

## Bayesian model selection analysis of WMAP3

Article (Unspecified)

Parkinson, D., Mukherjee, P. and Liddle, A. R. (2006) Bayesian model selection analysis of WMAP3. *Physical Review D*, 73 (12). pp. 1235231-1235237. ISSN 0556-2821

This version is available from Sussex Research Online: <http://sro.sussex.ac.uk/1104/>

This document is made available in accordance with publisher policies and may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher's version. Please see the URL above for details on accessing the published version.

### **Copyright and reuse:**

Sussex Research Online is a digital repository of the research output of the University.

Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable, the material made available in SRO has been checked for eligibility before being made available.

Copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

# A Bayesian model selection analysis of WMAP3

David Parkinson, Pia Mukherjee, and Andrew R. Liddle

*Astronomy Centre, University of Sussex, Brighton BN1 9QH, United Kingdom*

(Dated: June 20, 2006)

We present a Bayesian model selection analysis of WMAP3 data using our code CosmoNest. We focus on the density perturbation spectral index  $n_S$  and the tensor-to-scalar ratio  $r$ , which define the plane of slow-roll inflationary models. We find that while the Bayesian evidence supports the conclusion that  $n_S \neq 1$ , the data are not yet powerful enough to do so at a strong or decisive level. If tensors are assumed absent, the current odds are approximately 8 to 1 in favour of  $n_S \neq 1$  under our assumptions, when WMAP3 data is used together with external data sets. WMAP3 data on its own is unable to distinguish between the two models. Further, inclusion of  $r$  as a parameter weakens the conclusion against the Harrison–Zel’dovich case ( $n_S = 1$ ,  $r = 0$ ), albeit in a prior-dependent way. In appendices we describe the CosmoNest code in detail, noting its ability to supply posterior samples as well as to accurately compute the Bayesian evidence. We make a first public release of CosmoNest, now available at [www.cosmonest.org](http://www.cosmonest.org).

PACS numbers: 98.80.-k

astro-ph/0605003

## I. INTRODUCTION

The recent three-year results from WMAP [1] have provided the first firm indications that the spectral index of primordial density perturbations,  $n_S$ , differs from the Harrison–Zel’dovich case  $n_S = 1$ . The likelihood function in models with varying  $n_S$  suggests that  $n_S = 1$  is excluded at around three to four sigma, in cosmologies with no significant tensor contribution to the microwave anisotropies. However, the WMAP team stress that their result, based on a chi-squared per degrees of freedom argument, needs to be checked using the more sophisticated technique of Bayesian model selection [2, 3, 4]. That is the aim of the present paper, building on our previous analysis of the WMAP first-year data using our code CosmoNest [5].

We will consider two different scenarios. The first concerns the spectral index alone, under the assumption that there are no primordial gravitational waves (parametrized by the tensor-to-scalar ratio  $r$ ). As there is presently no indication for gravitational waves, this analysis addresses the question of whether  $n_S \neq 1$  should be considered part of the standard cosmological model. Secondly, we consider the plane of slow-roll inflation models parametrized by  $n_S$  and  $r$ , representing the simplest class of inflation models (for an extensive review, see Ref. [6]). This latter analysis determines the extent to which slow-roll inflation models have benefitted from the new data.

## II. BAYESIAN MODEL SELECTION

One of the most important classes of statistical problems in science, and particularly in cosmology, is determining the best fit to data in the case where the underlying model (i.e. the set of parameters to be varied) is unknown. Typically, each parameter represents some physical effect that may influence the data, but one can-

not simply include all possible physical effects simultaneously, as the data may be insufficiently constraining and all parameters become undetermined and biases get introduced [7]. In the Bayesian framework, the solution is model selection statistics, which set up a tension between goodness of fit to the data and model complexity. A model selection statistic does not care about the preferred values of the parameters defining the model, but is a property of the model itself, where here ‘model’ means both a choice of the set of parameters to be varied and the prior ranges for those parameters.

A key application of model selection is to provide a robust criterion for judging when data requires the addition of new parameters. Many of the most pressing questions in contemporary cosmology are of this type, such as whether the dark energy density evolves with redshift, or whether primordial gravitational waves exist. For the present data compilation following the WMAP3 announcement, it is the spectral index which is placed in the most interesting position — does WMAP3 convincingly exclude the possibility that  $n_S$  is precisely unity, as conjectured by Harrison and by Zel’dovich [8] long before the inflationary mechanism was discovered? In model selection terms, does the improved fit that a varying  $n_S$  allows justify its inclusion as an extra variable fit parameter?

