where the joint 2D posterior is plotted. Recall that the 2-$\sigma$ 1D marginalized constraints are close to the projection of 2D joint 1-$\sigma$ (see e.g. Ref. [57]). In both the (prior-dependent) posterior constraints and (prior-independent) PLR case, $\xi_* = 0$ is just outside the 2-$\sigma$ limit, indicating that the likelihood marginally prefers a positive $\xi_*$. This mild preference in the likelihood is not sufficient for the Bayesian evidence to require the inclusion of an extra parameter.

In Fig. 12 we show the PLR-derived constraints on the $\eta_*$ parameter. The analysis using a log prior on $\epsilon_*$ prefers a non-zero value of $\eta_*$, while repeating the analysis with uniform priors weakens these results. Since the
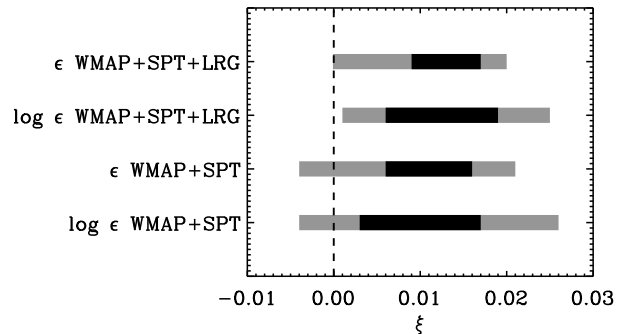


FIG. 11: Profile likelihood ratio-derived constraints on the $\xi$ parameter. Black lines show the 1-$\sigma$ intervals and grey lines the 2-$\sigma$ intervals. The dashed vertical line is a reference indicating the zero point. This result is consistent with Fig. 4 and indicates that the likelihood favors non-zero $\xi_*$ at about the 2-$\sigma$ level.
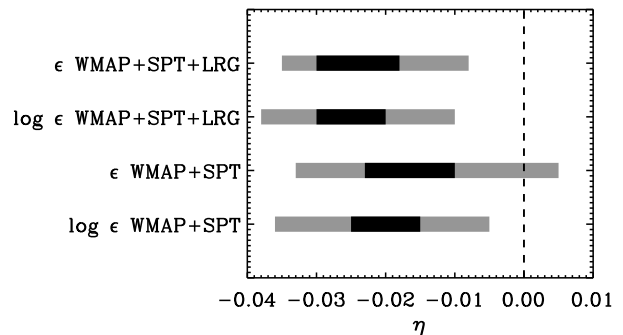


FIG. 12: Profile likelihood ratio-derived constraints on the $\eta$ parameter. Black lines show the 1-$\sigma$ intervals and grey lines the 2-$\sigma$ intervals. The dashed vertical line is a reference indicating the zero point. From this plot we conclude that including $\eta_*$ as an extra parameter improves the likelihood such that $\eta_* = 0$ is excluded, at about the 2-$\sigma$ limit.

PLR is supposed to be prior-independent, we interpret this difference as being due to numerical effects in the computation of the PLR, as we will discuss shortly.

First, let us compare these results with Fig. 4 where the joint 2D posterior is plotted. In both cases, for the combination WMAP+SPT+LRG, $\eta_* = 0$ is outside the 2-$\sigma$ limit, indicating that the likelihood prefers a negative $\eta_*$. The PLR for $\epsilon_*$ (not shown) indicates that $\epsilon_* = 0$ is always within the 1-$\sigma$ error-bar; the PLR distribution is one-sided, decreasing as $\epsilon_*$ increases. Again, this is consistent with Fig. 4. The fact that only upper limits on $\epsilon_*$ can be placed reflect the fact that there is no detection of tensor modes. But $n_s < 1$ is needed to fit the data, driving the likelihood to favor negative $\eta_*$ values (see Eq. 18).

