Comparison of cosmological models using Bayesian theory

Moncy V. John
Department of Physics, St. Thomas College, Kozhencherri 689641, India

J. V. Narlikar
Inter-University Centre for Astronomy and Astrophysics, Post Bag 4, Ganeshkhind, Pune 411007, India

(Received 2 January 2001; published 15 January 2002)

Using the Bayesian theory of model comparison, a new cosmological model due to John and Joseph [M. V. John and K. Babu Joseph, Phys. Rev. D **61**, 087304 (2000)] is compared with the standard $\Omega_\Lambda \neq 0$ cosmological model. Their analysis based on the recent apparent magnitude-redshift data of type Ia supernovas found evidence against the new model; our more careful analysis finds instead that this evidence is not strong. On the other hand, we find that the angular size-redshift data from compact (milliarcsecond) radio sources do not discriminate between the two models. Our analysis serves as an example of how to compare the relative merits of cosmological models in general, using the Bayesian approach.

DOI: 10.1103/PhysRevD.65.043506 PACS number(s): 98.80.Es, 02.50.−r

I. INTRODUCTION

In a recent publication [1], it was argued, by modifying an earlier ansatz by Chen and Wu [2], that the total energy density $\tilde{\rho}$ for the universe should vary as a^{-2} where a is the scale factor of its expansion. If the total pressure is $\tilde{p}$, then this argument leads to $\tilde{\rho}+3\tilde{p}=0$ for the universe. This deduction was made possible by the use of some dimensional considerations in line with quantum cosmology. The reasoning is as follows: Taking the comoving coordinate grid as dimensionless, we attribute a distance dimension to the scale factor a. Since there is no other fundamental energy scale available, one can always write $\tilde{\rho}$ as Planck density ($\rho_{pl} = c^5/\hbar G^2 = 5.158\times 10^{93}$ g cm^{-3}) times a dimensionless product of quantities. The variation of $\tilde{\rho}$ with a can now be written as

$$\tilde{\rho} \propto \rho_{pl}\left[\frac{l_{pl}}{a}\right]^n,$$

where $l_{pl}=(\hbar G/c^3)^{1/2}=1.616\times 10^{-33}$ cm is the Planck length. It is easy to see that $n<2$ $(n>2)$ will lead to a negative (positive) power of $\hbar$ appearing explicitly on the right hand side of the above equation. It was pointed out that such an $\hbar$-dependent total energy density would be quite unnatural in the classical Einstein equation for cosmology, much later than the Planck time. However, the case $n=2$ is just right to survive the semiclassical limit $\hbar \to 0$. Thus it was argued that if we take quantum cosmology seriously, then $\tilde{\rho}\propto a^{-2}$ or equivalently $\tilde{\rho}+3\tilde{p}=0$, for a conserved $\tilde{\rho}$. Solving the Friedmann equations gives a coasting evolution for the universe: i.e.,

$$a = m\,t,$$

where $m=\sqrt{k/(\tilde{\Omega}-1)}$ is a proportionality constant; $\tilde{\Omega}$ is the total density parameter and $k=0,\pm 1$ is the spatial curvature constant.

It shall be noted that $\tilde{\rho}+3\tilde{p}=0$ is an equation of state appropriate for strings or textures and that it is unrealistic to consider the present universe as string-dominated. But in [1], it was shown that this ansatz will lead to a realistic cosmology if we consider that $\tilde{\rho}$ is comprised of more than one component, say, ordinary matter (relativistic or nonrelativistic) with equation of state $p_m = w\,\rho_m$ and a cosmological constant Λ, which is time-varying. Let ρ_Λ denote the energy density arising from Λ and $p_\Lambda = -\rho_\Lambda$ be the corresponding pressure. With

$$\tilde{\rho}=\rho_m+\rho_\Lambda, \quad \tilde{p}=p_m+p_\Lambda,$$

the condition $\tilde{\rho}+3\tilde{p}=0$ will give

$$\frac{\rho_m}{\rho_\Lambda}=\frac{2}{1+3w},$$

and this gives a realistic model for the universe. It was also shown that this simplest cosmological model is devoid of the problems such as the horizon, flatness, monopole, cosmological constant, size, age of the universe and the generation of density perturbations on scales well above the present Hubble radius in the pure classical epoch. The solution of the cosmological constant, age and density perturbation problems deserve special mention since these are not solvable in an inflationary scenario. Moreover, the evolution of temperature in the model is nearly the same as that in the standard big bang model and if we assume the values $\Omega_m = 4/3$ and $\Omega_\Lambda = 2/3$, then there is no variation in the freezing temperature with the latter model, and this will enable nucleosynthesis to proceed in an almost identical manner. It also may be noted that an almost similar model which predicts the above values for the density parameters was proposed earlier [3], from some more fundamental assumptions based on entirely different grounds.

