Statistical Significance of Lorentz invariance violation from GRB 160625B

Shalini Ganguly[1*] and Shantanu Desai[1†]
[1] *Department of Physics, Indian Institute of Technology, Hyderabad, Telangana-502285, India*

Recently Wei et al [1] have found evidence for a transition from positive time lags to negative time lags in the spectral lag data of GRB 160625B. They have fit these observed lags to a sum of two components: an intrinsic time lag due to astrophysical mechanisms and an energy-dependent speed of light due to violation of Lorentz invariance, which could be a signature of quantum gravity. Here, we examine the statistical significance of the evidence for this claim using the same data by comparing it against the null hypothesis, viz. the time-lags are induced only by intrinsic delays. We use three different model comparison techniques: a frequentist test and two information based criteria (AIC and BIC). From the frequentist model comparison test, we find that evidence for Lorentz violation is favored at 3.05σ and 3.74σ for linear and quadratic models respectively and do not cross the 5σ discovery threshold. We find that ΔAIC and ΔBIC have values $\gtrsim 10$ for the quadratic Lorentz violating model pointing to "decisive evidence" against Lorentz invariance violation compared to only astrophysically induced intrinsic emission. Another concern however is that none of the three models (including the null hypothesis) provide a good fit to the data, which implies that there is additional physics or systematic errors, which are not accounted for while fitting the data to these models.

PACS numbers: 97.60.Jd, 04.80.Cc, 95.30.Sf

I. INTRODUCTION

In special relativity, the speed of light, c, is constant and has the same value in all inertial frames of reference. However, this *ansatz* is no longer true in Lorentz violating standard-model extensions [2] and also several quantum gravity and string theory models (see [3, 4] for reviews). In these models, Lorentz invariance is expected to be broken at very high energies close to the Planck scale, and the speed of light is dependent on the energy of the associated photon [5]. Although many astrophysical sources such as AGNs [6, 7], pulsars [8] etc. have been used to search for Lorentz violation-induced light speed variation, most of these searches have been done with Gamma-Ray Bursts (GRBs) [9–14] and references therein. These searches have been performed using single photons from an ensemble of GRBs as well as with a single GRB. Results from searches for Lorentz violation prior to 2006 or so can be found in the reviews in [3, 4] and references therein. We briefly enumerate some of the key results in the searches for this Lorentz violation since then.

Ellis et al [9] considered a statistical sample of about 60 GRBs at a range of redshifts and modeled the observed time-lag as sum of a constant intrinsic offset and an additional offset due to energy-dependent speed of light. They found 4σ evidence that the higher energy photons arrive earlier than the lower energy ones. The estimated lower limit was about 0.9×10^{16} GeV. However, when an additional systematic offset was added to enforce the χ^2/DOF for the null hypothesis to be of order unity, the statistical significance reduced to about 1σ. Abdo et al [10] used the fact that GeV photons detected by Fermi-LAT from GRB090510 arrived within a one-second interval to set a limit on the Lorentz violating scale of greater than the Planck energy. Chang et al. [11] used the continuum spectrum of about 20 short GRBs detected by the SWIFT satellite to set a constraint on the quantum gravity energy scale of $> 5 \times 10^{14}$ GeV. Zheng and Ma [12] followed a similar procedure as in [9] and fit the observed time lag of 8 GRBs to the sum of an unknown intrinsic offset and an energy dependent time-lag from linear Lorentz violation. They found a linear correlation between the observed time-lag and the energy and redshift-dependent time-lag calculated for every GRB. The slope of this relation was used to obtain a value of about 10^{18} GeV for the linear Lorentz violating scale and the intercept was used to obtain the intrinsic astrophysical offset. Xu et al [13, 14] confirmed this earlier prediction of [12] with 11 GRBs from Fermi-LAT using the same procedure and also found a linear relation between the observed time-lag and the redshift-corrected Lorentz violation factor and obtained a value of 3.6×10^{17} GeV for the energy scale of Lorentz violation. Therefore, we can see that many of these searches for Lorentz invariance violation in the past decade, by looking for energy-dependent speed of light in the arrival times of GRB photons have led to conflicting conclusions.

