The Gaussian linear model provides a unique way to obtain the posterior probability distribution as well as the Bayesian evidence analytically. Considering the expansion rate data, the Gaussian linear model can be applied for ΛCDM, wCDM and a non-flat ΛCDM. In this paper, we simulate the expansion data with various precision and obtain the Bayesian evidence, then it has been used to discriminate the models. The data uncertainty is in range \( \sigma \in (0.5, 10)\% \) and two different sampling rate have been considered. Our results indicate that it is possible to discriminate \( w = -1.02 \) (or \( w = -0.98 \)) model from the ΛCDM \( (w = -1) \) with \( \sigma = 0.5\% \) uncertainty in expansion rate data. Finally, we perform a parameters inference in both the MCMC and Gaussian linear model, using current available expansion rate data and compare the results.

I. INTRODUCTION

ΛCDM cosmology offers a simple and consistent concordance model which has agreed with observations for several decades \([1, 2]\). Separately, the ΛCDM model gives excellent agreement with cosmic microwave background data \([3, 4]\), as well as late time measurements of cosmic expansion \([5]\). However, recent observations have suggested a growing cosmic tension in the reporting of the Hubble constant \([6–8]\), among other tensions. Together with the long standing cosmic tension in the reporting of the Hubble constant \([6–8]\), the ΛCDM dance model which has agreed with observations for several decades \([1, 2]\). Separately, the ΛCDM cosmology offers a simple and consistent concordance model which has agreed with observations for several decades \([1, 2]\). Among all tensions and issues in the ΛCDM, the so called H\(_0\) tension, is the most severe one. A Lot of efforts have been undertaken so far to tackle the problem without any reliable and satisfactory solution (to see more details refer to \([20]\)). Notice that, along with all model modifications scenarios that have been considered so far, describing a data set in a model independent approach might be very useful in this case \([21–23]\).

For a typical cosmological setup, one naturally investigates a Markov chain Monte Carlo (MCMC) approach to infer parameter values using the latest observational data sets. However, given the plethora of cosmological models being proposed, this approach only gives the constrain on the free parameters and says nothing about model comparison. For this purpose, some statistical measures, under some simplifying assumptions, including the Akaike information criterion (AIC), the Bayesian information criterion (BIC) and Deviance information criterion (DIC) have been used for model selection. On the other hand the Bayesian evidence provides a robust and reliable measure for model selection. Unfortunately, computation of the quantity involves a high-dimension integration over the product of likelihood and prior, which is computationally very expensive. To overcome this, some numerical approaches like the nested sampling \([24–26]\) and the Savage-Dickey density ratio \([27]\) have been developed. However, when the model is linear in its free parameters, the Gaussian linear model (GLM) provides an analytic solution for the Bayesian evidence. The formalism for a Gaussian or flat prior has been presented in \([28]\). Moreover, the Bayesian model selection has been utilized in \([29, 30]\) to understand reliability of the Bayesian evidence. In this work, we consider three important cosmological models which are linear in their parameters and apply the GLM method to understand how precision of the expansion rate data affects the significance of the model discrimination through Bayesian evidence.

The structure of this paper is as follows: In Sec. II, we give the basic formalism of the Gaussian linear model (GLM) and introduce the analytic formula to obtain the posterior distribution as well as the evidence. In Sec. III, the details of simulated data in our models are given. In addition, we present the results of applying the GLM on these data in the section. Then in Sec. IV, we describe current available Hubble data from different observations and perform an MCMC parameter inference to obtain the best value of parameters as well as their uncertainties. We also apply the GLM method to the observational data and compare the results with those of the MCMC. Finally, we conclude and discuss the main points of our finding in section Sec. V.