The PLR results for $\xi_*$ (Figure 11) are largely independent of the prior, as expected, but some prior dependence

can be appreciated in the PLR for $\eta_*$ (Figure 12), particularly on the side of the errorbars towards zero (*i.e.* large $\eta_*$). In fact, there is a correlation between $\epsilon_*$ and $\eta_*$, with $\eta_*$ becoming less negative for large $\epsilon_*$ (Figs. 4 and 6). A log prior on $\epsilon_*$ undersamples the likelihood for large $\epsilon_*$, thus underestimating the upper limit of $\eta_*$ in the PLR. The effect is stronger for the cases where larger values of $\epsilon_*$ are allowed by the data, where the log prior severely penalizes the sampling. This undersampling effect is much smaller for the case WMAP+SPT+LRG – which places tighter constraints on $\epsilon_*$– compared to the CMB-only case.

## D. The Akaike information criterion

Let us now turn to the Akaike information criterion. The Bayesian evidence is the only model-selection statistic with a self-consistent probabilitistic interpretation [46]. However, only relatively recently has it become possible to actually perform the required computationally-intensive numerical integrals in a reasonable time, in order to evaluate the Bayesian evidence in most practical applications in cosmology; suitable numerical techniques became available only recently [42, 43, 58]. Before then, approximate model-selection criteria were used. A popular example is the Akaike information criterion (AIC) [56] which has the advantage of being extremely easy and fast to compute. The AIC is based on the Kullback-Leibler information entropy [59, 60] and is defined as

$$\text{AIC} \equiv -2 \ln \mathcal{L}_{\max} + 2k \,, \qquad (20)$$

where $\mathcal{L}_{\max}$ is the maximum likelihood achievable by the model and $k$ is the number of model parameters.

We present the results from the AIC statistic for the models and data combinations considered here, in order to allow a direct comparison with past work where AIC was used; this makes it possible to gauge (albeit approximately) how much the new data have improved constraints on models. A comparison of AIC with the results from Bayesian evidence ratios can also be used to estimate how well the approximations involved in the AIC work. In fact Ref. [61] (which introduced the AIC to cosmology) found that, with the CMB data available at the time, the AIC and Bayesian evidence gave significantly different conclusions.

Tables III and IV present the AIC computed according to Eq. 20 for each model, with a uniform and logarithmic prior on $\epsilon_*$ respectively. The numerical uncertainty on the AIC values listed here was estimated to be $\sim 0.6$ (uncertainty in the maximum likelihood $\sim 0.3$). Given this level of numerical uncertainty the two tables are fully consistent with each other. Note that the model with lowest AIC is the one in which we vary only over $\eta_*$, and it is therefore taken as reference. The ratio $R$ of probabilities of minimizing the information loss between two models can be estimated via $R \approx \exp[(\text{AIC}_1 - \text{AIC}_2)/2]$.

| Model | $\Delta$AIC | Probability ratio |
|---|---|---|
| $\epsilon_*, \eta_*, \xi_*$ | 0.069 | 0.97 |
| $\epsilon_*, \eta_*$ | 2.7 | 0.26 |
| $\eta_*$ only | 0 | 1 |
| $\epsilon_*$ only | 5.9 | 0.051 |

TABLE III: Akaike information criterion for the models considered in this work, using WMAP+SPT+LRG data and uniform priors on $\epsilon_*$. The first column lists the models. The second column lists the difference between the AIC for each model computed with equation (20) and the lowest AIC, which happens to be the one for the "$\eta_*$ only" model. The third column lists the probability ratio for minimizing information loss between each model and the "$\eta_*$ only" model. The numerical uncertainty on the AIC was estimated to be $\sim 0.6$.

| Model | $\Delta$AIC | Probability ratio |
|---|---|---|
| $\epsilon_*, \eta_*, \xi_*$ | 0.75 | 0.69 |
| $\epsilon_*, \eta_*$ | 2.0 | 0.37 |
| $\eta_*$ only | 0 | 1 |
| $\epsilon_*$ only | 5.7 | 0.057 |

TABLE IV: Akaike information criterion for the models considered in this work, using WMAP+SPT+LRG data and log priors on $\epsilon_*$. The columns are as in Table III. The numerical uncertainty on the AIC was estimated to be $\sim 0.6$.