One might wonder why we should bother with a model selection analysis of a result which a parameter estimation analysis says is already at three to four sigma level. The answer is that this significance level is exactly where model selection techniques are at their most crucial. It has long been recognized in the statistics community that Bayesian model selection analyses can give results in contradiction with inferences based on ‘number of sigmas’; this is known as Lindley’s ‘paradox’ [9] and is nicely summarized by Trotta [10]. Basically, Bayesian inference is inconsistent with the idea that there is a universal threshold, such as 95%, beyond which results should be seen as

definitive; instead such a threshold should depend both on the data properties and the prior parameter ranges of the models being compared. The Lindley paradox usually manifests itself for results with significance in the range two to four sigma [10], which as it happens is exactly where WMAP3 has placed  $n_S$ .

In a full implementation of Bayesian inference, the key statistic is the Bayesian evidence  $E$  (also known as the marginalized likelihood), which has the literal interpretation of the probability of the data given the model [2, 3, 4]. According to Bayes theorem, it therefore updates the prior model probability to the posterior model probability. It is simply the average of the likelihood over the prior parameter space. Often, the quantity of interest is the ratio of evidences of two models  $M_1$  and  $M_0$ , called the Bayes factor and denoted  $B_{10}$ , which indicates how the relative model probabilities have been updated by the data. The evidence has been exploited in a range of cosmological studies [5, 10, 11].

Computing the evidence is more challenging than calculating parameter uncertainties, as it requires knowledge of the likelihood throughout the prior parameter volume rather than only in the vicinity of its peak. So far brute force methods such as thermodynamic integration, though accurate, have proved to be computationally very intensive [12], while approximate information criterion based methods often lead to results which do not agree and hence can be ambiguous [7, 13]. We have recently developed an implementation of an algorithm due to Skilling known as *Nested Sampling* [14], which we call CosmoNest [5], which is able to carry out such calculations efficiently. It is a Monte Carlo method, but not a Markov chain one. We describe the code extensively in Ref. [5] and in the appendices of this article.

In assessing the significance of a model comparison, a useful guide is given by the Jeffreys' scale [2]. Labelling as  $M_1$  the model with the higher evidence, it rates  $\ln B_{10} < 1$  as 'not worth more than a bare mention',  $1 < \ln B_{10} < 2.5$  as 'substantial',  $2.5 < \ln B_{10} < 5$  'strong' to 'very strong' and  $5 < \ln B_{10}$  as 'decisive'. Note that  $\ln B_{10} = 5$  corresponds to odds of 1 in about 150, and  $\ln B_{10} = 2.5$  to odds of 1 in 13.

### III. APPLICATION TO WMAP3

Throughout we use a data compilation of the WMAP3 TT, TE and EE anisotropy power spectrum data [1], together with higher  $\ell$  CMB temperature power spectrum data from ACBAR [15], CBI [16], VSA [17], and Boomerang 2003 [18], and also matter power spectrum data from SDSS [19] and 2dFGRS [20]. Following the approach of Ref. [21], we use the updated beam error module, and do not marginalize over the amplitude of SZ fluctuations. For the higher  $\ell$  CMB data, we neglect those bands that overlap in  $\ell$  range with WMAP (as in Ref. [1]), so that they can be treated as independent measurements.

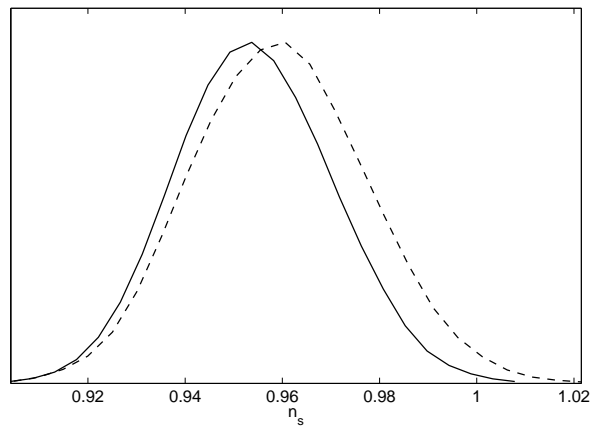


FIG. 1: Marginalized likelihood of  $n_S$  for WMAP alone (dashed) and WMAP+all (solid), obtained using CosmoNest.