II. GAUSSIAN LINEAR MODEL

In this section, we briefly provide the basic formalism of the GLM. In this scenario, a database is modeled by a function which is linear in its parameters

\[
f(x, \theta) = \sum \theta_j x^j(x),
\]

where

\[
\theta_j \sim \mathcal{N}(0, \sigma^2)
\]
where $X(x)$ is an arbitrary function of $x$ and $\theta_i$s are the free parameters. Notice that the base functions can very well be a non-linear function of $x$. Assuming a $n_{\text{obs}}$ dimension database as $(x_i, y_i, \tau_i)$, the likelihood function is given by

$$ p(y|\theta) = \mathcal{L}_0 \exp\left[ -\frac{1}{2} (\theta - \theta_0)^T L (\theta - \theta_0) \right] $$

where,

$$ \mathcal{L}_0 = \frac{1}{(2\pi)^{n_{\text{obs}}/2} \Pi_{i} \tau_i} \exp\left[ -\frac{1}{2} (b - A\theta_0)^T L (b - A\theta_0) \right], $$

and

$$ F_{ij} = X'(x_i), \ A = F_{ij}, \ b = y_i, \ L = A'A, \ \theta_0 = L^{-1}A'b. $$

Here the maximum likelihood occurs at $\theta_0$ and $L$ denotes the likelihood Fisher matrix.

In order to perform the Bayesian parameter inference, we need to define a prior on the free parameters. We consider a Gaussian prior as

$$ p(\theta) = \frac{|\Sigma_{\text{pri}}|^{-1/2}}{(2\pi)^{n_{\text{par}}/2}} \exp\left[ -\frac{1}{2} (\theta - \theta_{\text{pri}})^T \Sigma_{\text{pri}}^{-1} (\theta - \theta_{\text{pri}}) \right], $$

where $\theta_{\text{pri}}$ ($\Sigma_{\text{pri}}$) is the mean (covariance matrix) of the prior and $n_{\text{par}}$ denotes the number of free parameters. Using the Bayes theorem, the posterior distribution is proportional to

$$ p(\theta|y) \propto \exp\left[ -\frac{1}{2} (\theta - \theta_{\text{pos}})^T \Sigma_{\text{pos}}^{-1} (\theta - \theta_{\text{pos}}) \right], $$

where

$$ \Sigma_{\text{pos}} = \Sigma_{\text{pri}}^{-1} + L $$

and

$$ \theta_{\text{pos}} = (\Sigma_{\text{pri}}^{-1} + L)^{-1} (\Sigma_{\text{pri}}^{-1} \theta_{\text{pri}} + L\theta_0) $$

In addition to the Bayesian parameter inference, the GLM provides the Bayesian evidence which is a key quantity in model comparison. The Bayesian evidence includes an integration over all parameters space and is given by

$$ p(y) = \int d\theta p(\theta)p(\theta|y). $$

Since both the prior and likelihood are a multivariate Gaussian in the GLM, the integral has an analytical solution. The Bayesian evidence in the GLM is given by,

$$ p(y) = \mathcal{L}_0 |\Sigma_{\text{pri}}|^{-1/2} |\Sigma_{\text{pos}}|^{1/2} \exp(D), $$

where

$$ D = \frac{1}{2} \left[ (\theta_{\text{pos}}^T \Sigma_{\text{pos}}^{-1} + \theta_0^T L (\Sigma_{\text{pri}}^{-1} + L)^{-1} (\Sigma_{\text{pri}}^{-1} \theta_{\text{pri}} + L\theta_0) \right. $$

$$ - (\theta_{\text{pos}}^T \Sigma_{\text{pos}}^{-1} \theta_{\text{pri}} + \theta_0^T L \theta_0), $$

### TABLE I. The Jeffreys’ scales for interpreting the Bayes factor.

| $\ln B_{01}$ | Strength of evidence |
|-------------|----------------------|
| $< 1$       | Inconclusive          |
| 1           | Weak evidence         |
| 1.5         | Moderate evidence     |
| $> 5$       | Strong evidence       |

### TABLE II. The Kass-Raftery scales for interpreting the Bayes factor.

| $\ln B_{01}$ | Strength of evidence |
|-------------|----------------------|
| 0 to 1      | Inconclusive          |
| 1 to 3      | Positive evidence     |
| 3 to 5      | Strong evidence       |
| $> 5$       | Very strong evidence  |