We list this ratio for each model with respect to the reference "$\eta_*$ only" model. As expected the "$\epsilon_*, \eta_*, \xi_*$" model has the highest likelihood, but it is penalized in the AIC for having more parameters than the other models, making it virtually equivalent to the "$\eta_*$ only" case. This is consistent with the PLR analysis; see Figure 11 where the maximum likelihood value for $\xi_*$ is $\sim 0.01$ but the error bars extend almost to zero. The maximum likelihoods for the "$\epsilon_*, \eta_*$" and "$\eta_*$ only" models are close to one another (given our level of numerical uncertainty) but the "$\epsilon_*, \eta_*$" is penalized for having an additional parameter.

The AIC for the "$\epsilon_*$ only" model is considerably larger than that for the "$\eta_*$ only" model. Both models have the same number of parameters, but the likelihood for the latter model is larger at its maximum, implying that it is a better fit to the data. This is in agreement with the PLR findings: a model with $\eta_* = 0$ is disfavored by the likelihood at about the 2-$\sigma$ level when using the WMAP+SPT+LRG data combination (see Figure 12), and is again consistent with the results shown in Figures 4 and 6.

The AIC is sometimes used as a heuristic model-selection statistic, with a penalty involving the number of model parameters going some way towards implementing Occam's razor. We should note here that comparing it to the results from the Bayesian evidence shows that this penalty is not sufficiently conservative. By using the AIC as a model-selection statistic, we would have

wrongly concluded that the "$\epsilon_*$, $\eta_*$, $\xi_*$" model is to be preferred over the "$\epsilon_*$, $\eta_*$" model. While the AIC leads to the correct conclusion that the single parameter "$\eta_*$ only" model is preferred over the "$\epsilon_*$, $\eta_*$" model, the AIC contradicts the evidence in concluding that "$\epsilon_*$ only" is disfavored with respect to the two-parameter model.

## V. SUMMARY AND CONCLUSIONS

In this paper we have applied Bayesian model selection to slow roll reconstruction of the inflationary dynamics, as described by equation (5) truncated at different orders. We investigated whether current cosmological data require the inclusion of high-order slow roll parameters (e.g. $\xi$), or whether they can be adequately described by just the lowest-order slow roll parameters (i.e. $\epsilon$ and $\eta$).

The self-consistent statistic to answer this question is the Bayesian evidence. In order to compute it efficiently we used the publicly available MULTINEST [42, 43] extension to CosmoMC [44]. MULTINEST also computes posterior distributions that can be used for parameter estimation; this enables us to constrain the slow roll parameters for each of the models as shown in Figures 4 and 5. We compared models where the slow roll expansion is truncated at different orders, as described in section III C, using the latest compilation of CMB and large scale structure data. The slow roll approximation can induce systematic biases which are relevant for imminent precision datasets, such as CMB data from the Planck satellite. This problem can be avoided through the numerical solution of the mode equations, for which we used the publicly available MODECODE [40].

Our approach was to treat the slow roll expansion as a phenomenological description of the inflationary dynamics, valid only during the period in which the scales that are accessible to observations evolve and freeze out. We made no assumptions about reheating or the evolution of the universe after the end of inflation, so the prior only excludes models unable to produce sufficient inflation to account for the generation of perturbations on physical scales directly accessible to observations. This roughly corresponds to requiring that a given model ensure at least $N_{\min} \approx 12$ e-folds of inflation after the moment at which the pivot scale becomes of the size of the horizon.

Even if we don't employ the slow roll approximation and impose only a minimal requirement on the duration of inflation, we see that the data prefer models which satisfy the slow roll conditions. We have simply required that models achieve just enough e-folds to generate the directly observable cosmological perturbations. However, if we increase the minimal number of e-folds in the prior, the slow roll-allowed parameters are driven to smaller values: a longer period of inflation requires that the potential is flat over a wider field range of field values, as discussed in Ref. [16].

Slow Roll Reconstruction as implemented here will not recover sharp features like steps, or a modulated poten-

tial. However, even if the power spectrum was generated by one of these potentials, they can only be well-described within the Hamilton-Jacobi hierarchy if a large number of higher order terms are retained, and it is not clear that the expansion would converge in this limit. Moreover, given that the calculated evidence values demonstrate that a one-parameter model is a good fit to the data, it is not clear that models with sharp features would be favored over the simplest slow roll models, given the overall formalism of Slow Roll Reconstruction.