Most recently however, Wei et al [1] (W17) made a convincing case pertaining to the evidence for a transition from positive to negative time lag in the spectral lag data for GRB 160625B, by using the data from Fermi-LAT and Fermi-GBM. By modeling the time lag as sum of intrinsic time-lag (due to astrophysical processes) and energy-dependent speed of light due to Lorentz invariance violation (LIV), which kicks in at high energies, they argued that this observation constitutes a robust

*E-mail: gangulyshalini1@gmail.com
†E-mail: shntn05@gmail.com

evidence for LIV. Subsequently, constraints on Lorentz violation standard model extension coefficients have been obtained using this data [15]. However, no quantitative assessment of the significance of this claim was made in these papers. Given the potential path-breaking nature of this result, it is important to provide such a test and check if it meets the 5σ criterion for discovery, usually used in particle physics [16]. In this work we examine the statistical significance of this claim by using three different model-comparison tests, namely frequentist hypothesis test, as well as information-criterion based tests.

The outline of this paper is as follows. We provide a succinct introduction to the model comparison techniques used in Section II. We briefly review the observations, data analysis and conclusions reached by W17 in Section III. We then discuss the results from our model comparison tests using the same data in Section IV. Our conclusions can be found in Section V.

II. INTRODUCTION TO MODEL COMPARISON TECHNIQUES

In recent years a number of both Bayesian and frequentist model-comparison techniques (originally developed by the Statistics community) have been applied to a variety of problems in astrophysics, cosmology, and particle physics to address controversial issues. The aims of these techniques is two-fold. One is to find out which among the two hypothesis is favored. A second goal is to assess the statistical significance or p-value of how well the better model is favored. We note however that in many of these applications, not all the techniques used reach the same conclusions. Also the significances from the different techniques could be different. For our purpose, we shall employ multiple available techniques at our disposal to address how significant is the evidence for energy-dependent speed of light. We briefly recap these techniques below. More details on each of these (from a physics/astrophysics perspective) can be found in various reviews [17–19].

- **Frequentist Test**: The first step in a frequentist model comparison test involves constructing a χ^2 between a given model and the data and then finding the best-fit parameters for each model. Then from the best-fit χ^2 and degrees of freedom, one calculates the goodness of fit for each model, given by the χ^2 probability or goodness of fit [20]:

$$P(\chi^2,\nu)=\frac{1}{2^{\frac{\nu}{2}}\Gamma(\nu/2)}(\chi^2)^{\frac{\nu}{2}-1}\exp\Big(-\frac{\chi^2}{2}\Big). \quad (1)$$

where Γ is the incomplete Gamma function and ν is the total degrees of freedom.

The best-fit model is the one with the larger value of χ^2 goodness of fit. If the two models are nested, then from Wilk's theorem [21], the difference in χ^2 between the two models satisfies a χ^2 distribution with degrees of freedom equal to the difference in the number of free parameters for the two hypotheses [17]. Frequentist tests have been used a lot in astrophysics, from testing claims of sinusoidal variations in G as a function of time [22] to classification of GRBs [23].

- **Akaike Information Criterion**: The Akaike Information Criterion (AIC) is used for model comparison, when we need to penalize for any additional free parameters to avoid overfitting. AIC is an approximate minimization of Kullback-Leibler information entropy, which estimates the distance between two probability distributions [19]. For our purpose, we use the first-order corrected AIC, given by [18]:

$$\mathrm{AIC}=\chi^2+2\mathrm{p}+\frac{2\mathrm{p}(\mathrm{p}+1)}{\mathrm{N}-\mathrm{p}-1}, \quad (2)$$

where N is the total number of data points and p is the number of free parameters. A preferred model in this test is the one with the smaller value of AIC between the two hypothesis. From the difference in AIC (Δ AIC), there is no formal method to evaluate a p-value [35]. Only qualitative strength of evidence rules are available depending on the value of ΔAIC [24].

- **Bayesian Information Criterion**: The Bayesian Inference Criterion (BIC) is also used for penalizing the use of extra parameters. It is given by [18]:

$$\mathrm{BIC}=\chi^2+\mathrm{p}\ln\mathrm{N}. \quad (3)$$

Similar to AIC, the model with the smaller value of BIC is the preferred model. The significance is estimated qualitatively in the same way as for AIC. Both AIC and BIC have been used for comparison of cosmological models [24–26].