In order to apply the GLM in the cosmological context, we should have a linear model. Considering the $\Lambda$CDM model, the Hubble parameter as a function of redshift can be written as

$$ H^2(z) = 100^2 [\Omega_m h^2 (1 + z)^3 + h^2 - \Omega_m h^2] $$

$$ = 100^2 [\Omega_m h^2 ((1 + z)^3 - 1) + h^2], $$

where $H_0 = 100h$ is the current expansion rate of the universe and $\Omega_m$ is the matter density parameter. Interestingly, the second format is linear in its parameters and can be seen as a GLM with

$$ X^1(z) = (1 + z)^3 - 1, \ X^2(z) = 1 $$

and free parameters

$$ \theta_1 = \Omega_m h^2, \ \theta_2 = h^2 $$

### III. SIMULATED DATA AND RESULTS
In addition to the ΛCDM, the wCDM and non-flat ΛCDM (NACDM) also can be written as a GLM model,

\[ H^2(z) = 100^2\Omega_m h^2((1+z)^3 - (1+z)^3(1+w)) + h^2(1+z)^3(1+w), \]

for the wCDM and

\[ H^2(z) = 100^2\Omega_m h^2((1+z)^3 - 1) + \Omega_k h^2((1+z)^2 - 1) + h^2, \]

for the NACDM. In this case, the \( \Omega_k \) is the curvature density parameter. Since these models are GLM, it is easy to find the value of parameters which maximize the likelihood, the mean and covariance of posterior distribution as well as the Bayesian evidence.

To obtain the posterior distributions as well as the evidence, we need to define a prior on the free parameters. To avoid any possible prior bias, we consider a Gaussian wide prior on the free parameters. We use a multivariate Gaussian with mean and covariance matrix as

\[ \theta_{\text{pri}} = (\Omega_m h^2 = 0.13, \Omega_k h^2 = 0., h^2 = 0.45) \]

\[ \begin{bmatrix} 0.05 & 0 & 0 \\ 0 & 0.1 & 0 \\ 0 & 0 & 0.1 \end{bmatrix} \]

where the first, second and third row present covariance of the \( \Omega_m h^2, \Omega_k h^2 \) and \( h^2 \) respectively. We examine different mean and covariance matrix to check the robustness of our results. As long as the priors are wide enough, our results are the same and there is no prior bias.

### A. ΛCDM and wCDM

In order to realize how precision and sampling rate of an expansion rate database affects the model comparison, we simulate the expansion rate data with different precision and use the GLM method to perform a model comparison. To compare the ΛCDM and wCDM, two sampling rates have been considered. In the first case, we simulate 100 data points in range \( z \in (0, 3) \) with uncertainty \( \sigma \in (0.5, 10\%) \) using the ΛCDM model. For the second sampling strategy, we simulate 50 data points in the redshift range \( z \in (0, 2) \) and a similar uncertainty range as the first one. Now, we consider the simulated data and compute the evidence for both ΛCDM and wCDM. In our analysis, the EoS, is selected in a range \( (w \in [-0.9, -1.1]) \). Moreover, we consider a Gaussian distribution for the uncertainty and so the data points are randomly generated at each simulation. Taking this into account, we have a distribution of the Bayes factor. In order to consider such a statistical fluctuation, we have generated 40 simulated data sets and compute the mean of Bayes factor. Notice that, we examine other values for the number of data sets and the results are quite the same by considering more data sets. The mean Bayes factor \( \ln B_{01} = \ln B_{10} - \ln B_w \) for two strategies, have been shown in Figs (1) and (2) respectively. The value of Bayes factor for each cell is shown as a numerical value on the cell.

From Fig. (1), it is clear that discriminating ΛCDM from wCDM with uncertainty larger than \( 10^{(-1.4)} \sim 4\% \) for a wide range of EoS is almost impossible. With \( 3\% \) uncertainty, we see a strong evidence only for \( w \sim -0.9 \) or \( w \sim -1.1 \), which is already disfavored by other observations. On the other hand, with \( \sigma \leq 1\% \), the chance of discriminating increases significantly. For example with \( \sigma = 0.5\% \), a 2\% deviation in the EoS of dark energy \( (w = 1.02 \text{ or } w = 0.98) \) could be detected with a strong evidence. Notice that, as we mentioned above, the base model for simulated data is the ΛCDM and the Bayes factor is computed for the ΛCDM and wCDM. Contrary, if we consider the wCDM as the base model for simulated data and compute the Bayes factor as \( \ln B_{01} = \ln B_{10} - \ln B_w \), the results are the same. In this case, a positive Bayes factor indicates more evidence in favor of the wCDM.