The Bayesian evidence ratios are presented in Figures 9, 7 and 8. This analysis showed that the CMB data are well-described without the inclusion of $\xi_*$, and that there is "strong" evidence (on the Jeffreys scale) that the inclusion of this slow roll parameter is disfavored (i.e., a model with only $\epsilon_*$ and $\eta_*$ has strong evidence compared to a model with $\epsilon_*$, $\eta_*$ and $\xi_*$). The inclusion of LSS LRG power spectrum slightly weakens the results but the evidence remains "strong". This might be due to the combination of heterogeneous data sets; for example, something as simple as a systematic mismatch in normalization can be spuriously fit by an extra parameter. We used both logarithmic and uniform priors on $\epsilon_*$ and found that the numerical value for the evidence ratio depends slightly on the form of the prior for $\epsilon_*$. Despite this dependence, the evidence against the inclusion of $\xi_*$ remains "strong", making this conclusion robust to prior choice. These results are consistent with results both from minimally-parametric as well as Bayesian approaches to the reconstruction of the primordial power spectrum (for example, see Refs. [14, 62]).

We also compared the single-parameter models, in which only $\epsilon_*$ or $\eta_*$ are free parameters, with models containing both parameters, finding that there is no strong indication that more than one slow roll parameter is needed to fit the data. With a log $\epsilon_*$ prior the parameter space is heavily weighted towards small $\epsilon_*$, and approaches the "Low-$\epsilon$" scenario: $\epsilon_*$ decouples from the hierarchy, effectively making it a nuisance parameter in the absence of a tensor detection. With a uniform $\epsilon_*$ prior, the evidence "decisively" prefers a single parameter model over a two-parameter model, but does not indicate whether the parameter should be $\epsilon_*$ or $\eta_*$. A detection of (or strong upper limit on) primordial tensors would make it possible to differentiate the "$\epsilon_*$ only", "$\eta_*$ only" and "$\epsilon_*$, $\eta_*$" models; such a constraint would also be needed in order to obtain consistent model-selection results between logarithmic and uniform priors on $\epsilon_*$.

Since our Bayesian results are derived from phenomenological (rather than physical) priors, we also considered a prior-independent statistic, the profile likelihood ratio (PLR), which can used to derive prior-independent confidence levels which can then be compared with the prior-dependent posterior confidence intervals. Using the PLR, we found that the inclusion of $\xi_*$ as a parameter improves the likelihood just at the 2-$\sigma$ level. The inclusion of $\eta_*$ also improves the likelihood, so $\eta_* = 0$ is disfavored at slightly more than the 2-$\sigma$ level.

The PLR results are consistent with the Bayesian posterior constraints, indicating that for parameter estimation the prior dependence is mild.

In our analysis, we have uncovered a numerical pitfall in computing PLR from Bayesian samplers (see e.g. Ref. [63]). Indeed, MULTINEST uses the uncertainty on the Bayesian Evidence as a criterion for convergence. Therefore the computed PLR may not have the required coverage properties (*i.e.* errors are underestimated) due to the fact that it may undersample some regions of parameter space which give only a small contribution to the evidence. This happens for example when using a logarithmic prior on $\epsilon_*$, where the large $\epsilon_*$ region is poorly sampled. We found that this undersampling effect can be greatly reduced by running a Markov Chain Monte Carlo analysis with much more stringent convergence criteria than the COSMOMC defaults.

Finally, we computed the Akaike Information Criterion for the different models, finding that the results from AIC were consistent with both the PLR and the Bayesian posterior intervals. However, unlike the PLR, the AIC is not based purely on the likelihood, and attempts to correct for model complexity by penalizing the likelihood by the number of extra parameters. Previously the AIC has been used as an approximate model-selection statistic, as it is much easier to compute numerically than the Bayesian evidence. Comparing the model selection conclusions from the AIC with the Bayesian evidence, we see that the AIC tends to have a lower threshold for favoring the introduction of extra parameters. Thus, it tends to be less conservative than the Bayesian evidence in deciding whether the improvement in likelihood is sufficient to prefer a more complex model over a simpler one.

We expect that the SRR as implemented here will lead to significant improvements [64] in our knowledge of the inflationary dynamics when applied to future data, such as CMB data from the Planck satellite [8] and the LSS power spectrum as probed by the Baryon Oscillation Spectroscopic Survey (BOSS) [65] and the Euclid satellite [66].

