Mon. Not. R. Astron. Soc. **410,** 2488–2496 (2011) doi:10.1111/j.1365-2966.2010.17614.x

Should we doubt the cosmological constant?

M. C. March,[1] G. D. Starkman,[2] R. Trotta and P. M. Vaudrevange[2★]
[1] *Imperial College London, Astrophysics Group, Blackett Laboratory, Prince Consort Road, London SW7 2AZ*
[2] *CERCA/ISO & Department of Physics, Case Western Reserve University, 10900 Euclid Avenue, Cleveland, OH 44106, USA*

Accepted 2010 September 1. Received 2010 August 31; in original form 2010 May 26

ABSTRACT
While Bayesian model selection is a useful tool to discriminate between competing cosmological models, it only gives a relative rather than an absolute measure of how good a model is. Bayesian doubt introduces an unknown benchmark model against which the known models are compared, thereby obtaining an absolute measure of model performance in a Bayesian framework. We apply this new methodology to the problem of the dark energy equation of state, comparing an absolute upper bound on the Bayesian evidence for a presently unknown dark energy model against a collection of known models including a flat Lambda cold dark matter (ΛCDM) scenario. We find a strong absolute upper bound to the Bayes factor B between the unknown model and ΛCDM, giving $B \lesssim 5$. The posterior probability for doubt is found to be less than 13 per cent (with a 1 per cent prior doubt) while the probability for ΛCDM rises from an initial 25 per cent to almost 70 per cent in light of the data. We conclude that ΛCDM remains a sufficient phenomenological description of currently available observations and that there is little statistical room for model improvement.

Key words: methods: statistical – supernovae: type Ia – cosmological parameters – dark energy.

1 INTRODUCTION

One of the most important questions in cosmology is to identify the fundamental model underpinning the vast amount of observations nowadays available. The so-called 'cosmological concordance model' is based on the cosmological principle (i.e. the Universe is isotropic and homogeneous, at least on large enough scales) and on the hot big bang scenario, complemented by an inflationary epoch. This remarkably simple model is able to explain with only half a dozen free parameter observations spanning a huge range of time- and length-scales. Since both a cold dark matter (CDM) and a cosmological constant (Λ) component are required to fit the data, the concordance model is often referred to as 'the ΛCDM model'. It is, however, important to keep in mind that at this stage the ΛCDM model is not a model in the sense attributed to the word by particle physicists, but rather a phenomenological scenario that appears to be able to explain the vast majority of observations with a great economy of free parameters.

In the classical approach to statistics, models (or hypotheses) can never be proved true, only falsified. Popper (2002), for example, argued that theories always remain 'infinitely improbable' regardless of the amount of evidence gathered in their favour. However, in the context of Bayesian inference support can be accrued for a model if the observed data verify predictions made by the model but not by competing models (see Jaynes 2003). This is the subject of Bayesian model selection (see e.g. Trotta 2007; Trotta 2008 for applications to the cosmological context): given a set of competing models, the Bayes factor gives a measure of the relative performance of each model in explaining the data. This programme naturally prefers models that provide a good fit with the fewest number of free parameters, thus implementing a quantitative version of Occam's razor.

Although Bayesian model selection can identify the best model in a given set of known models, it has no way of indicating whether the absolute quality of the preferred model is high or low. However, it seems desirable to be able to gauge the absolute performance of a model in a Bayesian sense, rather than just its relative performance with respect to known competitors. In particular, this seems crucial for deciding whether the set of known models includes the true model.

The purpose of this paper is to build on the notion of Bayesian doubt introduced by Starkman, Trotta & Vaudrevange (2008) to develop and apply a Bayesian technique for model discovery, focusing in particular on the nature of dark energy. The structure of this paper is as follows. In Section 2 we recall the notion of Bayesian doubt and introduce a new procedure for estimating an upper bound for the Bayes factor in favour of doubt. We next employ this procedure in Section 3 to assess the state of our knowledge of the dark energy equation of state, focusing on the status of the current ΛCDM concordance model. We present our results in terms of the posterior probability for doubt and for ΛCDM in Section 4 and discuss our conclusions in Section 5.

2 BAYESIAN MODEL DISCOVERY

In this section we review the concept of Bayesian doubt and explain how this can lead to model discovery.

★E-mail: r.trotta@imperial.ac.uk

2.1 The notion of Bayesian doubt

Bayesian doubt, as introduced by Starkman et al. (2008), is an extension of Bayesian model selection. It seeks to determine a scale quantifying the absolute quality of a model, as opposed to the relative performance of two models, given by their Bayes factor. The key idea of Bayesian doubt is that the general statistical characteristics of what would be recognized as a 'good' model are known, even if the specifics of the model are not.

We begin by introducing a hypothetical unknown model X which has the characteristics of what would be considered a good model, to be defined below. This idealized good model then acts as a benchmark against which known models can be compared using standard Bayesian model selection. Following Starkman et al. (2008), we define 'doubt', $\mathcal{D}$, as the posterior probability of this unknown model:

$$\mathcal{D} \equiv p(X|d) = \frac{p(d|X)p(X)}{p(d)} = \left(1 + \frac{\sum_i p(d|\mathcal{M}_i)p(\mathcal{M}_i)}{p(d|X)p(X)}\right)^{-1} \quad (1)$$

where $\{\mathcal{M}_i\}$ $(i = 1, \ldots, N)$ is the set of N known models and d are the data. In the above expression, $p(X)$ is the prior probability for the model X, in other words, the prior probability that our list of known models does not contain the true model. $p(\mathcal{M}_i)$ is the prior probability of model $\mathcal{M}_i$ and $p(d|\mathcal{M}_i)$ is the Bayesian evidence for model $\mathcal{M}_i$, given by

$$p(d|\mathcal{M}_i) = \int \mathrm{d}\theta_i \, p(d|\theta_i, \mathcal{M}_i)p(\theta_i|\mathcal{M}_i), \quad (2)$$

where θ_i are the parameters of model $\mathcal{M}_i$. $p(d|\theta_i, \mathcal{M}_i)$ is the likelihood function for model $\mathcal{M}_i$, and $p(\theta_i|\mathcal{M}_i)$ is the prior probability of the parameters of model $\mathcal{M}_i$.

Once we have chosen the level of prior doubt by defining the value of $p(X)$, based on a principle of indifference we assume for simplicity that the prior probabilities for the known N models are all equal, i.e.

$$p(\mathcal{M}_i) = \frac{1}{N}\left[1 - p(X)\right]. \quad (3)$$

We single out the ΛCDM model as one of the set of known models, and, looking ahead, refer to it as our baseline model. Therefore it is useful to rewrite equation (1) as

$$\mathcal{D} = \left(1 + \frac{\langle B_{i\Lambda}\rangle}{B_{X\Lambda}}\left(\frac{1 - p(X)}{p(X)}\right)\right)^{-1}, \quad (4)$$

where we have introduced the Bayes factor

$$B_{ij} \equiv \frac{p(d|\mathcal{M}_i)}{p(d|\mathcal{M}_j)} \quad (5)$$

and the average Bayes factor between ΛCDM and each of the known models

$$\langle B_{i\Lambda}\rangle \equiv \frac{1}{N}\sum_{j=1}^{N} B_{j\Lambda}. \quad (6)$$

(Note that the sum over models $\mathcal{M}_j$ includes $j = \Lambda$ and therefore $\langle B_{i\Lambda}\rangle \geqslant 1/N$.)