Besides these techniques, the ratio of Bayesian evidence (or odds ratio) [27] has also been extensively used for model comparison in astrophysics and particle physics [25, 27–29]. However, there have been criticisms regarding the usage of odds ratio for model comparison, since the Bayesian evidence depends on the priors chosen for the parameters [30, 31]. We shall not consider Bayesian evidence in this work.

III. SUMMARY OF W17

W17 have used the spectral lag method to look for energy-dependent time lags in the arrival of photons of a particular GRB (namely GRB 160625B) using data from Fermi-LAT and Fermi-GBM, for which a remarkable transition from positive to negative time lags was observed in the arrival of higher energy photons. The observation of photons from the same source is aimed at providing tighter constraints on Lorentz invariance violation factor. We now briefly describe the *ansatz* made by W17 to fit the spectral lag data.

The observed time lags of photons of varying energies can be written down as :

$$\Delta t_{obs} = \Delta t_{int} + \Delta t_{LIV}, \quad (4)$$

where Δt_{int} is the intrinsic time lag between the emission of photon of a particular energy and the lowest energy photon from the GRB and Δt_{LIV} is the time-lag due to Lorentz invariance violation (hereafter, LIV). The uncertainty associated with Δt_{int} is the most, as it depends upon the internal dynamics of the GRB itself which cannot be obtained from observations. W17 posited the following model for the intrinsic emission delay:

$$\Delta t_{int}(E)(\mathrm{sec}) = \tau \left[\left(\frac{\mathrm{E}}{\mathrm{keV}} \right)^{\alpha} - \left(\frac{\mathrm{E_0}}{\mathrm{keV}} \right)^{\alpha} \right], \quad (5)$$

where E_0=11.34 keV; whereas τ and α are free parameters. The remaining time lag has been attributed to the Lorentz violation effect, occurring at a considerably higher energy (closed to Planck scale) and can be written as [32]:

$$\Delta t_{LIV} = -\frac{1+n}{2H_0} \frac{E^n - E_0^n}{E_{QG,n}^n} \int_0^z \frac{(1+z')^n dz'}{\sqrt{\Omega_M (1+z')^3 + \Omega_\Lambda}}, \quad (6)$$

where $E_{QG,n}$ is the Lorentz-violating or quantum gravity scale, above which Lorentz violation kicks in; H_0 is the Hubble constant. Studying the time lag of photons from a single source does not eliminate the necessity of taking into consideration the delay due to the intrinsic GRB mechanisms. Yet it does provide a more statistically robust method to fit the hypotheses to the observed data. The study of several GRBs beforehand has shown that it is plausible to conclude that there is a positive correlation between the photon energy and its intrinsic time delay [9, 13].

GRB 160625B on the other hand, had three sub-bursts, where the time lag increased upto a certain photon energy after which it dramatically decreased. Using a fitting engine, `McFit` (which uses Monte Carlo approach), W17 obtained the best fit parameters and their uncertainties corresponding to their proposed model consisting of the combined effects of intrinsic time lag and Lorentz violation. As can be found in W17, the χ^2_{dof} values are 2.39 and 2.25 for the linear (n=1) and quadratic (n=2) cases of LIV, respectively. These are however poor fits to the data, corresponding to χ^2 probabilities of 3×10^{-6} and 1.2×10^{-5} respectively. Using the best-fit values of $\log E_{QG,1}$ and $\log E_{QG,2}$ and their 1σ error bars, they obtained a 1σ confidence-level limit on LIV as

$$E_{QG,1} \geqslant 0.5 \times 10^{16} \text{ GeV} \qquad (n=1)$$
$$E_{QG,2} \geqslant 1.4 \times 10^{7} \text{ GeV} \qquad (n=2)$$

The main highlight of W17 was that they did not take into account a constant offset for intrinsic time lag as in Ellis et al [9]. Instead, they proposed a power law function for the intrinsic time lag which fit well with the observed data. They analyzed photons of different energy from the same source as compared to photons from several sources and simultaneously fit for the intrinsic and LIV-induced time lag; and their estimation of behavior of intrinsic time lag helped derive better limits of LIV.