### Appendix A: Bounds on the coefficients of the slow roll expansion.

In Table V we list the mean values and 68% confidence levels for the slow roll parameters at the pivot scale, for all the scenarios we considered. Note that the biggest improvement attained by including the SPT and LRG data is in the measurement of $\eta_*$, for which the data indicates a non-zero value, but this improvement is not highly significant ($\sim 1\sigma$). Note also that when $\xi_*$ is included, it is measured to be very small, with large uncertainties; this already hints at the fact that its inclusion is not required in order to fit the data.

| Run | Parameter | WMAP | | WMAP+SPT | | WMAP+SPT+LRG | |
|---|---|---|---|---|---|---|---|
| | | mean | $\sigma$ | mean | $\sigma$ | mean | $\sigma$ |
| $\epsilon_*$ only | $\epsilon_*$ | - | - | $7.25 \times 10^{-3}$ | $2.07 \times 10^{-3}$ | $8.52 \times 10^{-3}$ | $1.84 \times 10^{-3}$ |
| $\log_{10} \epsilon_*$ only | $\log_{10} \epsilon_*$ | - | - | $-2.33$ | $7.94 \times 10^{-1}$ | $-2.11$ | $1.98 \times 10^{-1}$ |
| $\eta_*$ only | $\eta_*$ | - | - | $-2.23 \times 10^{-2}$ | $5.40 \times 10^{-3}$ | $2.55 \times 10^{-2}$ | $4.85 \times 10^{-3}$ |
| $\epsilon_*$, $\eta_*$ | $\epsilon_*$ | $6.71 \times 10^{-3}$ | $4.96 \times 10^{-3}$ | $4.20 \times 10^{-3}$ | $3.39 \times 10^{-3}$ | $3.21 \times 10^{-3}$ | $2.76 \times 10^{-3}$ |
| | $\eta_*$ | $3.76 \times 10^{-3}$ | $1.89 \times 10^{-2}$ | $-1.10 \times 10^{-2}$ | $1.08 \times 10^{-2}$ | $-1.72 \times 10^{-2}$ | $8.42 \times 10^{-3}$ |
| $\log_{10} \epsilon_*$, $\eta_*$ | $\log_{10} \epsilon_*$ | $-6.19$ | $2.22$ | $-6.30$ | $2.16$ | $-6.30$ | $2.12$ |
| | $\eta_*$ | $-1.67 \times 10^{-2}$ | $7.79 \times 10^{-3}$ | $-2.15 \times 10^{-2}$ | $5.75 \times 10^{-3}$ | $-2.49 \times 10^{-2}$ | $4.93 \times 10^{-3}$ |
| $\epsilon_*$, $\eta_*$, $\xi_*$ | $\epsilon_*$ | $6.20 \times 10^{-3}$ | $5.02 \times 10^{-3}$ | $4.40 \times 10^{-3}$ | $3.77 \times 10^{-3}$ | $3.13 \times 10^{-3}$ | $2.75 \times 10^{-3}$ |
| | $\eta_*$ | $-6.13 \times 10^{-3}$ | $2.52 \times 10^{-2}$ | $-2.27 \times 10^{-2}$ | $1.50 \times 10^{-2}$ | $-3.06 \times 10^{-2}$ | $1.13 \times 10^{-2}$ |
| | $\xi_*$ | $4.15 \times 10^{-3}$ | $7.43 \times 10^{-3}$ | $7.63 \times 10^{-3}$ | $4.58 \times 10^{-3}$ | $8.86 \times 10^{-3}$ | $4.09 \times 10^{-3}$ |
| $\log_{10} \epsilon_*$, $\eta_*$, $\xi_*$ | $\log_{10} \epsilon_*$ | $-6.58$ | $2.11$ | $-6.38$ | $2.10$ | $-6.44$ | $2.12$ |
| | $\eta_*$ | $-3.86 \times 10^{-2}$ | $2.03 \times 10^{-2}$ | $-3.90 \times 10^{-2}$ | $1.16 \times 10^{-2}$ | $-4.33 \times 10^{-2}$ | $1.00 \times 10^{-2}$ |
| | $\xi_*$ | $1.04 \times 10^{-2}$ | $9.07 \times 10^{-3}$ | $1.04 \times 10^{-2}$ | $6.19 \times 10^{-3}$ | $1.21 \times 10^{-2}$ | $5.64 \times 10^{-3}$ |

TABLE V: Mean values and 68% CL obtained for the relevant parameters of each model, after marginalizing over all other parameters. The designations of the models are as listed in section III C; we also specify whether a log or a uniform prior was assumed on $\epsilon_*$.