Rather than looking at $\mathcal{D}$ directly, one can also consider the relative change in doubt $\mathcal{R}$, given by the ratio of posterior to prior doubt:

$$\mathcal{R} \equiv \frac{\mathcal{D}}{p(X)} = \left(p(X) + (1 - p(X))\frac{\langle B_{i\Lambda}\rangle}{B_{X\Lambda}}\right)^{-1}. \quad (7)$$

A necessary condition for doubt to grow ($\mathcal{R} > 1$) is

$$\frac{\langle B_{i\Lambda}\rangle}{B_{X\Lambda}} \ll 1, \quad (8)$$

i.e. that the Bayes factor between model X and ΛCDM be much larger than the average Bayes factor between the known models and ΛCDM.

However, for ΛCDM to be genuinely doubted it is not sufficient that $\mathcal{R} > 1$. One has also to require that the probability for ΛCDM itself decreases, i.e. that $p(\Lambda|d) < p(\Lambda)$. Applying again Bayes theorem, one finds that the ratio of the posterior probability for ΛCDM to its prior probability is given by

$$\mathcal{R}_\Lambda \equiv \frac{p(\Lambda|d)}{p(\Lambda)} = \left((1 - p(X))\langle B_{i\Lambda}\rangle + p(X)B_{X\Lambda}\right)^{-1}. \quad (9)$$

Hence to gather genuine doubt against ΛCDM we require that both conditions $\mathcal{R} > 1$ and $\mathcal{R}_\Lambda < 1$ be fulfilled.

2.2 Upper bound on the evidence of the unknown model

In order to apply Bayesian doubt to the problem of cosmological model selection, it is necessary to estimate the evidence of the unknown model, $p(d|X)$. The approach suggested by Starkman et al. (2008) was to calibrate the value of $p(d|X)$ on simulated data sets from the best among the known models. This has been shown to lead to model discovery for a toy linear model. However, in the cosmological context it would be computationally very expensive to implement, even given fast algorithms to compute the evidence, such as MultiNest (Feroz, Hobson & Bridges 2009), Population Monte Carlo (Kilbinger et al. 2010), a modification of the VEGAS algorithm (Serra, Heavens & Melchiorri 2007) or the Savage–Dickey density ratio (SDDR; Trotta 2007).

In this paper, we put forward a different, more economical approach, which aims at computing an absolute upper bound for $p(d|X)$. Since our aim is to investigate the dark energy sector, in the following we focus on the dark energy equation of state, $w(z)$. We cannot, of course, compute the evidence for X explicitly since its parametrization of $w(z)$ is unspecified. Since the unknown model X is to provide a benchmark value for the evidence of the known models, it should be designed to provide a good fit to the available data, including cosmic microwave background (CMB), matter power spectrum (mpk) and Type Ia supernovae (SNIa) observations. Therefore, the unknown model should have a high degree of flexibility. At the same time, we do not wish to incur the Occam's razor penalty coming from the high number of free parameters usually associated with a very flexible model. This is because we are seeking to build a phenomenological description for $w(z)$ which, if model X is to be a 'good' model, should arise from an underlying, presently unknown theory with a small number of free parameters.

In order to have the advantages of a flexible (and therefore well-fitting) unknown model (i.e. low χ^2/d.o.f.), without incurring a penalty for having a large number of free parameters, we define the evidence of the unknown model via the upper bound on the Bayes factor between the ΛCDM baseline model and a stand-in model with a very flexible $w(z)$, i.e. a model that allows a large number of degrees of freedom in the dark energy sector that will be used as a substitute for an unknown physical model (see Section 3.2). The absolute upper bound on the Bayes factor $B_{X\Lambda}$ between the unknown model X and ΛCDM (denoted by a subscript Λ) is given by (see Gordon & Trotta 2007 and references therein for details),

$$B_{X\Lambda} < \overline{B}_{X\Lambda} = \exp\left(-\frac{1}{2}\left(\chi^2_X - \chi^2_\Lambda\right)\right). \quad (10)$$

We have defined the best-fitting chi-squared as −2 the log-likelihood at the best-fitting point, θ_i^*:

$$\chi_i^2 = -2 \ln p(d|\theta_i, \mathcal{M}_i)\Big|_{\theta_i=\theta_i^*}, \quad (11)$$

where $i = X, \Lambda$.

The bound of equation (10) arises by putting a posteriori the prior probability for the parameters of the stand-in model into a delta-function located at the observed maximum likelihood value, i.e. by replacing $p(\theta_X|\mathcal{M}_X)$ in equation (2) with $\delta(\theta_X - \theta_X^*)$. While this prior choice has no Bayesian justification (for it is inappropriate to use a posteriori information to determine the prior), it does lead to an absolute upper bound on the relative evidence between the baseline ΛCDM model and the unknown model. In order to calculate the absolute bound of equation (10), all that is needed is the difference between the best-fitting log-likelihood (or chi-squared) of the two models, $\Delta\chi^2 \equiv \chi_X^2 - \chi_\Lambda^2$, which can be easily computed. Since the ΛCDM model is nested within the unknown model (i.e. the unknown model reverts to ΛCDM if the maximum likelihood value from the data is such that $w(z) = -1$ at all redshifts for the unknown model), it follows that $\Delta\chi^2 \leqslant 0$. Therefore it is clear that by construction $\overline{B}_{X\Lambda} > 1$ always, i.e. that our unknown model is always at least as good as ΛCDM.

By inspecting equation (10), one might be tempted to think that this upper bound on the Bayes factor merely translates in Bayesian terms the old goodness-of-fit χ^2 test. For if ΛCDM is a 'bad' model (on whatever scale one wishes to define this), the value of χ^2_Λ will be large and thus the Bayes factor in favour of the unknown model will be large, as well. Thus one might think that equation (10) simply rephrases the well-known rule of thumb of χ^2/d.o.f. ∼ 1. However, this is not the case, for the χ^2/d.o.f. ∼ 1 rule only applies asymptotically (for $n \to \infty$ number of data points) and only if the data points are independent, Gaussian distributed. Those conditions are almost invariably *not* met in the cosmological context. For instance, it is not even clear how one would define the concept of degrees of freedom for the CMB data, given that the C_ℓ values are not independent and are not Gaussian distributed. In the case of SNIa observations, the χ^2/d.o.f. ∼ 1 criterion is satisfied for ΛCDM by construction, for the value of the intrinsic dispersion for the SNe is adjusted in such a way to *require* this to be the case, see e.g. Kowalski et al. (2008). Therefore one cannot meaningfully use this kind of absolute goodness-of-fit tests on such a data set.

Instead, the upper bound given by equation (10) does not require any assumption about asymptotic behaviour, nor that the data are Gaussian distributed, nor independent. One only needs to be able to compute the log-likelihood at the best-fitting point, including relevant correlations as necessary.

Finally, the upper bound of equation (10) could also be computed using the highest best-fitting log-likelihood of all the known models, at no extra computational cost. This would give the absolute upper bound achievable among the class of known models. Although we do not pursue this approach in this paper, we recommend including in any Bayesian model comparison a model X with evidence obtained via this procedure, for this will give an estimate of the maximum possible level of doubt that can arise from the known models with their assigned priors.