However, it should be remarked that the argument given above, which leads to this cosmology, is heuristic and not based on formal reasoning. It should be taken only as a guiding principle. Also we note that it has some unusual consequences like the necessity of continuous creation of matter from vacuum energy, though it was argued in [1,2] that such creation will be too inaccessible to observation.

But it was mentioned in [1] that, in spite of those successes in predicting observed values, the recent observations of the magnitudes of 42 high-redshift type Ia supernovas [4] are a setback for the model. A statement was explicitly made to the effect that the predictions of Ω_m and Ω_Λ for the present model are outside the error ellipses given in the $\Omega_m - \Omega_\Lambda$ plot in [4] and it was claimed that this discrepancy is a serious problem. In this paper, we study this issue in detail to see how strong is the evidence against this model when compared with the standard model with a constant $\Lambda \neq 0$, discussed in [4,5]. Jackson and Dodgson [6,7] have examined the latter model in the light of Kellerman's [8] and Gurvits' [9] compilations of angular size-redshift data for ultracompact (milliarcsecond) radio sources. Gurvits' compilation of such data, which are measured by very long-baseline interferometry (VLBI), is claimed to have no evolution with cosmic epoch. Several authors (for, e.g., [10]) have made use of these data to test their cosmological models. In the present paper, we also analyze Gurvits' data to test the new model. Using the Bayesian theory of statistics, we compare the new model discussed above with the standard model with a non-zero cosmological constant, using both the apparent magnitude-redshift data and the angular size-redshift data. It is found that there is no strong evidence against the new model when the apparent magnitude-redshift data are considered. This is contradictory to the statement made in [1]. The angular size-redshift data, on the other hand, are found to provide equal preference to the standard model and the new one.

The remainder of this paper takes the new theory as given and compares it with other standard cosmological models. The analysis shall be viewed as an example of using Bayesian theory to test the relative merits of cosmological models, a method which is claimed to have many positive features when compared to indirect arguments using parameter estimates. As such, the technique described here has wider applicability than just to the comparison of two cosmological models.

The paper is organized as follows. In Sec. II, we discuss the Bayesian theory of model comparison for the general case. Section III discusses comparison of the two models using apparent magnitude-redshift data, and in Sec. IV we compare the models with the angular size-redshift data. Section V comprises a discussion of the results.

II. BAYESIAN THEORY OF MODEL COMPARISON

The Bayesian theory of statistics [11,12] is historically the original approach to statistics, developed by great mathematicians such as Gauss, Bayes, Laplace, Bernoulli, etc., and has several advantages over the currently used long-run relative frequency (frequentist) approach to statistics, especially in problems like those in astrophysics, where the notion of a statistical ensemble is highly contrived. The frequentist definition of probability can only describe the probability of a true random variable, which can take on various values throughout an ensemble or a series of repeated experiments. In astrophysical and similar problems, ensembles and repeated experiments are rarely possible and we speak about the probability of a hypothesis, which can only be either true or false, and hence is not a random variable. The Bayesian theory will help assign probabilities for such hypotheses by considering the (often incomplete) data available to us. For example, Laplace used Bayesian theory to estimate the masses of planets from astronomical data, and to quantify the uncertainty of the masses due to observational errors [13]. In fact, this theory finds application in all those problems where one can only have a numerical encoding of one's state of knowledge.

In the Bayesian theory of model comparison, it is common to report model probabilities via odds, the ratios of probabilities of the models. The posterior (i.e., after consideration of the data) odds for the model M_i over M_j are

$$O_{ij} = \frac{p(M_i|D,I)}{p(M_j|D,I)},$$

where $p(M_i|D,I)$ refers to the posterior probability for the model M_i, given the data D and assuming that any other information I regarding the models under consideration is true. Using Bayes's theorem, one can write the above equation as

$$O_{ij} = \frac{p(M_i|I)\mathcal{L}(M_i)}{p(M_j|I)\mathcal{L}(M_j)}, \tag{1}$$

where $p(M_i|I)$ is called the prior probability; i.e., any probability assigned to the model M_i before consideration of the data, but assuming the information I to be true. When I does not give any preference to one model over the other, these prior probabilities are equal so that