Nevertheless, no estimate of the significance of LIV-induced time lag compared to the null hypothesis of only astrophysically-induced time lag was done in W17. Also all the models proved to be a bad fit to the data. Given the potential path-breaking nature of the result (even though it is based on only one GRB), it is important to independently reproduce the results and estimate the statistical significance of this result, using multiple model comparison methods. This is what has been dealt with in the next section.

IV. ANALYSIS

The first step in frequentist analysis involves parameter estimation for a given hypothesis by minimizing χ^2 between the given model and data. We fit the data to the same three hypotheses as in W17: the time lags are only due to intrinsic astrophysical mechanisms given by Eqn. 5; followed by the hypothesis that the observed time lags consist of sum of intrinsic and LIV-induced time lags for linear (n=1 LIV) as well as quadratic models (n=2 LIV). Once we obtain the best-fit parameters, we then proceed to carry out model comparison using multiple techniques by treating the only intrinsic emission case as the null hypothesis.

The best-fit values for the predicted models were obtained by minimizing the χ^2 functional [20] between the observed model and the data and using the observed errors in the time-lag as the errors in the ordinate. We have assumed that the error bars between the different data points are uncorrelated. We also neglect the error bars in the X-axis. [36] In Eqn. 6, we used $H_0 = 67.3$ km/sec/Mpc and Ω_m=0.315. These are same as those in W17 and inferred from Planck 2015 observations [33].

We fit the 37 spectral lag-energy measurements of GRB 160625B (data obtained from Table 1 of W17 to the three different hypotheses to obtain the optimum values for τ, α and E_{qg}. The best fit values obtained from χ^2 minimization are summarized in Table 1. These mostly agree with the values obtained by W17 [37]. The best-fit curves for all the three models along with the observed spectral lag data are shown in Figure 1. We see that for energies less than 15 MeV, Δt_{obs} is correlated with energy. However above E $\sim$ 15.7 MeV, the observed time lag not only develops a negative correlation but it becomes abruptly negative. The subsequent points after that again have positive values but display a gradual negative correlation with energy.

The best-fit values of χ^2/ DOF for no LIV, LIV (n=1), LIV (n=2) are equal to 2.6, 2.37, and 2.23 respectively (cf. Table II.) For a reasonably good fit, χ^2/DOF has to be close to one [20]. Therefore, none of the models provide a decent fit to the data. The goodness of fit for each of these models is shown in Table II and is less than 10^{-5}.

We then proceed to carry out model comparison, by using the case of no LIV (or only intrinsic astrophysical emission) as the null hypothesis. Among the three models, we compare the χ^2 probability, $P(\chi^2, \nu)$ given by Eqn. 1. From Table II, we see that the model with n=2 LIV has the largest value of $P(\chi^2, \nu)$ and hence can be considered the best model amongst the three. In order to evaluate the statistical significance compared to the null hypothesis, we invoke Wilk's theorem, since the model of no LIV is nested within the n=1 LIV and n=2 LIV models and can be recovered for $E_{QG} = \infty$. To evaluate the significance, we make use of the fact that the difference in χ^2 between the no LIV case and the n=1 LIV and n=2 LIV models follows a χ^2 distribution with degrees of freedom equal to one [17]. From this, we calculate the p-value of n=1, by integrating the χ^2 probability distribution from $\Delta\chi^2$ value between the two models to infinity. The p- values for n=1 LIV and n=2 LIV, compared to no LIV is equal to 0.0014 and 9.2×10^{-5} respectively. One way to interpret the p value (for n=1), is that assuming the null hypothesis is true, the probability that we would see data that favors the model with n=1 LIV simply by chance is 0.0014. We then define significance as the number of standard deviations that a Gaussian variable would fluctuate in one direction to give the same p-value [34]. We find that the significances of n=1 LIV and n=2 LIV correspond to 3.05σ and 3.74σ respectively. Therefore, the frequentist significance does not cross the 5σ discovery threshold used in particle physics.