2.3 Behaviour of doubt and posterior probability for ΛCDM

In the following, we will adopt the absolute upper bound $\overline{B}_{X\Lambda}$ of equation (10) as an estimator for the Bayes factor of the unknown model X, and explore the consequence in terms of doubt and in terms of the posterior probability for ΛCDM. It is clear from equations (4) and (9) that for a given level of prior doubt $p(X)$, the posterior models' probabilities are controlled uniquely by the two quantities $\langle B_{i\Lambda}\rangle$ and $\overline{B}_{X\Lambda}$. The result can be expected to fall within one of the three scenarios below, which we will examine from two points of view: using doubt $\mathcal{D}$ and using the upper bound to the Bayes factor $\overline{B}_{X\Lambda}$ as measures of doubt. While there is something to be said for employing $\overline{B}_{X\Lambda}$ (whose value can be translated into a strength of evidence via the Jeffreys' scale, given in Table 1) as a criterion for goodness of fit, it turns out that doubt can shed some light on to how large $\overline{B}_{X\Lambda}$ should be to have genuine doubt without referring to the (in some sense arbitrary calibrated) Jeffreys' scale.

Table 1. Empirical scale for evaluating the strength of evidence from the Bayes factor B_{ij} between two models (so-called 'Jeffreys' scale'). The right most column gives our convention for denoting the different levels of evidence above these thresholds, following Gordon & Trotta (2007).

$\|\ln B_{ij}\|$	Odds	Strength of evidence
<1.0	$\lesssim$3: 1	Inconclusive
1.0	∼3: 1	Weak evidence
2.5	∼12: 1	Moderate evidence
5.0	∼150: 1	Strong evidence

(i) ***Case 1*:** $\overline{B}_{X\Lambda} \gg 1$ and $\langle B_{i\Lambda}\rangle \sim 1$: in this case, the unknown model has a much better evidence than ΛCDM, which in turn has about the same evidence as the other known models. As the Bayes factor $\overline{B}_{X\Lambda} > 1$, we should expect there to be a significant amount of doubt, $\mathcal{D} \approx 1$. And indeed, from equation (1) the doubt is, assuming $p(X) \ll 1$

$$\mathcal{D} \approx \left(1 + \frac{1}{p(X)\overline{B}_{X\Lambda}}\right)^{-1} \approx 1, \quad (12)$$

for $p(X)\overline{B}_{X\Lambda} \gg 1$. In other words, we are inclined to believe that there is a better model that we have not yet thought of if the Bayes factor between the unknown model and ΛCDM is sufficiently large to override the smallness of the prior doubt, $\overline{B}_{X\Lambda} > 1/p(X)$ (note the independence of the Jeffreys' scale). The change in the probability for ΛCDM itself is given by, from equation (9),

$$\mathcal{R}_\Lambda \approx \left(1 + p(X)\overline{B}_{X\Lambda}\right)^{-1}. \quad (13)$$

While the doubt grows ($\mathcal{D} \to 1$) the probability for ΛCDM declines, $\mathcal{R}_\Lambda \ll 1$. In this case, one is led to genuinely doubt ΛCDM.

(ii) ***Case 2*:** $\overline{B}_{X\Lambda} \gg 1$ and $B_{i\Lambda} \ll 1 (i \neq \Lambda)$: in this case, ΛCDM is clearly the best of the known models, as the Bayes factors between the known models and ΛCDM are all small. Again, as the Bayes factor $\overline{B}_{X\Lambda} \gg 1$ favours the unknown model, we should be doubting our list of models. As $\langle B_{i\Lambda}\rangle \approx 1/N$, we find

$$\mathcal{D} \approx \left(1 + \frac{1}{Np(X)\overline{B}_{X\Lambda}}\right)^{-1} \approx 1 \quad (14)$$

for $p(X)\overline{B}_{X\Lambda} \gg 1/N$. This seems to contradict the result of Case 1. However, as we noted above, the condition that $\mathcal{D} \approx 1$ is only necessary but not sufficient for doubt to arise. We need to examine the relative change in probability for ΛCDM which is given by

$$\mathcal{R}_\Lambda \approx \left(\frac{1}{N} + p(X)\overline{B}_{X\Lambda}\right)^{-1}. \quad (15)$$

Requiring $\mathcal{R}_\Lambda < 1$ leads to the stronger condition $p(X)\overline{B}_{X\Lambda} \gg 1$, as in Case 1. If the latter condition is not fulfilled, doubt will grow at the expense of the probability of the other known models, as the prior probability mass which was spread among the N known models according to equation (3) gets redistributed between X and ΛCDM.

(iii) ***Case 3*:** $\overline{B}_{X\Lambda} \sim 1$: in this case, the upper bound on the Bayes factor between the unknown model and ΛCDM is of order unity. This means that we should have no reason to doubt our set of models. The expression for doubt equation (1) simplifies to

$$\mathcal{D} \approx \left(1 + \frac{\langle B_{i\Lambda}\rangle}{p(X)}\right)^{-1}. \quad (16)$$

In order to reach a high level of doubt $\mathcal{D} \approx 1$, we would need $\langle B_{i\Lambda}\rangle / p(X) \approx 0$. Clearly, this is only the case if we allow for $p(X) \gg \langle B_{i\Lambda}\rangle \leqslant 1/N$, i.e. if we are starting off with a prior doubt which is larger than the indifference prior on the known models, which is usually not the case. Otherwise, if the Bayes factor $\overline{B}_{X\Lambda}$ is larger than the prior doubt $p(X)$, we can regard our list of models as reasonably complete, and perform Bayesian model comparison among the list of known models. Of course, this procedure must be repeated once new data arrive (see Starkman, Trotta & Vaudrevange 2009 for the procedure that this entails). Note that again we do not need to refer to the Jeffreys' scale, but need to compare the average Bayes factor $\langle B_{i\Lambda}\rangle$ with our prior doubt $p(X)$.

In summary, we are led to doubt the current baseline ΛCDM model only if the rule of thumb

$$p(X)\overline{B}_{X\Lambda} \gg 1 \quad (17)$$

is satisfied, which corresponds to either Case 1 or Case 2 when the condition for $\mathcal{R}_\Lambda < 1$ is also fulfilled. If equation (17) is satisfied, we are guaranteed that doubt will grow and at the same time the probability for the ΛCDM model will decrease, thus signaling the opportunity for model discovery. All this is accomplished without referring to Jeffreys' scale.

3 APPLICATION OF DOUBT TO THE DARK ENERGY EQUATION OF STATE

3.1 The known models

We take the flat ΛCDM model as our baseline model, described by the usual six-parameter set $\theta = \{A_s, n_S, \omega_b, \omega_c, \Omega_\Lambda, H_0\}$, where A_s is the amplitude of scalar fluctuations, n_S is the spectral index, ω_b the physical baryon density, ω_c the cold dark matter density, Ω_Λ the density parameter for the cosmological constant and H_0 the Hubble constant today. We assume purely adiabatic fluctuations throughout this paper.