$$O_{ij} = \frac{\mathcal{L}(M_i)}{\mathcal{L}(M_j)} \equiv B_{ij}. \tag{2}$$

B_{ij} is called the Bayes factor. $\mathcal{L}(M_i)$ denotes the probability $p(D|M_i)$ to obtain the data D if the model M_i is the true one and is referred to as the likelihood for the model M_i. The models under consideration will usually have one or more free parameters (like the density parameters Ω_m, Ω_Λ, etc. in the case of cosmological models), which we denote as α, β, ... $\mathcal{L}(M_i)$ can be evaluated for models with one parameter as

$$\mathcal{L}(M_i) \equiv p(D|M_i) = \int d\alpha\, p(\alpha|M_i)\mathcal{L}_i(\alpha), \tag{3}$$

where $p(\alpha|M_i)$ is the prior probability for the parameter α, assuming the model M_i to be true. $\mathcal{L}_i(\alpha)$ is the likelihood for α in the model and is usually taken to have the form

$$\mathcal{L}_i(\alpha) \equiv \exp[-\chi^2(\alpha)/2], \tag{4}$$

where

$$\chi^2 = \sum_k \left(\frac{\hat{A}_k - A_k(\alpha)}{\sigma_k}\right)^2 \tag{5}$$

is the χ^2 statistic. Here $\hat{A}_k$ are the measured values of the observable A, $A_k(\alpha)$ are its expected values (from theory) and σ_k are the uncertainties in the measurements of the observable.

Generalization to the case of more than one parameter is straightforward. As a specific case, consider a model M_i with two parameters, α and β, having flat prior probabilities; i.e., we assume to have no prior information regarding α and β except that they lie in some range $[\alpha, \ \alpha+\Delta\alpha]$ and $[\beta, \ \beta+\Delta\beta]$, respectively. Then $p(\alpha|M_i)=1/\Delta\alpha$, $p(\beta|M_i)=1/\Delta\beta$ and hence

$$\mathcal{L}(M_i)=\frac{1}{\Delta\alpha}\frac{1}{\Delta\beta}\int_{\alpha}^{\alpha+\Delta\alpha}d\alpha\int_{\beta}^{\beta+\Delta\beta}d\beta\exp[-\chi^2(\alpha,\beta)/2]. \tag{6}$$

It is instructive to rewrite this equation as

$$\mathcal{L}(M_i)=\frac{1}{\Delta\alpha}\int_{\alpha}^{\alpha+\Delta\alpha}d\alpha\mathcal{L}_i(\alpha).$$

In this case,

$$\mathcal{L}_i(\alpha)=\frac{1}{\Delta\beta}\int_{\beta}^{\beta+\Delta\beta}d\beta\exp[-\chi^2(\alpha,\beta)/2]$$

is called the marginal likelihood for the parameter α.

A. Interpretation of the Bayes factor

The interpretation of the Bayes factor B_{ij}, which is given by Eq. (2) and which evaluates the relative merits of model M_i over model M_j, is as follows [14]: If $1<B_{ij}<3$, there is evidence against M_j when compared with M_i, but it is not worth more than a bare mention. If $3<B_{ij}<20$, the evidence against M_j is definite but not strong. For $20<B_{ij}<150$, this evidence is strong and for $B_{ij}>150$, it is very strong.

III. COMPARISON USING REDSHIFT-MAGNITUDE DATA

For a Friedmann-Robertson-Walker (FRW) model which contains matter and a cosmological constant, the likelihood for these parameters, i.e., $\mathcal{L}_i(\Omega_m,\Omega_\Lambda)$ can be assigned using the redshift-apparent magnitude data in the following manner [14]. Before consideration of the data, let us agree that Ω_m lies somewhere in the range $0<\Omega_m<3$, Ω_Λ in the range $-3<\Omega_\Lambda<3$ and accept this as the only prior information I. Let $\hat{\mu}_k$ be the observed best-fit distance modulus for the supernova number k, s_k its uncertainty and $\hat{z}_k$ is the cosmological redshift, with w_k its uncertainty. We can write the expression for χ^2 as