The prior ranges for the other parameters were chosen as in Ref. [5]:  $0.018 \leq \Omega_b h^2 \leq 0.032$ ,  $0.04 \leq \Omega_{\text{cdm}} h^2 \leq 0.16$ ,  $0.98 \leq \Theta \leq 1.1$ ,  $0 \leq \tau \leq 0.5$ , and  $2.6 \leq \ln(A_s \times 10^{10}) \leq 4.2$ . Here  $\Theta$  is a measure of the sound horizon at decoupling, and the other symbols have their usual meaning.

When we quote Bayes factors, model  $M_0$  is always taken to be the Harrison-Zel'dovich case. We normalize to this case, which means positive numbers indicate models preferred against this case.

#### A. The spectral index

For the spectral index  $n_S$ , we will throughout assume a prior range  $0.8 < n_S < 1.2$ , as in Refs. [5, 12]. The model selection results presented for  $n_S$  must therefore be understood in light of this prior. As the allowed regions are well contained within this prior, it is trivial to recompute the Bayes factor if this range is extended; e.g. if it is doubled then  $\ln B_{10}$  is reduced by  $\ln 2 \simeq 0.7$ .

Qualitative understanding of our results can be obtained from studying the marginalized distributions for  $n_S$ , shown in Fig. 1, obtained using CosmoNest. For reasons explained later, we did not include marginalization over the SZ effect in obtaining these likelihoods, which shifts them somewhat towards  $n_S = 1$  as compared to those of the WMAP team [1]. A rapid guide to the expected result can be obtained by employing a gaussian approximation to the marginalized posterior distribution for  $n_S$ . As shown by Trotta [10], the Bayes factor can be computed in this approximation as a function of  $\lambda$ , being the 'number of sigmas' of the putative detection, and the information content  $I \equiv \log_{10}(\Delta n_S / \hat{\sigma})$  which measures the reduction of the allowed parameter volume between the prior and posterior (where  $\Delta n_S$  is the prior width and  $\hat{\sigma}$  the standard deviation of the posterior). For WMAP3 alone and our choice of prior  $\lambda \simeq 2.3$  and  $I \simeq 1.4$ . Employing Eq. (18) of Ref. [10] yields the esti-

mate  $\ln B_{10} = 0.4$ , i.e. the varying  $n_S$  model is preferred but only very mildly. However, this expression assumes that the gaussian form holds quite far into its tail, which may not be valid, and so we proceed to results from the full numerical calculation (which we anyway had to do to obtain these marginalized likelihoods as a by-product, as explained in Appendix B).

Using CosmoNest, we ran the Harrison–Zel’dovich (HZ) and spectral index cases ( $n_S$ ) to find the difference in evidence for WMAP3, with and without the external CMB and large-scale structure data sets. We found that using WMAP alone,  $\ln B_{10} = 0.34$ , in agreement with the estimate above. However, when the extra data sets are included,  $\ln B_{10} = 1.99$ , which is substantial but not strong evidence for the necessity of  $n_S$  as an extra parameter. The results are given in Table I.

The difference of 1.65 in  $\ln E$  between the two datasets can be understood simply from the marginalized likelihoods shown in Fig. 1. Although the curves are similar near the peaks, and the maximum likelihood value has shifted only by about 0.005, the difference in the mean and more importantly the variance have a large effect in the tails. The probability of the  $n_S = 1$  value is about 4 times smaller in the WMAP+all case, which would, all else being considered equal, translate into a change of 1.4 in  $\ln E$  explaining most of the difference. Nevertheless, this shift is not particularly significant on the Jeffreys’ scale.

### B. The inflationary plane

A non-zero value of  $n_S - 1$  is commonly interpreted as a strong indication in favour of inflation. However slow-roll inflationary models predict not just a non-zero  $n_S - 1$ , but also a non-zero value for the tensor-to-scalar ratio  $r$ . A proper model comparison motivating inflation should therefore examine not the spectral index model but the two-parameter extension of HZ into the  $n_S$ – $r$  plane. There is no simple way to make an estimate of the outcome in this case.

At this point we run into the issue of the choice of prior for  $r$ . While it is uncontroversial to choose a uniform prior for  $n_S$ , whose value is more or less known,  $r$  is instead a parameter whose order-of-magnitude is currently unknown (sometimes called a ‘scale’ parameter). We will consider two possibilities. The first, Case 1, is that  $r$  has a uniform prior in the range  $[0, 1]$ , which is the assumption used by Spergel et al. [1] for parameter estimation. The second, Case 2, considers the Jeffreys’ prior which states that for scale parameters the prior should be uniform in  $\ln(r)$  rather than  $r$ . For the problem to be well-defined this needs to be cut off at both ends. We use the same upper limit, and as a lower limit take the smallest conceivable inflation scale of the electroweak scale, which would yield  $r \sim \rho^{1/2}/m_{\text{Pl}}^2 \sim 10^{-34}$  (where  $\rho$  is the energy density). The prior range is therefore  $-80 \leq \ln r \leq 0$ . Results are shown in Table I.