We define the other models in the known models list by increasing the complexity of the baseline model in successive steps. First, we add only a non-zero curvature parameter, $\Omega_\kappa \neq 0$, with a flat prior in the range $-1.0 \leqslant \Omega_\kappa \leqslant 1.0$. This is the 'astronomer's prior', adopted and justified in Vardanyan, Trotta & Silk (2009). Alternatively, another model is obtained by adding only an effective equation of state parameter for dark energy, $w \neq -1$, with a flat prior in the range $-1.3 \leqslant w \leqslant -0.33$ while keeping $\Omega_\kappa = 0$ fixed. This choice of the prior range is motivated by the fact that w must be $<-1/3$ to achieve an accelerated expansion. Also, while current data strongly constrain models with a phantom equation of state ($w < -1$), it seems too harsh to completely exclude this class of models from the prior, and therefore we extend the lower range somewhat into the phantom domain. We emphasize that this is a somewhat subjective choice – other equally motivated choices are possible, and indeed would effectively constitute a new choice of model. We comment in Section 4.2 how our conclusions would change if one were to add further possibilities (including different choices for the prior on w) to the list of known models. Finally, a fourth model with eight free parameters is obtained by adding both $\Omega_\kappa \neq 0$ and $w \neq -1$ with the above priors to the ΛCDM baseline model.

It is also possible to estimate analytically how our results would change under a modification of the prior ranges defined above for Ω_κ and w. This is useful to assess the robustness of our conclusions under a reasonable change of prior. As the ΛCDM model is nested within the three models with Ω_κ and/or w free, the Bayes factor between ΛCDM and each of the more complex models is given by the SDDR formula (see Trotta 2007 for details):

$$B_{\Lambda j} = \frac{p(\theta = \theta_0 | d, \mathcal{M}_j)}{p(\theta = \theta_0 | \mathcal{M}_j)}, \quad (18)$$

where θ is either Ω_κ or w and θ_0 is the value of the extra parameters for which the more complex model reverts to ΛCDM, i.e. $\theta_0 = 0$ for $\theta = \Omega_\kappa$ and $\theta_0 = -1$ for $\theta = w$. From the SDDR expression, it is clear that the Bayes factor is directly proportional to the prior width on the extra parameter, as for a flat prior $p(\theta = \theta_0 | \mathcal{M}_j) = 1/\Delta\theta$, and $\Delta\theta$ is the chosen prior range in the variable considered (Ω_κ or w). Therefore, under a change of the prior range (provided the support of the likelihood function is well within the prior), the log of the Bayes factor scales proportionally to the log of the prior width, with a larger range for the extra parameters favouring the ΛCDM model due to the Occam's razor effect. As a consequence, to achieve a significant shift (say, $\Delta \ln B \sim 2.5$) on the Jeffreys' scale the prior range would have to be rescaled by a factor $\exp(2.5) \sim 12$, which is clearly unreasonable for both Ω_κ and w. Therefore we expect our results to be fairly robust with respect to reasonable changes in the prior ranges chosen.

One could in principle further increase the complexity of the known models, e.g. by adopting more complex descriptions for $w(z)$, such as the so-called Chevallier–Polarski–Linder (CPL) parametrization in terms of the parameters (w_0, w_a). However, those models have in general a lower evidence than ΛCDM, as they are penalized for their wasted parameter space, see e.g. Liddle et al. (2006b) and Kilbinger et al. (2010). As a consequence, they are expected not to contribute significantly to $\langle B_{i\Lambda}\rangle$, and therefore their influence on posterior doubt would be minor, see Section 4.2 for details. One could also add to the list alternative explanations for the apparent acceleration of the Universe, such as for example modified gravity models, provided one can compute their evidence numerically (Heavens, Kitching & Verde 2007). As the main goal of this paper is to introduce the methodology related to Bayesian doubt, we however restrict our considerations to the four models listed above.

Finally, in this work we do not address the problem of the fine tuning of the value of the cosmological constant itself. All models we consider here suffer equally from the fine-tuning problem, i.e. the fact that the measured value of the cosmological constant is some 120 orders of magnitude smaller than the 'natural' scale set by the Planck mass if Λ arises from quantum fluctuations of the vacuum. Anthropic reasoning in the context of the Multiverse has been invoked to explain the smallness of the cosmological constant, and while Bayesian reasoning could be brought to bear on the effectiveness of such an 'explanation', we shall not consider this aspect further in this paper.

3.2 Parametrization of the unknown model

Our discussion so far has been completely general, sidestepping the crucial issue of how to evaluate equation (10) for the unknown model. In order to make further progress, we have to make some assumptions regarding the class of alternative models the unknown model X is supposed to come from.

As we are interested in the dark energy sector, we will assume that the phenomenology of model X is such that it only leads to modifications to the right-hand side of Einstein equations. In other words, we do not investigate models that modify general relativity except for those whose only impact is a change in the effective energy–momentum tensor. Under this assumption, a model X is fully specified once we give its redshift-dependent equation of state of dark energy $w(z)$. Note that we also implicitly assume that the Universe is well described by a Friedmann–Robertson–Walker (FRW) isotropic cosmology. If one wished to include a more general class of alternative models from which to draw X, one could do so by parametrizing their phenomenology in a suitable way. One could define even more general classes of alternative models, for example by fitting parametrized functions to the observations. However, we do not pursue this approach here, because such a modelling of the data would be devoid of any physical insight and would achieve a purely descriptive fit to the observations. To see why this is not desirable, one only has to push this approach to its extreme consequences: given any data collection, there is always a 'model' that fits the data perfectly. This model is obtained by simply choosing the value of the 'theory' to be identical to the observed value for each of the observations. Of course, nobody would ever consider such a model to be a valid scientific theory, because we demand that the latter should have explanatory power, not be a simple description of the data. Therefore, it seems sensible to require from the outset that our unknown model X be part of a class of *physical* theories, with phenomenological parameters that are linked with the physical framework of the class of models considered (here, FRW isotropic Universes with time-varying dark energy equation of state and otherwise standard cosmology).

Therefore we are left with the task of parametrizing $w(z)$ as a function of redshift, and then use its functional form to compute the $\Delta\chi^2$ between the unknown model and the ΛCDM baseline model. To this purpose, we employ the parametrized post-Friedman (PPF) prescription developed by Hu & Sawicki (2007) and Hu (2008). The PPF prescription was originally introduced to describe the behaviour of theories of modified gravity in a metric framework that describes leading order deviations from general relativity (subject to certain assumptions). However, it was also found be well suited for describing the evolution of dark energy models that cross the so-called 'phantom divide', $w = -1$. Crossing this phantom divide in models with fixed sound speed would lead to divergences in the pressure perturbations. Hence models that are phenomenologically described by a time-varying $w(z)$ that crosses $w = -1$ must be described microphysically by a theory of scalar fields with a varying speed of sound, e.g. DGP-type models.

3.3 Numerical implementation and data sets

Below, we investigate the behaviour of doubt for different combinations of cosmological data sets. In particular, we are interested in studying doubt as the constraining power of the combined data increases.

We modified the COSMOMC (Lewis & Bridle 2002) parameter estimation package to sample the additional parameters $w_i \equiv w(z_i)$, where z_i are uniformly spaced at $n = 10$ redshift value, ranging from $z = 0, \ldots, 1.5$. Fang et al. (2008b) and Fang, Hu & Lewis (2008a) wrote a plugin to CAMB (Lewis, Challinor & Lasenby 2000) that implements the PPF prescription and is freely available for download,[1] which we adopted for this work. The PPF module uses cubic splines to interpolate w between these points, and assumes $w(z > 1.5) \equiv w(z = 1.5)$.