We have obtained AIC and BIC difference values for these models as opposed to the null hypothesis (cf. Table II). The model with lesser AIC and BIC value is preferred but for our purpose of model comparison against the null hypothesis, we are mostly interested in the difference of AIC and BIC values. Both n=1 and n=2 LIV models have smaller AIC/BIC values compared to the null hypothesis. The ΔAIC and ΔBIC values for n=1 LIV is about 8.5 and 6.9 respectively, which do not correspond to "decisive evidence", according to the qualitative scales indicated in Shi et al [24]. For n=2 LIV, ΔAIC and ΔBIC correspond to 12.9 and 11.7 respectively, which denotes that *empirical support for the models is essentially none* or that *evidence against the models is very strong* in comparison to the null hypothesis.

TABLE I: Best-fit values of the models for the three hypotheses considered. The equation for the time lag with no LIV is described in Eqn. 5. The equations for the two LIV models correspond to Eqn. 4.

	No LIV[a]	**(n=1)**[b]	**(n=2)**[c]
α	0.059	0.175	0.122
τ **(sec)**	5.86	1.24	2.13
E_{qg}/GeV		4.7×10^{15}	1.47×10^{7}

[a] *No Lorentz Invariance*
[b] *Lorentz Invariance up to linear (n=1) order*
[c] *Lorentz Invariance up to quadratic (n=2) order*

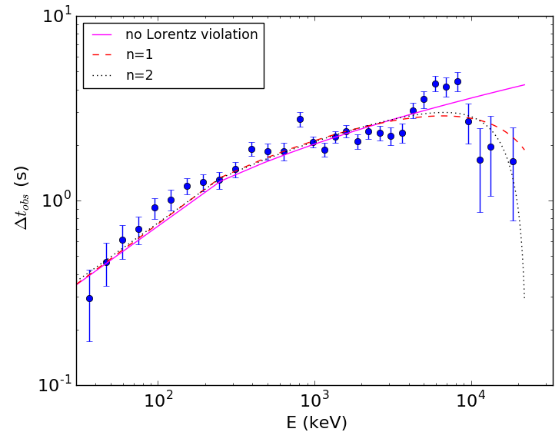

Figure 1 : Summary of the best fit LIV models for $n = 1$ and $n = 2$ along with no Lorentz violation superposed on top of the spectral lag data from GRB 160625B. We note that one data point at (E,Δt) = (15708 keV, -0.223 sec) has been omitted for brevity. All the spectral lag data points have been obtained from Table 1 of W17.

Although, all the three model comparison tests point to n=2 LIV case as the best-fit model, one possible concern is that χ^2/DOF is greater than two and the χ^2 goodness of fit is less than about 10^{-5} for all the three models. This implies that there is additional physics missing or that the data contains unknown systematic error or an intrinsic scatter about the models.

V. CONCLUSIONS

About a year ago Fermi-GBM and Fermi-LAT detected a remarkable Gamma-Ray Burst GRB160625B with three isolated sub-bursts with a total duration of about 770 seconds. This GRB is the only burst so far with a well-defined transition from positive to negative time lags between photons of different energies

This spectral time-lag data was fit by W17 [1] to a model consisting of an intrinsic time lag caused by the astrophysical mechanism related to the GRB emission (see Eqn. 5) and a delay due to energy-dependent speed of light, caused by the violation of Lorentz invariance. This Lorentz violation factor is a function of redshift and also whether a linear or quadratic

TABLE II: Statistical significance of Lorentz invariance violation (LIV) for the two models (linear and quadratic LIV) as opposed to the null hypothesis, i.e. no Lorentz invariance violation using four different model comparison methods. The frequentist significance does not yet cross the 5σ threshold. Both ΔAIC and ΔBIC have values > 10 for the quadratic LIV, pointing to "decisive evidence" using the qualitative strength of evidence rules. However, all the three models have large values of χ^2/DOF. So none of them (in an absolute sense) provide a good fit to the observed data.