Datasets	Model	$\ln B_{10}$
WMAP only	HZ	0.0
	$n_S$	$0.34 \pm 0.26$
WMAP + all	HZ	0.0
	$n_S$	$1.99 \pm 0.26$
	$n_S + r$ (uniform prior)	$-1.45 \pm 0.45$
	$n_S + r$ (log prior)	$1.90 \pm 0.24$

TABLE I: Evidence differences for the different models and different data sets, as discussed in the text.

For Case 1, CosmoNest calculations indicate  $\ln B_{10}$  of  $-1.45$ . The large amount of unused prior parameter space in the  $n_S$ – $r$  plane means that this model is somewhat disfavoured as compared to HZ.

For Case 2, a calculation is not in fact really necessary, since the vast majority of the prior space lies in the region where  $r$  is observationally negligible, and hence generates the same likelihood as a model where the spectral index alone varies. We confirm this explicitly by computing evidence over a limited range  $-5 < \ln r < 0$  and extrapolating the result down to  $\ln r = -80$ .

We conclude therefore that the evidence in the inflationary plane does carry significant prior dependence, bracketted by the values we have found under Case 1 and Case 2. Given the present shape of the likelihood, the evidence for the inflation model will not be as large as for the spectral index model under any prior choice, and may be significantly less. For a uniform prior on  $r$ , the inflation model is actually rated below Harrison–Zel’dovich.

### C. Systematic effects

The evidence computation we have described takes into account only statistical uncertainties. However one should also consider the possible effect of systematic uncertainties, and there are some indications that these are present at a level which would have some impact on our conclusions, despite the very careful job that the WMAP team have done. We highlight some of these issues here.

There is some effect from the precise choice of dataset used. All the dataset combinations quoted in Ref. [1] give very similar constraints on  $n_S$ , though none corresponds precisely to the dataset compilation we are employing. Curiously though, the dataset WMAP+all on the LAMBDA archive,[29] which adds two supernovae datasets to our compilation, gives an  $n_S$  value about one-sigma lower than any other dataset quoted, which would be expected to lead to a stronger result for the Bayes factor. However it is puzzling that this data compilation gives a lower  $n_S$  (and optical depth  $\tau$ ) than do any of the separate datasets from which it is compiled.

There is some uncertainty in how to treat the Sunyaev–Zel’dovich effect and gravitational lensing. The WMAP team allow only for the former, while Lewis has argued

[22] that the two effects are of the same order, and nearly cancel, at least as regards their effect on  $n_S$ , and that it is better to ignore both than to include only one. Accordingly we have not included the SZ correction, which increases  $n_S$  as compared to the WMAP3 analysis.

Another subtlety concerns the modelling of the beam. As discussed in Ref. [21] there are different options for doing this, which appear to have a slight effect on the constraint on  $n_S$ . We have followed the procedure described in that paper, rather than that of the main WMAP3 papers [1].

Yet more uncertainty surrounds the modelling of the recombination process. According to Ref. [23], inclusion of additional two-photon decays leads to significant differences as compared to the standard RECFAST treatment used in the WMAP papers. If confirmed, this is perhaps not too important for WMAP, but would certainly matter at Planck sensitivity (Antony Lewis, private communication).

Also, the reionization optical depth  $\tau$  and  $n_S$  are correlated. The constraint on  $\tau$  comes mainly from the estimate of the power in the low  $\ell$  multipoles of CMB polarization. Substantial foregrounds are present in polarization data, so that their removal using just the frequency information gathered by WMAP can be tricky. Foreground subtraction uncertainties could therefore affect  $\tau$  and hence  $n_S$ .

Finally, we note that the inclusion of Lyman alpha power spectrum data (not used in the WMAP3 papers) seems to have a marginally significant effect. According to the analysis of Ref. [24], inclusion of this data shifts  $n_S$  upwards by around one-sigma while leaving the uncertainty unchanged. Similar results are obtained in Ref. [25] though the trend is less clear as they round their quoted results at the second decimal place.

While individually none of the above would have a very major effect on model selection conclusions, that there are so many clearly urges caution in interpreting a result whose statistical significance remains rather marginal.