We adopted the 307 SNeIa from the 'Union' data set compiled by Kowalski et al. (2008). The CMB data and likelihood used were the *Wilkinson Microwave Anisotropy Probe (WMAP)* 5-yr data set (Dunkley et al. 2009). Tegmark et al. (2006) provided the data and likelihood code for the mpk using Sloan Digital Sky Survey (SDSS) Data Release (DR)4. The evidence for the known models is computed using the publicly available MULTINEST code (Feroz & Hobson 2008; Trotta et al. 2008; Feroz et al. 2009), which implements the nested sampling algorithm, employed as an add-in sampler to COSMOMC (Lewis & Bridle 2002) and CAMB (Lewis et al. 2000).

The gist of nested sampling is that the multidimensional evidence integral of equation (2) is recast into a 1D integral. This is accomplished by defining the prior volume x as $\mathrm{d}x \equiv p(\theta)\mathrm{d}\theta$ so that

$$x(\lambda) = \int_{\mathcal{L}(\theta)>\lambda} p(\theta)\,\mathrm{d}\theta, \tag{19}$$

where the integral is over the parameter space enclosed by the iso-likelihood contour $\mathcal{L}(\theta) = \lambda$. So $x(\lambda)$ gives the volume of parameter space above a certain level λ of the likelihood. Then the Bayesian evidence, equation (2), can be written as

$$p(d) = \int_0^1 \mathcal{L}(x)\,\mathrm{d}x, \tag{20}$$

where $\mathcal{L}(x)$ is the inverse of equation (19). Samples from $\mathcal{L}(X)$ can be obtained by drawing uniformly samples from the likelihood volume within the isocontour surface defined by λ. The 1D integral of equation (20) can be obtained by simple quadrature, thus

$$p(d) \approx \sum_i \mathcal{L}(x_i) W_i, \tag{21}$$

where the weights are $W_i = (1/2)(x_{i-1} - x_{i+1})$. The standard deviation on the value of the log-evidence can be estimated as $(H/n_{\mathrm{live}})^{1/2}$, where H is the negative relative entropy and n_{live} is the number of live points adopted, which in our case is $n_{\mathrm{live}} = 4000$ (see Feroz & Hobson 2008 for details).

The best-fitting χ^2 required to evaluate equation (10) is obtained by performing a Metropolis–Hastings Markov chain Monte Carlo (MCMC) reconstruction of the posterior of the 16-parameter model comprising the ΛCDM parameters θ and the above 10-parameter description of $w(z)$. We gather a total of 5×10^5 samples in eight parallel chains and verify that the Gelman & Rubin mixing criterion (Gelman & Rubin 1992) is satisfied (i.e. $R \ll 0.1$, where R is the interchain variance divided by the intrachain variance).

MCMC is rather geared towards exploring the bulk of the posterior probability density, and is not particularly optimized to look for the absolute best-fitting value. This is especially true for high-dimensional parameter spaces. Therefore, we expect that the best-fitting χ^2 values recovered via MCMC for the 16D model X are going to be systematically higher than the true best fit. In order to estimate and correct for this numerical bias, we sampled via MCMC a 16D Gaussian of unit variance, recovered the best-fitting χ^2 and compared it with the true best-fitting value, repeating the procedure 5000 times. This gives an estimate of the numerical bias, under the assumption (which is valid locally) that the posterior distribution of model X is close to Gaussian in the immediate vicinity of the best fit. We found that the MCMC systematically overestimates the best-fitting χ^2 value by 0.94 ± 0.14, and therefore subtracted this

[1] http://camb.info/ppf/

Table 2. In the first three columns, we report the Bayes factors between the known models and ΛCDM for different combinations of data sets, where $\ln B_{j\Lambda} < 0$ favours ΛCDM. The fourth columns gives $\Delta\chi^2 = \chi^2_X - \chi^2_\Lambda$, the improvement in the best-fitting log-likelihood obtained by using model *X* (specified in the text) over ΛCDM. The last column gives the corresponding absolute upper bound to the Bayes factor between model *X* and ΛCDM. The table shows only statistical errors. There is a further systematic error from the assumed number of bins for $w(z)$ in the unknown model. This is estimated to be ±0.5 for $\Delta\chi^2$, which translates into a 25 per cent systematic uncertainty in $\overline{B}_{X\Lambda}$ (on top of the statistical uncertainty given in the table).

	$-1.3 < w < -0.3$	$w = -1.0$	$-1.3 < w < -0.3$	'Unknown' model *X*		$\langle B_{i\Lambda}\rangle$	$\overline{B}_{X\Lambda}$
	$\Omega_\kappa = 0.0$	$-1.0 < \Omega_\kappa < 1.0$	$-1.0 < \Omega_\kappa < 1.0$				
	$\ln B_{j\Lambda}$	$\ln B_{j\Lambda}$	$\ln B_{j\Lambda}$	$\Delta\chi^2$	$\ln \overline{B}_{X\Lambda}$		
CMB only	-0.30 ± 0.09	-2.25 ± 0.09	-2.79 ± 0.09	-1.02 ± 0.2	0.51 ± 0.2	0.48 ± 0.02	1.67 ± 0.2
CMB+SN	-0.85 ± 0.09	-2.52 ± 0.09	-3.33 ± 0.09	-3.05 ± 0.2	1.53 ± 0.2	0.39 ± 0.01	4.60 ± 0.7
CMB+mpk	-0.98 ± 0.08	-3.79 ± 0.08	-4.39 ± 0.08	-0.84 ± 0.2	0.42 ± 0.1	0.35 ± 0.01	1.52 ± 0.2
CMB+SN+mpk	-0.96 ± 0.09	-3.73 ± 0.09	-4.43 ± 0.09	-3.38 ± 0.3	1.69 ± 0.2	0.35 ± 0.01	5.42 ± 1.0

estimate from the recovered χ^2 best-fitting value for model *X*. The total error on the best-fitting χ^2 is computed by adding in quadrature the above error and the interchain variance of the recovered χ^2 for each of our eight chains. We also verified that the numerical bias in recovering the best-fitting χ^2 for a 6D parameter space (such as ΛCDM) is negligible in comparison.[2] There is also a source of systematic error coming form our choice of adopting a 10-bin model for w. We have verified that the best-fitting χ^2 changes by no more than 0.5 units if we halve or double the number of w bins for the *X* model, and we therefore add this systematic error to the statistical one.

4 RESULTS AND DISCUSSION

We now proceed to evaluate the doubt and the posterior probability of ΛCDM for various combinations of cosmological data sets.

4.1 Model comparison outcome including doubt

In Table 2, we present the estimated upper limit on the Bayes factor between ΛCDM and model *X* as well as the Bayes factors with respect to ΛCDM for the other known models. Among the known models, we confirm what many others have shown – that ΛCDM is the best-fitting known model, or at least that no other model is demonstrably better. Thus, we find an inconclusive model comparison result (according to the Jeffreys' scale, Table 1) when comparing ΛCDM and a model with a free (but constant) w. We also find moderate evidence ($\ln B \sim 2.5$) against spatially curved models when compared to a flat ΛCDM, in agreement with the more detailed findings of Vardanyan et al. (2009). Finally, there is moderate (tending to strong) evidence against the most complex of the known models – one exhibiting both $w \neq -1$ and $\Omega_\kappa \neq 0$. This is in good agreement with the results of previous more thorough analyses, e.g. Liddle, Mukherjee & Parkinson (2006a), Liddle et al. (2006b), Li, Li & Zhang (2010) and Kilbinger et al. (2010). From this, ordinary Bayesian model comparison concludes that ΛCDM is still the best of the known models (at least for the limited range of alternative models considered here).