In addition to the precision of each data point, the sampling rate of a database affects the model comparison. For a less cadence database, the results are presented in (2). Overall, the results are the same as the first sampling strategy but strength of the Bayes factor decreases. For example the extreme case
in the first sampling strategy is $\ln B_{01} \sim 80$ for $\sigma = 0.5\%$ and $w = -0.9$, while considering the second strategy, the number decreases to $\ln B_{01} \sim 33$ which is around 60% less than the former.

B. $\Lambda$CDM and the non-flat $\Lambda$CDM

As we mentioned above, the NACDM can also be written in the form of a GLM. In this case, we follow similar strategies as previous one. The parameter $\Omega_k h^2$ is selected in the range $\Omega_k h^2 \in (-0.05, 0.05)$ to simulate data points. In this case, we simulate the Hubble data using NACDM and compute the evidence in both NACDM and $\Lambda$CDM. In Figs (3) and (4) the mean Bayes factor (averaging over 40 databases) $\ln B_{01} = \ln B_N - \ln B_A$ have been shown for two sampling strategies. Here $B_N$ indicates the evidence of the NACDM model.

FIG. 3. The mean Bayes factor for the $\Lambda$CDM and NACDM using the first sampling strategy.

FIG. 4. The mean Bayes factor for the $\Lambda$CDM and NACDM using the second sampling strategy.

Our results indicate that with $\sigma > 3\%$ there is no chance to discriminate a flat and non flat $\Lambda$CDM models. The evidence become more significant at both positive and negative curvature with smaller uncertainties. We see very strong evidence for $\Omega_0 h^2 \sim 0.2$ with $\sigma < 1\%$ which is much higher for larger and smaller values of $\Omega_0 h^2$. Furthermore, interestingly, we see a negative evidence for a high accuracy data $\sigma < 1\%$ when $\Omega_0 h^2 \sim 0$. This is due to the Occam’s razor effect which favor a simpler model. The effect indicates that an extra free parameter, not being constrained significantly with the data, makes the evidence smaller compare to the model without that free parameter. In these cases, the Bayes factor favor the simpler model which is in our case the flat $\Lambda$CDM. Of course the constrain for $\Omega_0 h^2$ is significantly improved for larger and smaller value of $\Omega_0 h^2$ so the evidence become positive and the more complex model is favored.

The results for the second sampling strategy have been shown in (4). As the previous case, the strength of the evidence decreases for a less cadence sampling rate. In this case, the extreme case decreases around 75% in the second sampling rate.

IV. EVIDENCE FOR CURRENT AVAILABLE HUBBLE DATA

In this section, we apply the GLM approach on the current observational data. The expansion rate database in our analysis is a combination of the Hubble parameter measurement from cosmic chronometers and the BAO measurements. The database has been collected in [33]. In addition, the local $H_0$ measurement (the SHOES data) [5] has been added to the database. Considering the $\Lambda$CDM, wCDM and the non-flat $\Lambda$CDM, the results of MCMC analysis have been shown in the Table (III). To perform the MCMC analysis, we use the public python package pymc3 [34]. In this analysis, we consider the wide Gaussian prior on the free parameters introduced in section III. The 1σ uncertainty of each parameters has been shown along with its mean value. These values are estimated from a sample of parameters generated in the MCMC algorithm.

Now, we use the database and apply the GLM formalism to obtain the MLE, mean and covariance of posterior as well as the evidence. The results have been presented in Tab.(IV). As it is clear, the results from the GLM are quit in agreement with those of the MCMC. Notice that, in case of GLM, we have an analytic posterior distribution for the free parameters and compute the mean and 1σ uncertainty directly from the distributions.