#### IV. CONCLUSIONS

We have carried out a Bayesian model selection analysis of WMAP3 data, as advocated by the WMAP team. We have found that WMAP3 data do indeed give support for a varying spectral index when combined with other data, with the Bayes factor compared to the Harrison–Zel’dovich spectrum being approximately  $\ln B_{10} = 2$ . According to the Jeffreys’ scale, this should be regarded as significant, but neither strong nor decisive. It corresponds to probabilistic odds of about 8 to 1 against the Harrison–Zel’dovich model (i.e. the chance that  $n_S$  is equal to one is about that of tossing a coin three times and them all being heads). WMAP3 alone does not provide any discrimination between the models.

In computing our numbers, we have assumed throughout that the prior model probabilities are equal, so that

models are regarded as equally likely before the data came along. Anyone who prefers to make an alternative assumption is welcome to do so, and can readily follow the consequences using the evidence numbers we have supplied. For instance, a perfectly plausible standpoint might be that since inflation is a physical model, its predictions should be taken more seriously than pure HZ which is motivated only by symmetry considerations. Hence its prior model probability should be greater, perhaps tipping the post-data odds decisively against HZ. Readers are quite welcome to take that viewpoint, but should bear in mind that their conclusion then derives from a mixture of the data and their prior prejudice. From the data *alone*, the situation remains to be decisively resolved.

In a companion paper [26], we forecast the abilities of the Planck satellite to resolve this situation, in light of the WMAP3 results.

#### Acknowledgments

The authors were supported by PPARC. We thank Mike Hobson, Martin Kunz, Antony Lewis, Cédric Pahud, Hiranya Peiris, Douglas Scott, John Skilling, and Roberto Trotta for helpful discussions. We acknowledge use of the UK National Cosmology Supercomputer (COSMOS) funded by Silicon Graphics, Intel, HEFCE and PPARC.

#### APPENDIX A: THE NESTED SAMPLING ALGORITHM

Our implementation of the Nested Sampling algorithm is described in Ref. [5]. To summarize, the algorithm (as first developed in Ref. [14]) recasts the problem of calculating the evidence as a one-dimensional integral in terms of the remaining prior mass  $X$ , where  $dX = P(\theta|M)d\theta$ . So the integral is transformed

$$E = \int L(\bar{\theta})p(\bar{\theta})d\bar{\theta} \quad \rightarrow \quad E = \int L(X)dX \quad (\text{A1})$$

where  $L(X)$  is the likelihood  $P(D|\theta, M)$ . The algorithm samples the prior a large number of times, assigning a ‘prior mass’ probabilistically to each sample. The samples are ordered by likelihood, and the integration follows as the sum of the sequence,

$$E = \sum_{j=1}^m E_j, \quad E_j = \frac{L_j}{2}(X_{j-1} - X_{j+1}). \quad (\text{A2})$$

The scheme is illustrated in Figure 2.

In order to compute the integral accurately the prior mass is logarithmically sampled. We start by randomly placing  $N$  points uniformly in the prior space, where in a typical cosmological application  $N \sim 300$ . We then iteratively discard the lowest likelihood point  $L_j$ , replacing it

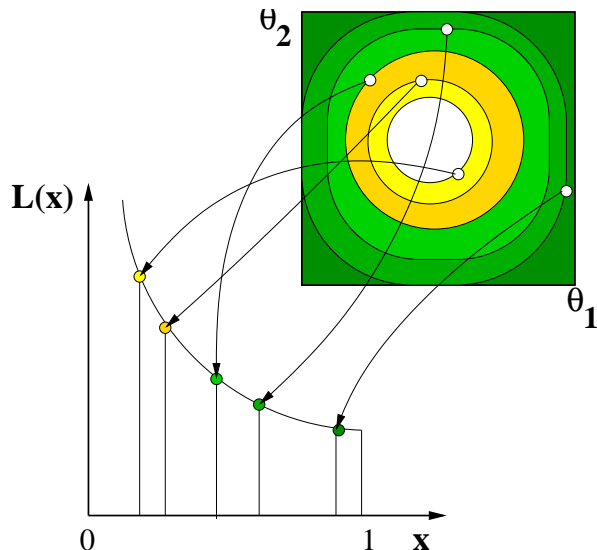


FIG. 2: Schematic of the Nested Sampling algorithm.

with a new point uniformly sampled from the remaining prior mass (i.e. with likelihood  $> L_j$ ). Each time a point is discarded the prior mass remaining  $X_j$  shrinks by a factor that is known probabilistically, and the evidence is incremented accordingly. In this way the algorithm works its way towards the higher likelihood regions.