Most importantly, in the table, we report the improvement in the best-fitting log-likelihood obtained over ΛCDM by using *X*, and use this to compute an absolute upper bound to the Bayes factor via equation (10). We notice that the improvement in the best fit is fairly modest for all the data sets considered, supportive of the general sentiment in the community that ΛCDM is in good agreement with available observations and that therefore there is little room for statistical improvement of the quality of fit. This is in part because it is very hard to improve the quality of fit by changing $w(z)$ – observables are usually a double integral of $w(z)$, and therefore insensitive to features in the equation of state (see e.g Huterer & Turner 1999; Maor, Brustein & Steinhardt 2001; Clarkson 2009). As a consequence, even a highly flexible $w(z)$ model such as the one we used here to describe *X* will lead to only small observable departures from the standard cosmological constant scenario. However, we do observe a stronger improvement in the best-fitting χ^2 using the *X* model for combinations of data sets including the SN data. This means that the fit of SN data can be slightly improved by our doubt model. It is important to keep in mind that such statements depend strongly on the statistics one employs to examine the models. For example, the standard likelihood function for CMB data is insensitive to most of the reported anomalies in the low-ℓ CMB (Bennett et al. 2010; Copi et al. 2010). Also, fluctuations in the best-fitting χ^2 at the level observed here in the SN data can easily be ascribed to systematic discrepancies between the data and the model being fitted to them, in the form of e.g. evolution of the Phillips correction parameters (Kessler et al. 2010).

The interesting consequence from the point of view of doubt is that this translates into strong upper limits for the Bayes factor between model *X* and ΛCDM (third from last column of Table 2). We find that the upper limit on the Bayes factor $\overline{B}_{X\Lambda}$ (last column of Table 2) for all the data combinations is less than ∼5, just above the 'weak evidence' threshold (see Table 1). From our discussion in Section 2.2, this means that the necessary condition for doubt to grow, $p(X)\overline{B}_{X\Lambda} \gg 1$, is not met for any reasonable doubt prior choice. We remind the reader at this point that our unknown model *X* has been designed in such a way as to exhibit the maximum possible evidence against ΛCDM. Therefore, if even such a model cannot achieve a significant level of evidence against ΛCDM, one can safely conclude that no other reasonable model will. Of course this conclusion depends both on the set of observations we have considered and on the particular likelihood function we have ascribed to that data. New statistical treatments can bring to light anomalies in the existing data, while new observations might contain new unexpected features.

Our results in terms of posterior probability for doubt and for the ΛCDM model are shown in Table 3, for two different assumptions regarding the level of prior doubt, $p(X) = 10^{-2}$ and $p(X) = 10^{-6}$. These two choices are representative of a range that we think might

[2] In principle, one might use a numerical maximization method to search for the extremum of the likelihood. However, in general likelihood surfaces might not be very smooth, making this option not very feasible, see also the comments on the corresponding mode of COSMOMC at http://cosmologist.info/cosmomc/readme.html.

Table 3. First two columns: posterior doubt for different data sets combinations and two prior doubt assumptions. Last column: posterior probability for the ΛCDM model when allowing for the possibility of a 1 per cent prior doubt on the completeness of our list of known models.

	Doubt $\mathcal{D}$ Prior doubt: $p(X) = 10^{-2}$	Doubt $\mathcal{D}$ Prior doubt: $p(X) = 10^{-6}$	Posterior for ΛCDM, $p(\Lambda\|d)$ [with $p(X) = 10^{-2}$ and $p(\Lambda) \approx 0.25$]
CMB only	$(3.41 \pm 0.3) \times 10^{-2}$	$(3.49 \pm 0.3) \times 10^{-6}$	0.51 ± 0.02
CMB+SN	$(10.8 \pm 1.0) \times 10^{-2}$	$(11.9 \pm 1.2) \times 10^{-6}$	0.58 ± 0.02
CMB+mpk	$(4.18 \pm 0.4) \times 10^{-2}$	$(4.32 \pm 0.4) \times 10^{-6}$	0.68 ± 0.01
CMB+SN+mpk	$(13.4 \pm 1.6) \times 10^{-2}$	$(15.3 \pm 2.1) \times 10^{-6}$	0.61 ± 0.02

bracket reasonable prior expectations: a prior doubt of 1 per cent is certainly not too large, while leaving a little space for updating our models beliefs in the light of data. A prior doubt of 10^{-6} reflects the fact that surely we have to allow for a one-in-a-million chance that our current list of known models might be incomplete, and that the true underlying dark energy model might still be undiscovered.

Table 3 contains the level of doubt, which is updated from the prior by using the results of Table 2 for the models' evidence. We find an increase in doubt by a factor of 3–15, depending on the data combination used. As noted above, data sets including the SNe show a more marked improvement in the level of doubt. This, however, is largely a consequence of the doubt model acquiring some of the probability mass of the known models other than ΛCDM, as discussed under Case 3 in Section 2.2. Indeed, the posterior probability of ΛCDM is observed to increase (last column of Table 3), from the initial prior value $p(\Lambda) \approx 0.25$ to almost 70 per cent for the most constraining data combination. This result is almost independent of the choice of prior doubt. The behaviour of the posterior probability for doubt and ΛCDM for a prior choice $p(X) = 10^{-2}$ is shown in Fig. 1, as a function of the data sets employed.

Our rationale for introducing a small prior doubt, $p(X) \ll 1$, is to leave the door open to doubt, rather than to consider the case where doubt starts off with a probability comparable to the known models. After all, our method already gives to the unknown model a strong advantage, by removing by construction any Occam's razor penalty from its evidence and giving it the maximum evidence the data can possibly allow a posteriori.

However, in order to assess the robustness of our conclusions, it is interesting to consider the extreme case when the prior doubt is set equal to the prior probability of the known models, i.e.

$$p(X) = p(\mathcal{M}_j) = \frac{1}{N+1}. \quad (22)$$

The resulting posterior probability for doubt and ΛCDM are given in Table 4. Despite the fact that this is an extremely favourable choice for doubt, we can see that the posterior doubt grows only by a factor of ∼2–4, while the probability of ΛCDM stays roughly constant, or increases somewhat. We conclude that even in this extreme case doubt cannot be genuinely favoured, for the condition $\mathcal{R}_\Lambda < 1$ is not satisfied.