[1] A. H. Guth, Phys.Rev. **D23**, 347 (1981).
[2] A. D. Linde, Phys.Lett. **B108**, 389 (1982).
[3] D. Baumann, (2009), 0907.5424.
[4] D. H. Lyth and A. Riotto, Phys.Rept. **314**, 1 (1999), hep-ph/9807278.
[5] C. L. Reichardt *et al.*, Astrophys. J. **694**, 1200 (2009), 0801.1491.
[6] J. W. Fowler *et al.*, Astrophys. J. **722**, 1148 (2010), 1001.2934.
[7] SDSS, K. N. Abazajian *et al.*, Astrophys. J. Suppl. **182**, 543 (2009), 0812.0649.
[8] Planck Collaboration, (2006), astro-ph/0604069.
[9] M. S. Turner, Phys.Rev. **D48**, 5539 (1993), astro-ph/9307035.
[10] E. J. Copeland, E. W. Kolb, A. R. Liddle, and J. E. Lidsey, Phys.Rev. **D48**, 2529 (1993), hep-ph/9303288.
[11] E. J. Copeland, E. W. Kolb, A. R. Liddle, and J. E. Lidsey, Phys.Rev.Lett. **71**, 219 (1993), hep-ph/9304228.
[12] A. R. Liddle and M. S. Turner, Phys.Rev. **D50**, 758 (1994), astro-ph/9402021.
[13] J. E. Lidsey *et al.*, Rev.Mod.Phys. **69**, 373 (1997), astro-ph/9508078.
[14] H. Peiris and R. Easther, JCAP **0610**, 017 (2006), astro-ph/0609003.
[15] H. Peiris and R. Easther, JCAP **0607**, 002 (2006), astro-ph/0603587.
[16] R. Easther and H. Peiris, JCAP **0609**, 010 (2006), astro-ph/0604214.
[17] H. V. Peiris and R. Easther, JCAP **0807**, 024 (2008), 0805.2154.
[18] P. Adshead and R. Easther, JCAP **0810**, 047 (2008), 0802.3898.
[19] A. Muslimov, Class.Quant.Grav. **7**, 231 (1990).
[20] D. Salopek and J. Bond, Phys.Rev. **D42**, 3936 (1990).
[21] D. Salopek and J. Bond, Phys.Rev. **D43**, 1005 (1991).
[22] J. Lidsey, Phys.Lett. **B273**, 42 (1991).
[23] A. R. Liddle, Phys.Rev. **D68**, 103504 (2003), astro-ph/0307286.
[24] L. A. Boyle and P. J. Steinhardt, Phys.Rev. **D77**, 063504 (2008), astro-ph/0512014.
[25] M. B. Hoffman and M. S. Turner, Phys.Rev. **D64**, 023506 (2001), astro-ph/0006321.
[26] W. H. Kinney, Phys.Rev. **D66**, 083508 (2002), astro-ph/0206032.
[27] R. Easther and W. H. Kinney, Phys.Rev. **D67**, 043511 (2003), astro-ph/0210345.
[28] W. H. Kinney, E. W. Kolb, A. Melchiorri, and A. Riotto, Phys.Rev. **D74**, 023502 (2006), astro-ph/0605338.
[29] B. A. Powell and W. H. Kinney, JCAP **0708**, 006 (2007), 0706.1982.
[30] B. A. Powell, K. Tzirakis, and W. H. Kinney, JCAP **0904**, 019 (2009), 0812.1797.
[31] M. Malquarti, S. M. Leach, and A. R. Liddle, Phys.Rev. **D69**, 063505 (2004), astro-ph/0310498.
[32] J. M. Cline and L. Hoi, JCAP **0606**, 007 (2006), astro-ph/0603403.
[33] S. M. Leach, A. R. Liddle, J. Martin, and D. J. Schwarz, Phys.Rev. **D66**, 023515 (2002), astro-ph/0202094.
[34] S. M. Leach and A. R. Liddle, Mon.Not.Roy.Astron.Soc. **341**, 1151 (2003), astro-ph/0207213.
[35] S. M. Leach and A. R. Liddle, Phys.Rev. **D68**, 123508 (2003), astro-ph/0306305.