$$\chi^2=\sum_k\left(\frac{\hat{\mu}_k-\mu_k}{\sigma_k}\right)^2. \tag{7}$$

Here,

$$\hat{\mu}_k=\mu_k+n_k=g_k-\eta+n_k, \tag{8}$$

with

$$\mu_k\equiv g_k-\eta=5\log\left[\frac{D_L(z;\Omega_m,\Omega_\Lambda,H_0)}{1\ \ \mathrm{Mpc}}\right]+25$$

being the redshift-apparent magnitude relation. The luminosity distance is $D_L(z;\Omega_m,\Omega_\Lambda,H_0)=cH_0^{-1}d_L(z;\Omega_m,\Omega_\Lambda)$, where c is the velocity of light, H_0 is the Hubble constant at the present epoch and d_L is the dimensionless luminosity distance. $g_k=g(\hat{z}_k)$ is the part of μ_k which depends implicitly on Ω_m and Ω_Λ and η is its H_0-dependent part. The latter quantity can be written as $\eta\equiv5\log(h/c_2)-25$ where $H_0=h\times100\ \ \mathrm{km\,s^{-1}\,Mpc^{-1}}$ and c_2 is the speed of light in units of $100\ \ \mathrm{km\,s^{-1}}$. The probability distribution for the value n_k in Eq. (8) is assumed to be a zero-mean Gaussian with standard deviation σ_k, where $\sigma_k^2=s_k^2+[\mu'(\hat{z}_k)]^2w_k^2$, in the absence of systematic or evolutionary effects.

One can evaluate $\mathcal{L}(\Omega_m,\Omega_\Lambda,\eta)$ in a manner similar to that in Eq. (4), where χ^2 now is a function of the three parameters Ω_m, Ω_Λ and η. The likelihood for Ω_m and Ω_Λ, denoted as $\mathcal{L}(\Omega_m,\Omega_\Lambda)$ can be obtained by the technique of marginalizing over η, if one assumes a flat prior probability for η in some appropriate range.

To do this, we define $s^{-1}=\sqrt{\Sigma_k(1/\sigma_k^2)}$ where s is the posterior uncertainty for η and let $1/\Delta\eta$ a flat prior probability be assigned to η. (These, being the same for all models, will get canceled when evaluating probability ratios.) Using these definitions, the marginal likelihood (defined at the end of Sec. II) for the density parameters is

$$\mathcal{L}(\Omega_m,\Omega_\Lambda)=\frac{1}{\Delta\eta}\int d\eta e^{-\chi^2/2}. \tag{9}$$

Evaluating this integral analytically [14], one assigns a likelihood for the parameters Ω_m and Ω_Λ in any one model as

$$\mathcal{L}(\Omega_m,\Omega_\Lambda)=\frac{s\sqrt{2\pi}}{\Delta\eta}e^{-q/2}, \tag{10}$$

where

$$q(\Omega_m,\Omega_\Lambda)=\sum_k\frac{(\hat{\mu}_k-g_k+\hat{\eta})^2}{\sigma_k^2}, \tag{11}$$

is of the form of a χ^2-statistic, with $\hat{\eta}$ the best fit (most probable) value of η, given Ω_m and Ω_Λ. The latter can be computed as [14]

$$\hat{\eta}(\Omega_m,\Omega_\Lambda)=s^2\sum_k\frac{(g_k-\hat{\mu}_k)^2}{\sigma_k^2}. \tag{12}$$

Now, we compare the model in [4,5] (model M_1, having parameters Ω_m, Ω_Λ and η) with the new model discussed in Sec. I (model M_2, having only the parameters Ω_m and η). The Bayes factor B_{12} can be written with the help of Eq. (2) and Eq. (3) as

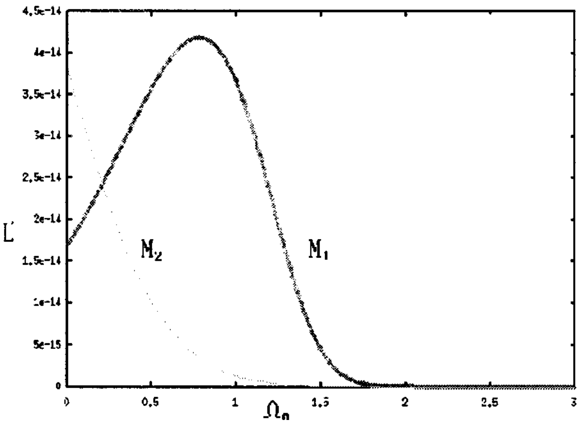

FIG. 1. L' vs Ω_m for both models, using the apparent magnitude-redshift data for type Ia supernova. The curves M_1 and M_2 correspond to the marginal likelihoods for Ω_m for the standard $\Omega_\Lambda \neq 0$ model and the new model, respectively (apart from some multiplicative constants, which cancel on taking ratios).