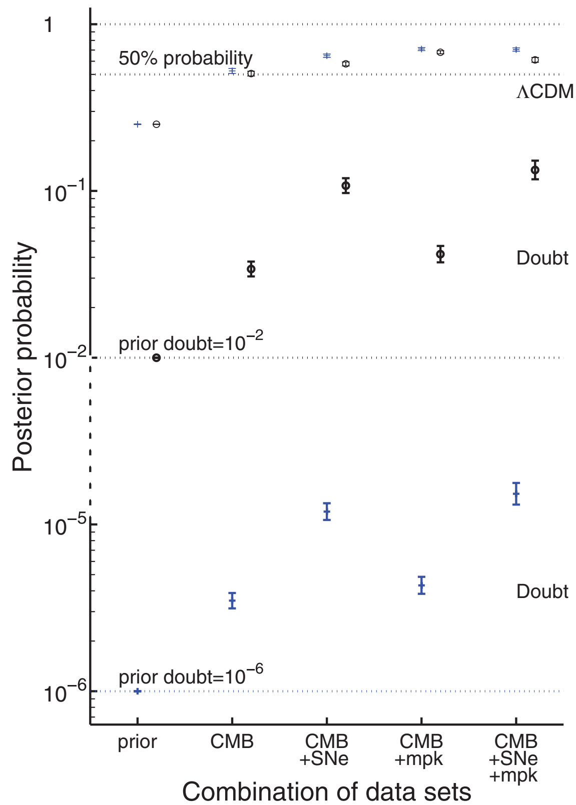

Figure 1. Posterior probability for doubt for the ΛCDM model as a function of different combinations of data sets, for two choices of (a small) prior doubt. The probability of ΛCDM increases from the initial 25 per cent to almost 70 per cent. The probability of doubt increases from the initial 1 per cent (10^{-6}) to just over 13 per cent (15×10^{-6}), mostly as a consequence of acquiring probability from the other three known models considered in the analysis. This increase in doubt is driven largely by the SN data. The fact that the probability for ΛCDM increases, as well, signals that ΛCDM remains the most valid statistical description of the data.

4.2 Impact of the addition of further known models

We now proceed to estimate the robustness of our findings with respect to expanding the set of known models. As has been mentioned above, the list of three alternative known models to ΛCDM we adopted in this work is far from complete. However, even if a larger number of models N were included in the known models list, it is reasonable to assume that the value of the average evidence between the known models and ΛCDM would scale approximately as $\propto 1/N$, for there is no other known model that presently can achieve a substantially higher evidence than ΛCDM (if this was the case, then this other best model would take the place of ΛCDM and become our baseline model which we seek to doubt – or rather the dominant model in our list of models where we intend to compute the doubt for the whole list). By equating equations (4) and (9) we can solve for the value of $\Delta\chi^2$ required for the posterior on doubt to be equal to the posterior of ΛCDM. This gives the approximate condition [assuming that $\langle B_{i\Lambda}\rangle \approx 1/N$ and that $p(X) \ll 1$]

$$\Delta\chi^2 \approx 2\ln(Np(X)). \quad (23)$$

So the value of $\Delta\chi^2$ required for posterior doubt to reach the posterior for ΛCDM scales logarithmically with the number of known models. Assuming a prior doubt $p(X) = 10^{-2}$ one obtains the

Table 4. Posterior doubt and posterior probability for ΛCDM for the case where doubt is given the same prior probability as the known models, i.e. $p(X) = 1/(N + 1)$. Even in this extremely favourable case for doubt, the posterior probability for ΛCDM does not decrease, showing that ΛCDM cannot be genuinely doubted.

	Doubt $\mathcal{D}$ Prior doubt: $p(X) = 0.2$	Posterior for ΛCDM, $p(\Lambda\|d)$ [with $p(X) = 0.2$ and $p(\Lambda) = 0.2$]
CMB only	0.46 ± 0.02	0.28 ± 0.02
CMB+SN	0.75 ± 0.02	0.16 ± 0.02
CMB+mpk	0.52 ± 0.03	0.34 ± 0.03
CMB+SN+mpk	0.79 ± 0.02	0.15 ± 0.02

Table 5. Improvement in the χ^2 of ΛCDM required for the unknown model X to have the same a posteriori probability as ΛCDM. First two columns: as a function of the number of known models, N, assuming a fixed prior doubt $p(X) = 10^{-2}$. Last two columns: assuming a fixed fractional prior doubt, $p(X) = p(\Lambda)/f$, and as a function of f. It is assumed that the evidence of the known models is much smaller than the evidence for ΛCDM.

Known models	Required $\Delta\chi^2$ for $p(X\|d) = p(\Lambda\|d)$		
N	$p(X) = 10^{-2}$	f	$p(X) = p(\Lambda)/f$
4	−6.4	4	−5.5
10	−4.6	10	−9.2
20	−3.2	10^2	−18.4
50	−1.4	10^3	−27.6

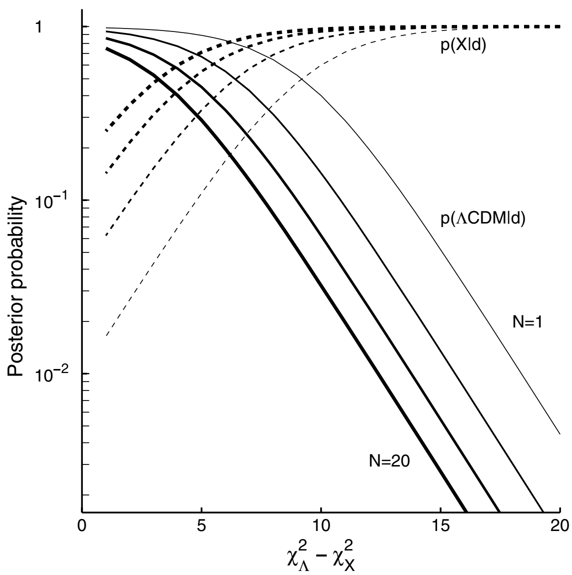

Figure 2. Posterior for doubt (dashed lines) and for ΛCDM (solid lines) as a function of $-\Delta\chi^2 = \chi^2_\Lambda - \chi^2_X$ assuming a fixed prior doubt $p(X) = 10^{-2}$. Different curves are for different numbers of known models, $N = 1, 4, 10$ and 20 (from thin to thick), assuming that $\langle B_{i\Lambda}\rangle \approx 1/N$.

values of $\Delta\chi^2$ listed in the first column of Table 5 as a function of N. As more known doubts are put on the table, it becomes easier to doubt ΛCDM. From this scaling, it would appear that the improvement of $\Delta\chi^2 = -3.4$ for model X reported in Table 2 for the data combination cmb+SN+mpk would lead to a larger probability of doubt than for ΛCDM if we had assumed a list of $N \gtrsim 20$ known models, rather than just three. As illustrated in Fig. 2, this effect is, however, a consequence of our choice of spreading the level of prior probabilities among the N known models, while assuming *a fixed* $p(X)$, see equation (3). As N increases, the prior for ΛCDM decreases while the prior doubt is kept constant. As a consequence, it becomes easier for the former to 'catch up' with the latter.

In order to avoid this spurious effect, one could choose to set the prior doubt as a fraction $1/f$ ($f > 1$) of the prior probability for ΛCDM, i.e. to require that the relative probability between X and Λ is constant a priori, independent of the number of known models. We thus replace the prescription of equation (3) by

$$p(\Lambda) = \frac{1}{N}(1 - p(X)) \tag{24}$$

$$p(X) = \frac{p(\Lambda)}{f} = (Nf + 1)^{-1} \tag{25}$$

and by equating the posterior doubt with the posterior for ΛCDM we obtain the following requirement for the $\Delta\chi^2$:

$$\Delta\chi^2 = -4 \ln f. \tag{26}$$

This is now independent of the number of known models N and it only depends logarithmically on the prior doubt fraction, f. From the last two columns of Table 5 we can see that even if doubt started off a factor of just $f = 4$ less probable than ΛCDM, a $\Delta\chi^2 = -5.5$ would be required in order for the unknown model X to become as probable as ΛCDM. Increasing the prior gap between doubt and ΛCDM (i.e. increasing f) only makes the requirements on the χ^2 improvement more taxing.