As the remaining prior mass shrinks by orders of magnitude, the challenging part is to find an efficient way to draw new points from the remaining prior volume. We do this by using the  $N - 1$  remaining points at each stage to define an ellipsoid that encompasses the extremes of the points and is aligned with their principal axes. The ellipsoid is expanded by a constant enlargement factor, in order to allow for the iso-likelihood contours not being exactly elliptical, as well as to take in the edges. New points are then selected uniformly within the expanded ellipse until one has a likelihood exceeding the old minimum.

The process is terminated when the integral has been computed to desired accuracy (see Ref. [5]). In the end the evidence contributed by the  $N - 1$  points remaining is added to the accumulated evidence.

The method is general, and the effects of topology and dimensionality are implicitly built into it.

## APPENDIX B: COSMONEST

The Cosmological Monte Carlo code (CosmoMC) developed by Lewis and Bridle [27] was created to perform an exploration of the cosmological parameter space, through the Monte Carlo Markov Chain process (for an overview of MCMC methods see Ref. [28]). While it is most commonly used with the Metropolis–Hastings algorithm, other sampling algorithms (such as Gibbs Sampling and Slice Sampling) can easily be implemented.

The Nested Sampling algorithm can be considered as just another Monte Carlo sampling algorithm. The important difference is that the generation of a chain, which in this case is not a Markov chain, is ancillary to its primary purpose of calculating the evidence accurately.

The Cosmological Nested Sampling code (CosmoNest) we have developed is an additional module that works as part of CosmoMC.

### 1. Evidence evaluation

CosmoMC has a ‘memory’ of only one point: the algorithm needs only to know where it is in order to decide where to go next. CosmoNest needs to know about the point it is discarding, but must also hold in its memory all the other  $N - 1$  live points, as well as knowing how far through the prior mass  $X$  it has progressed and what value of the evidence ( $E$ ) it has accumulated. The output of a CosmoNest run consists of the set of discarded minimum likelihood points, along with their  $X$  value, their likelihood, and total accumulated evidence to that point.

CosmoMC runs multiple chains for two purposes: increasing the speed of generating samples, and as a way of estimating the extent to which the chains have explored the parameter space (the Gelman–Rubin statistic). Here we run multiple iterations of CosmoNest to obtain an estimate of the uncertainty in the computed evidence.

### 2. Posterior samples

The sequence of discarded points from the Nested Sampling process is similar to the Markov chain produced by an MCMC process with one important difference: the MCMC points are sampled from the posterior whereas the Nested Sampling points are sampled from the prior with a known distribution in  $X$ . With the appropriate weightings, the ‘chain’ of discarded points (distributed uniformly in  $\ln X$ ) plus the remaining live points (distributed uniformly in  $X$  within the remaining volume) can be used to construct the posterior probability distribution of the parameters, as outlined in Ref. [14].

To summarize, from Bayes’ theorem

$$p(\theta|D) = \frac{L(D|\theta)\pi(\theta)}{E(D)}, \quad (\text{B1})$$

where  $p(\theta|D)$  is the posterior probability of a parameter point  $\theta$  given data  $D$ ,  $L$  is the likelihood and  $\pi$  the prior. So for an element  $i$  in the chain of discarded points, the posterior weighting is

$$p_i = \frac{L_i w_i}{E}, \quad (\text{B2})$$

where  $w_i = \frac{1}{2}(X_{i-1} - X_{i+1})$  is the prior mass associated with that particular point. The  $N - 1$  points finally

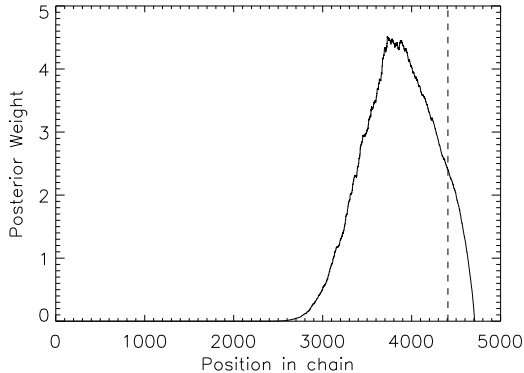


FIG. 3: The posterior weights  $p_i$  assigned to each point in one of our HZ runs. The  $x$ -axis is the element number in the chain, and the vertical dashed line indicates where the live points start to be used. This transition could be shifted to the right by running the code for longer so as to generate a longer chain of discarded points.

remaining also need to be included to avoid undersampling the centre of the distribution. They are taken as uniformly sampling the remainder of the prior space.