$$B_{12}=\frac{\mathcal{L}(M_1)}{\mathcal{L}(M_2)}$$

$$=\frac{\int d\Omega_m \int d\Omega_\Lambda p(\Omega_m,\Omega_\Lambda|M_1)\mathcal{L}_1(\Omega_m,\Omega_\Lambda)}{\int d\Omega_m p(\Omega_m|M_2)\mathcal{L}_2(\Omega_m)}. \tag{13}$$

With the information I at hand, one can assign flat prior probabilities $p(\Omega_m,\Omega_\Lambda|M_1)=1/18$ and $p(\Omega_m|M_2)=1/3$. Using Eqs. (6) and (10) we can write the above as

$$B_{12}=\frac{\int_{-3}^{3} d\Omega_\Lambda \int_0^3 d\Omega_m \exp[-q_1(\Omega_m,\Omega_\Lambda)/2]}{6\int_0^3 d\Omega_m \exp[-q_2(\Omega_m)/2]}. \tag{14}$$

Our first step in the evaluation of B_{12} is to find q given in Eq. (11), for both the models. For model 1, we have to use

$$g(z)=5\log\{(1+z)|\Omega_k|^{-1/2}\mathrm{sinn}[|\Omega_k|^{1/2}I(z)]\},$$

where $\Omega_k=1-\Omega_m-\Omega_\Lambda$ and $\mathrm{sinn}(x)=\sin x$ for $\Omega_m+\Omega_\Lambda>1$, $\mathrm{sinn}(x)=\sinh x$ for $\Omega_m+\Omega_\Lambda<1$ and $\mathrm{sinn}(x)=x$ for $\Omega_m+\Omega_\Lambda=1$. Also

$$I(z)=\int_0^z[(1+z')^2(1+\Omega_m z')-z'(2+z')(\Omega_\Lambda)]^{-1/2}dz'.$$

For model 2, the function $g(z)$ can be written as

$$g(z)=5\log\left\{m(1+z)\mathrm{sinn}\left(\frac{1}{m}\ln(1+z)\right)\right\},$$

where $m=\sqrt{2k/(3\Omega_m-2)}$ for the nonrelativistic era and $\mathrm{sinn}(x)=\sin x$ for $\Omega_m>2/3$, $\mathrm{sinn}(x)=\sinh x$ for $\Omega_m<2/3$ and $\mathrm{sinn}(x)=x$ for $\Omega_m=2/3$.

Using these expressions, Eq. (14) is numerically evaluated to obtain $B_{12}=3.1$. (In this calculation, we have used the data corresponding to the Fit C in [4], which involve 54 supernovas.) As per the interpretation of B_{ij} given in Sec. II A, the above is evidence against model 2, but it is only barely definite; the discrepancy is not a ‘‘serious problem’’ as had been stated in [1].

IV. COMPARISON USING ANGULAR SIZE-REDSHIFT DATA

For this purpose, we use Gurvits’ data and divide the sample which contains 256 sources into 16 redshift bins, as done by Jackson and Dodgson and shown in their Fig. 1 [7]. For model 1, we use the expression for angular size

$$\Delta\theta=\frac{d}{d_A}\equiv\frac{d}{(1+z)^{-1}(k/\Omega_k)\frac{c}{H_0}\mathrm{sinn}[|\Omega_k|^{1/2}I'(z)]}$$

$$=\frac{dH_0}{c}\frac{(1+z)}{(k/\Omega_k)^{1/2}\mathrm{sinn}[|\Omega_k|^{1/2}I'(z)]}, \tag{15}$$

where

$$I'(z)=\int_1^{1+z}\frac{dx}{x\left(\Omega_k+\Omega_m x+\frac{\Omega_\Lambda}{x^2}\right)^{1/2}}. \tag{16}$$