	No LIV[a]	**(n=1)**[b]	**(n=2)**[c]
Frequentist			
DOF	35	34	34
χ^2/DOF	2.6	2.37	2.23
χ^2GOF	2.2×10^{-7}	3.7×10^{-6}	1.5×10^{-5}
p-value		0.0014	9.2×10^{-5}
significance		3.05σ	3.74σ
Δ AIC		8.2	12.9
Δ BIC		6.9	11.7

[a] *No Lorentz Invariance*
[b] *Lorentz Invariance up to linear (n=1) order*
[c] *Lorentz Invariance up to quadratic (n=2) order*

model is considered (see Eqn. 6.). A joint fit was done to simultaneously determine the parameters of the intrinsic model and also the energy scale of Lorentz violation (or the quantum gravity scale). Their estimated quantum gravity scale is given by $\log(E_{QG}/\mathrm{GeV}) \sim 15.7$ and ~ 7.2 for a linear and quadratic LIV models respectively. However, the χ^2/DOF for both these models corresponds to 2.39 and 2.25 for linear and quadratic models. Both these models are therefore bad fits to the data with χ^2 probabilities of 3×10^{-6} and 1.2×10^{-5} respectively. No statistical significance was estimated compared to the null hypothesis of no Lorentz violation.

In this work, we redo the same analysis of the spectral lag data from GRB 160625B in order to estimate the statistical significance of the evidence of Lorentz violation. We fit the data to three different models. The first model posits that the time lag is only due to astrophysical emission and is considered the null hypothesis. The other two models involve a sum of the intrinsic mechanism and a linear as well as quadratic LIV model. The parameter estimation for all the three models was done by minimizing the χ^2, similar to what was done in W17. We then carried out three different model comparison tests. The first test involves the frequentist comparison test, where we compare the χ^2 probabilities, which is a proxy for the goodness of fit to determine the model.

Since the null hypothesis is nested within both the LIV models, we use Wilk's theorem to estimate the statistical significance of the two LIV models compared to the null hypothesis of no violation of Lorentz invariance. We find that the χ^2/DOF for the null hypothesis, n=1 LIV, n=2 LIV are equal to 2.6, 2.37, and 2.23 corresponding to χ^2 probabilities of 2.2×10^{-7}, 3.7×10^{-6}, and 1.15×10^{-5} respectively. Therefore, we find in agreement with the results of W17 that all these models are a bad fit to the data. We find that n=2 LIV model has the largest χ^2 goodness of fit and hence can be considered the best model among the three. When we use Wilk's theorem and consider the case of no Lorentz violation as the null hypothesis, we find that the p-values are 1.4×10^{-3} and 9.2×10^{-5} corresponding to 3.05σ and 3.74σ respectively. Therefore, the frequentist significance does not cross the 5σ threshold typically used in particle physics to announce a discovery. We also find that the ΔAIC and ΔBIC are equal to 8.2 and 6.9 in favor of the n=1 LIV model compared to the null hypothesis. For n=2 LIV model, we find that ΔAIC and ΔBIC are equal to 12.9 and 11.7. So the information criterion based values just cross the threshold for "decisive evidence" in favor of the n=2 LIV model.

Therefore, we conclude that the statistical significance of the two Lorentz violating models compared to no Lorentz violation does not cross the 5σ threshold. Only the information criterion based tests just cross the threshold for decisive evidence for the quadratic Lorentz violating model. Another concern however is that none of the three models (including the null hypothesis) provide a good fit to the data, which implies that there is additional Physics or systematic errors which are not accounted for in the time lag data. This certainly needs to be understood before a similar claim can be made with future data.

Acknowledgments

We are grateful to Sabine Hossenfelder for drawing our attention to this result.