In our analysis, we find $\ln B_{01} = \ln B_A - \ln B_N = 0.43$ which indicates an inconclusive evidence for considered models. In fact, this result was expected because of the low precision observational data points. Notice that the most data points from

| Model/Parameters | $h^2$ | $\Omega_0 h^2$ | $\Omega_k h^2$ |
|------------------|------|----------------|---------------|
| $\Lambda$CDM     | $0.518\pm0.015$ | $0.123\pm0.005$ | -             |
| NACDM            | $0.526\pm0.018$ | $0.135\pm0.016$ | $-0.034\pm0.044$ |

TABLE III. The mean and uncertainty of parameters in $\Lambda$CDM and NACDM using MCMC method.
The landscape of cosmological models has drastically increased in recent years with the combined open problems of cosmological tensions and the internal consistency issues of gravitational models. In this work, we explore the GLM in the context of three cosmological models, namely ΛCDM, non-flat ΛCDM and the wCDM, in order to explore the question of precision requirements for specific data sets to discriminate these models. We wish to assess the data set precision needed to differentiate between each pair of these cosmological models.

Vanilla ΛCDM is our base model for considering any modification to the concordance model. Here, we compare ΛCDM together with wCDM where the equation of state parameter is allowed to be any value. These two models are central to modern cosmology and have Friedmann equations represented by Ω_	ext{CDM} using the GLM method.

| Model/Parameters | $h^2$ | $\Omega_m h^2$ | $\Omega_b h^2$ |
|------------------|------|----------------|----------------|
| ΛCDM             | 0.517 ± 0.015 | 0.121 ± 0.005 | – |
| NACDM            | 0.526 ± 0.018 | 0.134 ± 0.016 | −0.038 ± 0.044 |

**TABLE IV.** The mean and uncertainty of parameters in ΛCDM and NACDM using the GLM method.

cosmic chronometer have uncertainty larger than 10% but uncertainty of the expansion rate data from BAO is less than 10% and the uncertainty of the most precise one is $\sigma \sim 3.5\%$.

**V. CONCLUSION**

The result of our investigation is that precision increases in recent years with the combined open problems of cosmological tensions and the internal consistency issues of gravitational models, and in this work we explore the GLM in the context of three cosmological models, namely ΛCDM, non-flat ΛCDM and the wCDM, in order to explore the question of precision requirements for specific data sets to discriminate these models. We wish to assess the data set precision needed to differentiate between each pair of these cosmological models.

Here, we find that indeed for an equation of state parameter veers away from the ΛCDM value, the approach indicates a high confidence for differentiating these models. Specifically, we see a strong evidence in favor of the ΛCDM when the data simulated with uncertainty $\sigma \sim 0.5$ and the rival model is the wCDM with $w = -1.02$ or $w = -0.98$. Moreover, we show how different sampling rates affect the Bayes factor. The pattern of the results are almost the same for our two sampling strategies but the value of the Bayes factor is smaller in the case of the less cadence sampling.

In Figs. (3,4) we repeat the analysis for the ΛCDM and a non-flat ΛCDM setting. Our results indicate that it is impossible to discriminate these models with $\sigma > 3\%$ even if the $\Omega_m h^2$ is in the range ($-0.05, 0.05$). The strength of the evidence become much more larger for $\Omega_m h^2$ not close to zero with uncertainty $\sigma < 3\%$. In addition, we see clear evidence of Occam’s razor effect when the curvature density is close to zero. To make this procedure as transparent as possible, we are making the code for this analysis available for others to use and improve upon.