Here d is the linear dimension of an object, d_A is the angular size distance, and Ω_k and $\mathrm{sinn}(x)$ are defined as in the case of model 1 in the last section. Similarly for model 2, we have

$$\Delta\theta=\frac{d}{d_A}=\frac{dH_0}{c}\frac{(1+z)}{m\,\mathrm{sinn}\left(\frac{1}{m}\ln(1+z)\right)}, \tag{17}$$

where m and $\mathrm{sinn}(x)$ are defined as in the earlier case of model 2. For the purpose of comparison, we only need to combine the unknown parameters d and H_0 to form a single parameter $p\equiv dH_0/c$. Thus model 1 has three parameters p, Ω_m and Ω_Λ whereas model 2 has only the parameters p and Ω_m. As in the previous case, we accept $0<\Omega_m<3$ and $-3<\Omega_\Lambda<3$ as the prior information I. With these ranges of values of Ω_m and Ω_Λ, p is found to give significantly low values of χ^2 only for the range $0.1<p<1$ in both the models, p being given in units of milliarcseconds. The formal expressions to be used are

$$\chi^2=\sum_k\left(\frac{\Delta\hat{\theta}_k-\Delta\theta_k}{\sigma_k}\right)^2 \tag{18}$$

and

$$B_{12}=\frac{\mathcal{L}(M_1)}{\mathcal{L}(M_2)}=\frac{\frac{1}{\Delta p}\frac{1}{\Delta\Omega_m}\frac{1}{\Delta\Omega_\Lambda}\int dp\int d\Omega_m\int d\Omega_\Lambda \exp[-\chi_1^2(p,\Omega_m,\Omega_\Lambda)]}{\frac{1}{\Delta p}\frac{1}{\Delta\Omega_m}\int dp\int d\Omega_m \exp[-\chi_2^2(p,\Omega_m)]}$$

$$=\frac{\int_{0.1}^{1}dp\int_0^3 d\Omega_m\int_{-3}^{3}d\Omega_\lambda \exp[-\chi_1^2/2]}{6\int_{0.1}^{1}dp\int_0^3 d\Omega_m \exp[-\chi_2^2/2]}. \quad (19)$$

The result obtained is $B_{12}\approx 1$. This may be interpreted as providing equal preference to both models.

V. DISCUSSION

While evaluating the Bayes factors using both kinds of data, we have assumed that our prior information I regarding the density parameters is $0<\Omega_m<3$ and $-3<\Omega_\Lambda<3$. The range of values of Ω_Λ considered in [4] is $-1.5<\Omega_\Lambda<3$ and in [7] it is $-4<\Omega_\Lambda<1$. Even if we modify the range of this parameter in our analysis to some reasonable extent, the main conclusions of the paper will remain unaltered. For example, if we accept $0<\Omega_m<3$ and $-1.5<\Omega_\Lambda<1.5$ as some prior information I', the Bayes factors in each case become 3.8 and 0.8, in place of 3.1 and 1, respectively. Instead, if we choose I' as $0<\Omega_m<3$ and $-6<\Omega_\Lambda<6$, the corresponding values are 1.55 and 1.4, respectively. These do not change our conclusions very much in the light of the discriminatory inequalities mentioned in Sec. II A.

In order to get an intuitive feeling why the standard (M_1) and new (M_2) models have comparable likelihoods, consider Figs. 1 and 2. Figure 1 is for the apparent magnitude-redshift data and plots the quantities $L'=\frac{1}{6}\int_{-3}^{3}d\Omega_\Lambda \exp[-q_1(\Omega_m,\Omega_\Lambda)/2]$ (curve labeled M_1) and $L'=\exp[-q_2(\Omega_m)/2]$ (curve labeled M_2) against Ω_m. From the definition of marginal likelihood given at the end of Sec. II and from Eqs. (9)–(14), it can be seen that these two curves correspond to the marginal likelihoods for the parameter Ω_m in models M_1 and M_2, respectively (apart from some multiplicative constants, which cancel on taking ratios). Similarly, Fig. 2, which is for the angular size-redshift data, plots $\mathcal{L}=[1/(6\times 0.9)]\int_{0.1}^{1}dp\int_{-3}^{3}d\Omega_\Lambda \exp[-\chi_1^2/2]$ (curve M_1) and $\mathcal{L}=(1/0.9)\int_{0.1}^{1}dp \exp[-\chi_2^2/2]$ (curve M_2) against Ω_m. Equation (19) allows us to interpret these terms as the marginal likelihoods for Ω_m in models M_1 and M_2, respectively. In fact, these curves rigorously show the integrands one must integrate over Ω_m to get the Bayes factors. Using the apparent magnitude-redshift data, a lower value of q (which is a modified χ^2 statistic) is obtained for model M_1 whereas for angular size-redshift data, lower χ^2 is claimed by model M_2. However, the areas under the curves are comparable in both cases and this shows why the Bayes factors are also comparable. This is one of the strong points of the Bayesian method, in contrast to frequentist goodness of fit tests, which consider only the best fit parameter values for comparing models [11].