[1] J.-J. Wei, B.-B. Zhang, L. Shao, X.-F. Wu, and P. Mészáros, Astrophys. J. Lett. **834**, L13 (2017), 1612.09425.
[2] J. D. Tasson, Reports on Progress in Physics **77**, 062901 (2014), 1403.7785.
[3] D. Mattingly, Living Reviews in Relativity **8**, 5 (2005), gr-qc/0502097.
[4] G. Amelino-Camelia, Living Reviews in Relativity **16**, 5 (2013), 0806.0339.
[5] G. Amelino-Camelia, J. Ellis, N. E. Mavromatos, D. V. Nanopoulos, and S. Sarkar, Nature (London) **395**, 525 (1998).
[6] M. Lorentz, P. Brun, and for the H. E. S. S. Collaboration, ArXiv e-prints (2016), 1606.08600.
[7] J. Ellis and N. E. Mavromatos, Astroparticle Physics **43**, 50 (2013), 1111.1178.
[8] P. Kaaret, Astron. & Astrophys. **345**, L32 (1999), astro-ph/9903464.
[9] J. Ellis, N. E. Mavromatos, D. V. Nanopoulos, A. S. Sakharov, and E. K. G. Sarkisyan, Astroparticle Physics **25**, 402 (2006), astro-ph/0510172.
[10] A. A. Abdo, M. Ackermann, M. Ajello, K. Asano, W. B. Atwood, M. Axelsson, L. Baldini, J. Ballet, G. Barbiellini, M. G. Baring, et al., Nature (London) **462**, 331 (2009), 0908.1832.
[11] Z. Chang, X. Li, H.-N. Lin, Y. Sang, P. Wang, and S. Wang, Chinese Physics C **40**, 045102 (2016), 1506.08495.
[12] S. Zhang and B.-Q. Ma, Astroparticle Physics **61**, 108 (2015), 1406.4568.
[13] H. Xu and B.-Q. Ma, Physics Letters B **760**, 602 (2016), 1607.08043.
[14] H. Xu and B.-Q. Ma, Astroparticle Physics **82**, 72 (2016), 1607.03203.
[15] J.-J. Wei, X.-F. Wu, B.-B. Zhang, L. Shao, P. Mészáros, and V. A. Kostelecký, ArXiv e-prints (2017), 1704.05984.
[16] L. Lyons, ArXiv e-prints (2013), 1310.1284.
[17] L. Lyons, ArXiv e-prints (2016), 1607.03549.
[18] A. R. Liddle, Mon. Not. R. Astron. Soc. **351**, L49 (2004), astro-ph/0401198.
[19] A. R. Liddle, Mon. Not. R. Astron. Soc. **377**, L74 (2007), astro-ph/0701113.
[20] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical recipes in FORTRAN. The art of scientific computing* (1992).
[21] S. S. Wilks, Annals Math. Statist. **9**, 60 (1938).
[22] S. Desai, EPL (Europhysics Letters) **115**, 20006 (2016), 1607.03845.
[23] S. Kulkarni and S. Desai, Astrophys. and Space Science **362**, 70 (2017), 1612.08235.
[24] K. Shi, Y. F. Huang, and T. Lu, Mon. Not. R. Astron. Soc. **426**, 2452 (2012), 1207.5875.
[25] D. L. Shafer, Phys. Rev. D **91**, 103516 (2015), 1502.05416.
[26] S. Desai and D. W. Liu, Astroparticle Physics **82**, 86 (2016), 1604.06758.
[27] R. Trotta, ArXiv e-prints (2017), 1701.01467.
[28] A. Heavens, Y. Fantaye, E. Sellentin, H. Eggers, Z. Hosenie, S. Kroon, and A. Mootoovaloo, ArXiv e-prints (2017), 1704.03467.
[29] J. Martin, C. Ringeval, R. Trotta, and V. Vennin, JCAP **3**, 039 (2014), 1312.3529.
[30] R. D. Cousins, Physical Review Letters **101**, 029101 (2008), 0807.1330.
[31] G. Efstathiou, Mon. Not. R. Astron. Soc. **388**, 1314 (2008), 0802.3185.
[32] U. Jacob and T. Piran, JCAP **1**, 031 (2008), 0712.2170.
[33] P. A. R. Ade et al. (Planck), Astron. Astrophys. **594**, A13 (2016), 1502.01589.
[34] C. Patrignani et al. (Particle Data Group), Chin. Phys. **C40**, 100001 (2016).
[35] See however [25] which posits a significance based on $\exp(-\Delta AIC/2)$
[36] We also verified that our results do not change significantly by including the errors in E which in this case correspond to the bin width, by modifying our χ^2 to include errors in X-axis following the same prescription as in [22].
[37] We do not report 1σ error bars, since our main goal is model comparison and not parameter estimation