[36] J. Lesgourgues, A. A. Starobinsky, and W. Valkenburg, JCAP **0801**, 010 (2008), 0710.1630.
[37] J. Hamann, J. Lesgourgues, and W. Valkenburg, JCAP **0804**, 016 (2008), 0802.0505.
[38] F. Finelli, J. Hamann, S. M. Leach, and J. Lesgourgues, JCAP **1004**, 011 (2010), 0912.0522.
[39] M. Kawasaki and T. Sekiguchi, JCAP **1002**, 013 (2010), 0911.5191.
[40] M. J. Mortonson, H. V. Peiris, and R. Easther, Phys.Rev. **D83**, 043505 (2011), 1007.4205.
[41] R. Easther and H. Peiris, (2011), 1112.0326.
[42] F. Feroz and M. Hobson, Mon.Not.Roy.Astron.Soc. **384**, 449 (2008), 0704.3704.
[43] F. Feroz, M. Hobson, and M. Bridges, Mon.Not.Roy.Astron.Soc. **398**, 1601 (2009), 0809.3437.
[44] A. Lewis and S. Bridle, Phys.Rev. **D66**, 103511 (2002), astro-ph/0205436.
[45] A. R. Liddle, P. Parsons, and J. D. Barrow, Phys.Rev. **D50**, 7222 (1994), astro-ph/9408015.
[46] R. T. Cox, Am. J. Phys **14**, 1 (1946).
[47] WMAP Collaboration, E. Komatsu *et al.*, Astrophys.J.Suppl. **192**, 18 (2011), 1001.4538.
[48] W. Valkenburg, J. Hamann, and L. M. Krauss, Phys.Rev. **D78**, 063521 (2008), 0804.3390.
[49] D. Larson *et al.*, Astrophys.J.Suppl. **192**, 16 (2011), 1001.4635.
[50] R. Keisler *et al.*, Astrophys.J. **743**, 28 (2011), 1105.3182.
[51] B. A. Reid *et al.*, Mon. Not. Roy. Astron. Soc. **404**, 60 (2010), 0907.1659.
[52] H. Jeffreys, *Theory of Probability* (Oxford University Press, 1961).
[53] S. Wilks, Ann.Math.Statist. **9**, 60 (1938).
[54] B. A. Reid, L. Verde, R. Jimenez, and O. Mena, JCAP **1001**, 003 (2010), 0910.0008.
[55] A. X. Gonzalez-Morales, R. Poltis, B. D. Sherwin, and L. Verde, (2011), 1106.5052.
[56] H. Akaike, IEE T. Automat. Contr. **19**, 716 (1974).
[57] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, (1992), Cambridge University Press.
[58] J. Skilling, AIP Conference Proceedings of the 24th

Interna- tional Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering **735**, 395 (2004).

[59] T. T. Takeuchi, Astrophys.Space Sci. **271**, 213 (2000), astro-ph/9909324.

[60] A. D. R. Burnham K. P., *Model selecton and multimodel inference, 2nd ed.* (Springer-Verlag, 2002).

[61] A. R. Liddle, Mon.Not.Roy.Astron.Soc. **377**, L74 (2007), astro-ph/0701113.

[62] M. Bridges, F. Feroz, M. Hobson, and A. Lasenby, (2008),

0812.3541.

[63] F. Feroz, K. Cranmer, M. Hobson, R. Ruiz de Austri, and R. Trotta, JHEP **1106**, 042 (2011), 1101.3296, 21 pages, 9 figures, 1 table/ minor changes following referee report. Matches version accepted by JHEP.

[64] Z. Huang, L. Verde, and F. Vernizzi, (2012), 1201.5955.

[65] D. Schlegel, M. White, and D. Eisenstein, (2009), 0902.4680.

[66] Euclid consortium, http://www.euclid-ec.org/.