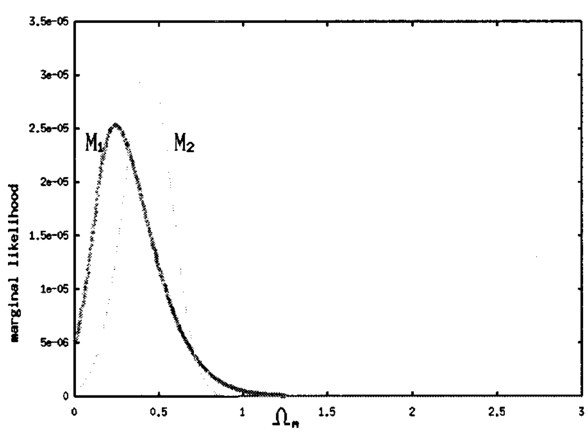

FIG. 2. Marginal likelihood vs Ω_m for both models, using the angular size-redshift data. The curves M_1 and M_2 correspond to the marginal likelihoods for Ω_m for the standard $\Omega_\Lambda\neq 0$ model and the new model, respectively.

These figures, however, show some feature that is disturbing for the new model. Figures 1 and 2 indicate best fit values of $\Omega_m=0$ and $\Omega_m=0.42$, respectively, for this model. In both cases it appears to rule out the value $\Omega_m=\frac{4}{3}$ that is needed to meet the constraints on nucleosynthesis, a condition which had been stated in the Introduction. Though, as mentioned above, Bayesian model comparison does not hinge upon the best fit values in evaluating relative merits of models, one would desire to have an agreement between predicted and observed parameter values. A natural option in such cases would be to compare the models by adjusting the prior probabilities regarding the parameters so that any additional information is accounted for. But we have not attempted this in our analysis.

The constant $\Omega_\Lambda\neq 0$ model we considered has one parameter in excess of the new model in both cases. It should be kept in mind that in the Bayesian method, simpler models with fewer parameters are often favored unless the data are truly difficult to account for with such models. Bayes's factors thus implement a kind of automatic and objective Occam's razor. In this context, it is interesting to check how the new model fares when compared with flat (inflationary) models where $\Omega_m+\Omega_\Lambda=1$, by which condition the number

TABLE I. Interpretation of results

Data	Model M_1	Model M_2	Bayes factor	Interpretation
$m-z$	Standard $\Omega_\Lambda \neq 0$ model	New model	$B_{12}=3.1$	Slightly definite but not strong evidence against the new model
$m-z$	Standard flat $\Omega_\Lambda \neq 0$ model	New model	$B_{12}=5$	Definite but not strong evidence against the new model
$\theta-z$	Standard $\Omega_\Lambda \neq 0$ model	New model	$B_{12}=1$	Both models are equally favored
$\theta-z$	Standard flat $\Omega_\Lambda \neq 0$ model	New model	$B_{21}=15$	Definite but not strong evidence against the flat model

of parameters of model M_1 is reduced by one. This makes the two models on a par with each other, with regard to the number of parameters. We have calculated the Bayes factor between this flat model M_1 and the new model M_2, using the apparent magnitude-redshift data and the result is $B_{12}=5.0$. This appears to be slightly more definite evidence against the new model than the corresponding result obtained in Sec. III ($B_{12}=3.1$). [However, inflationary models with a constant Λ-term suffer from the "graceful exit problem" for Λ; i.e., in order to explain how Λ manages to change from its grand unified theory (GUT) magnitude to $\approx 10^{-126}$ of its initial value, some extreme fine tuning would be required [15].] On the other hand, a comparison of the $\Omega_m+\Omega_\Lambda=1$ model with the new model using angular size-redshift data gives a value for the Bayes factor $B_{21}=15$, which shows that these data are more difficult to account for with the flat inflationary models than with the new one. The results we obtained, while using the information I, are summarized in Table I.