Figure 3 shows as an example of the posterior weights assigned in a particular run. The early points have negligible weight as their likelihood is low, and the late ones because the prior mass  $w_i$  per point becomes small. We see that in this case the live points have to be included to properly sample the centre of the distribution. The fractional contribution from live points can be reduced by running the code for longer. The structure of these weights should be contrasted with Metropolis–Hastings where all samples have integer weights (values greater than one accruing when new samples are rejected and instead the original sample duplicated).

Using this method we can reconstruct the posterior samples and compare to similar results from standard Metropolis–Hastings MCMC. We illustrate this in Fig. 4. Posteriors obtained from the two methods are in good agreement.

### 3. The information

The information  $H$  is defined as (minus) the logarithm of the amount the posterior is compressed inside the prior [14]. It is easy to compute from the posterior samples once the evidence has been calculated:

$$H \equiv \int \ln \left( \frac{dP}{dX} \right) dP = \int \frac{E}{L} \ln \left( \frac{E}{L} \right) dX. \quad (\text{B3})$$

The uncertainty on a single estimate of the evidence is dominated by the Poisson uncertainty in the number of steps (replacements) to reach the bulk of the posterior. This is given by  $\sigma^2(E) \approx H \ln[(N+1)/N]$ . For the priors

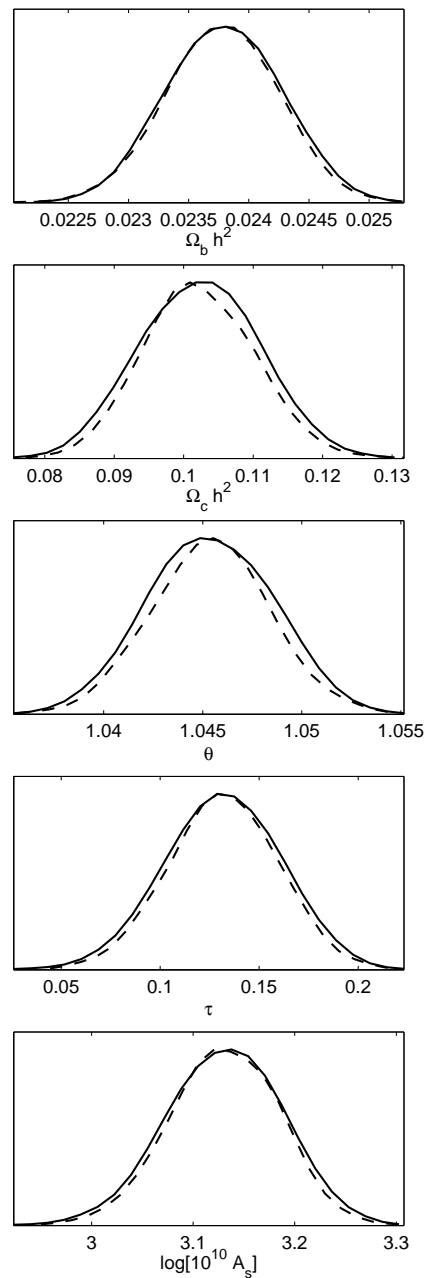


FIG. 4: Posterior samples from Nested Sampling (solid) compared to MCMC (dashed), for a  $\Lambda$ CDM HZ model using WMAP3 data only.

we have been considering  $H \simeq 10$ , and given our choice of  $N$ , this uncertainty turns out to be 0.15 to 0.2.

### 4. Public code release

The CosmoNest code is now freely available for public use, and can be downloaded from [www.cosmonest.org](http://www.cosmonest.org). Its use requires a working installation of the CosmoMC package of Lewis and Bridle [27].