---

[1] A. G. Riess et al. (Supernova Search Team), Astron. J. 116, 1009 (1998), arXiv:astro-ph/9805201 [astro-ph].
[2] S. Perlmutter et al. (Supernova Cosmology Project), Astrophys. J. 517, 565 (1999), arXiv:astro-ph/9812133 [astro-ph].
[3] N. Aghanim et al. (Planck), Astron. Astrophys. 641, A6 (2020), arXiv:1807.06269 [astro-ph.CO].
[4] S. Alam et al. (eBOSS), Phys. Rev. D 103, 083533 (2021), arXiv:2007.08991 [astro-ph.CO].
[5] A. G. Riess, S. Casertano, W. Yuan, L. M. Macri, and D. Scolnic, Astrophys. J. 876, 85 (2021), arXiv:1903.07603 [astro-ph.CO].
[6] E. Di Valentino et al., Astropart. Phys. 131, 102606 (2021), arXiv:2008.11283 [astro-ph.CO].
[7] E. Di Valentino et al., Astropart. Phys. 131, 102605 (2021), arXiv:2008.11284 [astro-ph.CO].
[8] A. G. Riess et al., (2021), arXiv:2112.04510 [astro-ph.CO].
[9] S. Weinberg, Rev. Mod. Phys. 61, 1 (1989).
[10] C. Bengaly, (2021), arXiv:2111.06869 [gr-qc].
[11] T. Nadolny, R. Durrer, M. Kunz, and H. Padmanabhan, JCAP 11, 009 (2021), arXiv:2106.05284 [astro-ph.CO].
[12] E. J. Copeland, M. Sami, and S. Tsujikawa, Int. J. Mod. Phys. D 15, 1753 (2006), arXiv:hep-th/0603057.
[13] D. Benisty and D. Staicova, (2021), arXiv:2107.14129 [astro-ph.CO].
[14] D. Benisty and D. Staicova, Astron. Astrophys. 647, A38 (2021), arXiv:2009.10701 [astro-ph.CO].
[15] T. Clifton, P. G. Ferreira, A. Padilla, and C. Skordis, Phys. Rept. 513, 1 (2012), arXiv:1106.2476 [astro-ph.CO].
[16] E. N. Saridakis et al. (CANTATA), (2021), arXiv:2105.12582 [gr-qc].
[17] S. Bahamonde, K. F. Diallektopoulos, C. Escamilla-Rivera, G. Farrugia, V. Gakis, M. Hendry, M. Hohmann, J. L. Said, J. Mifsud, and E. Di Valentino, (2021), arXiv:2106.13793 [gr-qc].
[18] R. Alves Batista et al., (2021), arXiv:2110.10074 [astro-ph.HE].
[19] A. Addazi et al., (2021), arXiv:2111.05659 [hep-ph].
[20] E. Di Valentino, O. Mena, S. Pan, L. Visinelli, W. Yang, A. Melchiorri, D. F. Mota, A. G. Riess, and J. Silk, Class. Quant. Grav. 38, 153001 (2021), arXiv:2103.01183 [astro-ph.CO].
[21] M. Vazirnia and A. Mehrabi, Phys. Rev. D 104, 123530 (2021), arXiv:2107.11539 [astro-ph.CO].
[22] A. Mehrabi and S. Basilakos, Eur. Phys. J. C 80, 632 (2020), arXiv:2002.12577 [astro-ph.CO].
[23] A. Mehrabi et al., Mon. Not. Roy. Astron. Soc. 509, 224 (2021), arXiv:2107.08820 [astro-ph.CO].
[24] W. J. Handley, M. P. Hobson, and A. N. Lasenby, Mon. Not. Roy. Astron. Soc. 450, L61 (2015), arXiv:1502.01856 [astro-ph.CO].
[25] J. S. Speagle, Monthly Notices of the Royal Astronomical Society 493, 3132–3158 (2020).
[26] J. Alsing and W. Handley, Mon. Not. Roy. Astron. Soc. 505, L95 (2021), arXiv:2102.12478 [astro-ph.IM].
[27] R. Trotta, Mon. Not. Roy. Astron. Soc. 378, 72 (2007), arXiv:astro-ph/0504022.
[28] S. Nesseris and J. Garcia-Bellido, JCAP 08, 036 (2013), arXiv:1210.7652 [astro-ph.CO].
[29] R. E. Keeley and A. Shafieloo, (2021), arXiv:2111.04231 [astro-ph.CO].
[30] H. Koo, R. E. Keeley, A. Shafieloo, and B. L’Huillier, (2021), arXiv:2110.10977 [astro-ph.CO].
[31] H. Jeffreys, Theory of Probability, 3rd ed. (Oxford, Oxford, England, 1961).
[32] R. E. Kass and A. E. Raftery, Journal of the American Statistical Association 90, 773 (1995), https://www.tandfonline.com/doi/pdf/10.1080/01621459.1995.10476572.
[33] O. Farooq, F. R. Madiyar, S. Crandall, and B. Ratra, Astrophys. J. 835, 26 (2017), arXiv:1607.03537 [astro-ph.CO].
[34] J. Salvatier, T. V. Wiecki, and C. Fonnesbeck, PeerJ Computer Science 2, e55 (2016).
[35] E. Di Valentino, A. Melchiorri, and J. Silk, Nature Astron. 4, 196 (2019), arXiv:1911.02087 [astro-ph.CO].