When compared to the frequentist goodness of fit test of models, which judge the relative merits of the models using the lowest value of χ^2 (even when it is obtained by some fine tuning or by having more parameters), the present approach has the advantage that it evaluates the overall performance of the models under consideration. The Bayesian method is thus a very powerful tool of model comparison and it is high time that the method is used to evaluate the plausibility of cosmological models cropping up in the literature. It is true that since we have only one universe, one can only resort to model making and then to comparing their predictions with observations. Again, since we cannot experiment with the universe, it is not meaningful to use the frequentist approach. We believe that the only meaningful way is to use the Bayesian approach in such cases. Here we have made a comparison between the model in [4,5] with the new model in [1]. It deserves to be stressed that the recent apparent magnitude-redshift observations on type Ia supernovas do not pose a "serious problem" to the new model, as had been claimed in [1]. The angular size-redshift data, on the other hand, do not discriminate between the general $\Omega_\Lambda \neq 0$ model and the new model and they provide definite but not strong evidence against the standard flat ($\Omega_m+\Omega_\Lambda=1$) model when compared to the new one.

Here it is essential to point out that Bayesian inference summarizes the weight of evidence by the full posterior odds and not just by the Bayes factor. Throughout our analysis above, we have assumed that the only prior information available to us is either I (stated in the beginning of Sec. III) or I' (stated in the beginning of Sec. V), which helps to make the posterior odds equal to the Bayes factor. However, when the Bayes factor is near unity, the prior odds $p(M_i|I)/p(M_j|I)$ in Eq. (1) become very important. The standard $\Omega_\Lambda \neq 0$ model and the standard flat (inflationary) models are plagued by the large number of cosmological problems (as mentioned in Sec. I) and the new model has the heuristic nature of its derivation and the problem with nucleosynthesis, setting (subjective) prior odds against each of them. In the context of having obtained comparable values for the Bayes factor, the Bayesian model comparison forces us to conclude, in a similar tone as in [14], that the existing apparent magnitude or angular size-redshift data alone are not very discriminating about these cosmological models. It is also worth remarking here that the Bayesian theory tells us how to adjust our plausibility assessments when our state of knowledge regarding a hypothesis changes through the acquisition of new data [11]. Concerning future observations, one would have to say that if the supernova test is extended to higher redshifts and if the astronomers are sure about the standard candle hypothesis, then the theories can be tested for such new data using Bayesian model comparison, using what we have now obtained as the prior odds. In this context, it also deserves serious consideration to extend the analysis done here to other cosmological data, such as those of cosmic microwave background radiation and primordial nucleosynthesis. Hopefully, further analysis and future observations will help to give more decisive answers to these questions.

ACKNOWLEDGMENTS

We are thankful to Professor K. Babu Joseph, Ninan Sajith Philip, and Dr. R. G. Vishwakarma for valuable discussions. Also the help in performing the computations rendered by Ninan Sajith Philip and Mathew V. Samuel is acknowledged with thanks.

[1] M. V. John and K. Babu Joseph, Phys. Rev. D **61**, 087304 (2000).
[2] W. Chen and Y. Wu, Phys. Rev. D **41**, 695 (1990).
[3] M. V. John and K. Babu Joseph, Phys. Lett. B **387**, 466 (1996); Class. Quantum Grav. **14**, 1115 (1997); M. V. John, Ph.D thesis, Cochin University of Science and Technology, Kochi, India, 2000, gr-qc/0007053.
[4] S. Perlmutter *et al.*, Astrophys. J. **517**, 565 (1999).
[5] A. G. Riess *et al.*, Astron. J. **116**, 1009 (1998).
[6] J. C. Jackson and M. Dodgson, Mon. Not. R. Astron. Soc. **278**, 603 (1996).
[7] J. C. Jackson and M. Dodgson, Mon. Not. R. Astron. Soc. **285**, 806 (1997).
[8] K. I. Kellerman, Nature (London) **361**, 134 (1993).
[9] L. I. Gurvits, Astrophys. J. **425**, 442 (1994).
[10] S. K. Banerjee and J. V. Narlikar, Mon. Not. R. Astron. Soc. **307**, 73 (1999).
[11] T. J. Loredo, in *Maximum Entropy and Bayesian Methods*, edited by P. F. Fougere (Kluwer, Dordrecht, 1990).
[12] G. D'Agostini, hep-ph/9512295.
[13] E. T. Jaynes, in *Maximum-Entropy and Bayesian Methods in Applied Statistics*, edited by J. H. Justice (Cambridge University Press, Cambridge, England, 1986).
[14] P. S. Drell, T. J. Loredo, and I. Wasserman, Astrophys. J. **530**, 593 (2000), and references therein.
[15] S. Weinberg, Rev. Mod. Phys. **61**, 1 (1989).