- 
- [1] D. N. Spergel et al. (WMAP Collaboration), [astro-ph/0603449](#); G. Hinshaw et al. (WMAP Collaboration), [astro-ph/0603451](#); L. Page et al. (WMAP Collaboration), [astro-ph/0603450](#).
- [2] H. Jeffreys, *Theory of Probability*, 3rd ed, Oxford University Press (1961).
- [3] D. J. C. MacKay, *Information theory, inference and learning algorithms*, Cambridge University Press (2003).
- [4] P. Gregory, *Bayesian Logical Data Analysis for the Physical Sciences*, Cambridge University Press (2005).
- [5] P. Mukherjee, D. Parkinson, and A. R. Liddle, *Astrophys. J. Lett.* **638**, L51 (2006), [astro-ph/0508461](#).
- [6] A. R. Liddle and D. H. Lyth, *Cosmological inflation and large-scale structure*, Cambridge University Press (2000).
- [7] A. R. Liddle, *Mon. Not. Roy. Astron. Soc.* **351**, L49, [astro-ph/0401198](#).
- [8] R. Harrison, *Phys. Rev.* **D1**, 2726 (1970); Ya. B. Zel'dovich, *Mon. Not. Roy. Astron. Soc.* **160**, 1p (1972).
- [9] D. Lindley, *Biometrika* **44**, 187 (1957).
- [10] R. Trotta, [astro-ph/0504022](#).
- [11] A. Jaffe, *Astrophys. J.* **471**, 24 (1996), [astro-ph/9501070](#); P. S. Drell, T. J. Loredo, and I. Wasserman, *Astrophys. J.* **530**, 593 (2000), [astro-ph/9905027](#); M. V. John and J. V. Narlikar, *Phys. Rev.* **D65**, 043506 (2002), [astro-ph/0111122](#); A. Slosar et al., *Mon. Not. Roy. Astron. Soc.* **341**, L29 (2003), [astro-ph/0212497](#); T. D. Saini, J. Weller, and S. L. Bridle, *Mon. Not. Roy. Astron. Soc.* **348**, 603 (2004), [astro-ph/0305526](#); P. J. Marshall, M. P. Hobson, and A. Slosar, *Mon. Not. Roy. Astron. Soc.* **346**, 489 (2003), [astro-ph/0307098](#); A. Niarchou, A. H. Jaffe, and L. Pogosian, *Phys. Rev.* **D69**, 063515 (2004), [astro-ph/0308461](#); B. A. Bassett, P. S. Corasaniti and M. Kunz, *Astrophys. J. Lett.* **617**, L1 (2004), [astro-ph/0407364](#); M. Bridges, A. N. Lasenby, and M. P. Hobson, [astro-ph/0511573](#); P. Mukherjee, D. Parkinson, P. S. Corasaniti, A. R. Liddle, and M. Kunz, [astro-ph/0512484](#); M. Kunz, R. Trotta, and D. Parkinson, [astro-ph/0602378](#).
- [12] M. Beltrán, J. Garcia-Bellido, J. Lesgourgues, A. R. Liddle, and A. Slosar, *Phys. Rev.* **D71**, 063532 (2005), [astro-ph/0501477](#).
- [13] J. Magueijo and R. D. Sorkin, [astro-ph/0604410](#).
- [14] J. Skilling, in *Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, ed. R. Fischer et al., Amer. Inst. Phys., conf. proc., 735, 395 (2004), (available at <http://www.inference.phy.cam.ac.uk/bayesys/>).
- [15] C. L. Kuo et al., *Astrophys. J.* **600**, 32 (2004), [astro-ph/0212289](#).
- [16] T. J. Pearson et al., *Astrophys. J.* **591**, 556 (2003), [astro-ph/0205388](#).
- [17] C. Dickinson et al., *Mon. Not. Roy. Astron. Soc.* **353**, 732 (2004), [astro-ph/0402498](#).
- [18] W. C. Jones et al., [astro-ph/0507494](#).
- [19] M. Tegmark et al., *Astrophys. J.* **606**, 702 (2004), [astro-ph/0310725](#).
- [20] W. Percival et al., *Mon. Not. Roy. Astron. Soc.* **327**, 1297 (2001), [astro-ph/0105252](#).
- [21] H. Peiris and R. Easther, [astro-ph/0603587](#).
- [22] A. Lewis, [astro-ph/0603753](#).
- [23] V. K. Dubrovich and S. I. Grachev, *Astron. Lett.* **31**, 359 (2005), [astro-ph/0501672](#).
- [24] U. Seljak, A. Slosar, and P. McDonald, [astro-ph/0604335](#).
- [25] M. Viel, M. G. Haehnelt, and A. Lewis, [astro-ph/0604310](#).
- [26] C. Pahud, A. R. Liddle, P. Mukherjee, and D. Parkinson, [astro-ph/0605004](#).
- [27] A. Lewis and S. Bridle, *Phys. Rev. D* **66**, 103511 (2002), [astro-ph/0205436](#), code available from <http://cosmologist.info/cosmomc>.
- [28] W. R. Gilks, S. Richardson, and D. J. Spiegelhalter, *Markov Chain Monte Carlo in Practice*, Chapman and Hall, London (1996); D. S. Sivia, *Bayesian Data Analysis*, Clarendon Press, Oxford (1996).
- [29] <http://lambda.gsfc.nasa.gov>