In summary, once the effect of adding extra models to the known models' list is corrected for by introducing the fractional prior doubt f, we find that the improvement in the χ^2 found for various combinations of data sets is insufficient to doubt ΛCDM. If the unknown model starts off being a factor of 4 less probable than ΛCDM, one would need an improvement in the χ^2 of about 5 units to reverse the situation in the posterior, which is quite a bit larger than the maximum χ^2 improvement observed from the data.

5 CONCLUSIONS

The aim of this paper was to extend the application of Bayesian model selection to define an absolute scale of goodness of fit for models, rather than just a relative one, such as the Jeffreys' scale. We showed how the notion of doubt can be used to evaluate the evidence in favour of a missing 'ideal' unknown model in the list of known cosmological models. We demonstrated how a useful absolute upper bound to the Bayesian evidence of an unknown model can be derived and how this can be implemented in the context of Bayesian model comparison.

Doubt can be incorporated in the framework of model comparison to help us decide whether our currently 'best' model is statistically adequate for the data at hand. Kunz, Trotta & Parkinson (2006) introduced the notion of Bayesian complexity to decide whether the available models are overcomplex with respect to the constraining power of the data. Bayesian doubt can act as a useful complement to Bayesian complexity, giving an indication of whether the current models are statistically insufficient to describe the data. Used in conjunction, doubt and complexity can thus extend the power and domain of applicability of Bayesian model comparison. Of course statistical considerations should never replace proper physical insight: all of our arguments are restricted to the statistical aspects of data modelling. But for the problem of dark energy, where most 'models' are of a phenomenological kind, it seems to us that a rigorous statistical framework can help deciding whether new theoretical explorations might be fruitful. Other domains where we expect doubt to be useful include the description of the spectral distribution of CMB anisotropies and the problem of anomalous alignments between multipoles in the CMB (Tegmark, de Oliveira-Costa & Hamilton 2003; Schwarz et al. 2004; Land & Magueijo 2005).

We have applied this methodology to the problem of dark energy, adopting a list of known models including possible extensions of the dark energy sector and non-zero curvature of the Universe. In principle, many more models could be added to the list of known models. However we argued that our results are robust against adding further models to the list of known models. We found that current CMB, mpk and SNIa data do not require the introduction of an alternative model to the baseline flat ΛCDM model. The upper bound of the Bayesian evidence for a presently unknown dark energy model against ΛCDM gives only weak evidence in favour of the unknown model. Since this is an absolute upper bound, we conclude that ΛCDM remains a sufficient phenomenological description of currently available observations.

ACKNOWLEDGMENTS

MCM was partially supported by a travel grant by the Royal Astronomical Society. MCM would like to thank CWRU for hospitality. PMV would like to thank Imperial College London for hospitality. PMV was supported by NASA ATP grant NNX07AG89G to Case Western Reserve University. GDS was supported by a grant from the US DoE to the theory group at CWRU. We would like to thank Pietro Berkes, Andrew Jaffe and Ben Wandelt for useful discussions. We thank an anonymous referee for useful suggestions. Numerical calculations were carried out at Case Western's High Performance Computer Cluster and at the Imperial College High Performance Computing Service. We acknowledge the use of the Legacy Archive for Microwave Background Data Analysis (LAMBDA). Support for LAMBDA is provided by the NASA Office of Space Science.

REFERENCES

Bennett C. L. et al., 2010, preprint (arXiv:1001.4758)
Clarkson C., 2009, preprint (arXiv:0911.2601)
Copi C. J., Huterer D., Schwarz D. J., Starkman G., 2010, preprint (arXiv:1004.5602)
Dunkley J. et al., 2009, ApJS, 180, 306
Fang W., Hu W., Lewis A., 2008a, Phys. Rev. D, 78, 87303
Fang W., Wang S., Hu W., Haiman Z., Hui L., May M., 2008b, Phys. Rev. D, 78, 103509
Feroz F., Hobson M. P., 2008, MNRAS, 384, 449
Feroz F., Hobson M. P., Bridges M., 2009, MNRAS, 398, 1601
Gelman A., Rubin D., 1992, Statistical Sci., 7, 457
Gordon C., Trotta R., 2007, MNRAS, 382, 1859
Heavens A. F., Kitching T. D., Verde L., 2007, MNRAS, 380, 1029
Hu W., 2008, Phys. Rev. D, 77, 103524
Hu W., Sawicki I., 2007, Phys. Rev. D, 76, 104043
Huterer D., Turner M. S., 1999, Phys. Rev. D, 60, 081301
Jaynes E. T., 2003, Probability Theory: The Logic of Science. Cambridge Univ. Press, Cambridge
Kessler R. et al., 2009, ApJS, 185, 32
Kilbinger M. et al., 2010, MNRAS, 405, 2381
Kowalski M., Rubin D., Aldering G., Agostinho R. J., Amadon A., 2008, ApJ, 686, 749
Kunz M., Trotta R., Parkinson D., 2006, Phys. Rev. D, 74, 023503
Land K., Magueijo J. a., 2005, Phys. Rev. Lett., 95, 071301
Lewis A., Bridle S., 2002, Phys. Rev. D, 66, 103511
Lewis A., Challinor A., Lasenby A., 2000, ApJ, 538, 473
Li M., Li X., Zhang X., 2010, Science China Phys. Mech. Astron., 2010, 53, 1613
Liddle A. R., Mukherjee P., Parkinson D., 2006a, Astron. Geophys., 47, 4.30
Liddle A. R., Mukherjee P., Parkinson D., Wang Y., 2006b, Phys. Rev. D, 74, 123506
Maor I., Brustein R., Steinhardt P. J., 2001, Phys. Rev. Lett., 86, 6
Popper K., 2002, The Logic of Scientific Discovery, 2nd edn. Routledge, London
Schwarz D. J., Starkman G. D., Huterer D., Copi C. J., 2004, Phys. Rev. Lett., 93, 221301
Serra P., Heavens A., Melchiorri A., MNRAS, 379, 169
Starkman G. D., Trotta R., Vaudrevange P. M., 2008, preprint (arXiv:0811.2415)
Starkman G. D., Trotta R., Vaudrevange P. M., 2010, MNRAS, 401, L15
Tegmark M., de Oliveira-Costa A., Hamilton A. J. S., 2003, Phys. Rev. D, 68, 123523
Tegmark M. et al., 2006, Phys. Rev. D, 74, 123507
Trotta R., 2007, MNRAS, 378, 72
Trotta R., 2008, Contemp. Phys., 49, 71
Trotta R., Feroz F., Hobson M., Roszkowski L., Ruiz de Austri R., 2008, J. High Energy Phys., 12, 24
Vardanyan M., Trotta R., Silk J., 2009, MNRAS, 397, 431

This paper has been typeset from a TEX/LATEX file prepared by the author.