

No Evidence for Extensions to the Standard Cosmological Model

Alan Heavens,^{1,*} Yabebal Fantaye,^{2,3} Elena Sellentin,^{1,4} Hans Eggers,^{5,6} Zafiiroh Hosenie,^{7,2,8}
Steve Kroon,⁹ and Arrykrishna Mootoovaloo^{10,2,8}

¹Imperial Centre for Inference and Cosmology (ICIC), Imperial College,

Blackett Laboratory, Prince Consort Road, London SW7 2AZ, United Kingdom

²African Institute for Mathematical Sciences, 68 Melrose Road, Muizenberg 7945, South Africa

³Department of Mathematics, Stellenbosch University, P/Bag X1, 7602 Matieland, South Africa

⁴Département de Physique Théorique, Université de Genève, Quai Ernest-Ansermet 24 CH-1211 Genève, Switzerland

⁵Department of Physics, Stellenbosch University, P/Bag X1, 7602 Matieland, South Africa

⁶National Institute for Theoretical Physics, Stellenbosch University, P/Bag X1, 7602 Matieland, South Africa

⁷Centre for Space Research, North-West University, Potchefstroom 2520, South Africa

⁸South African Astronomical Observatory, Observatory Road, Observatory, Cape Town 7935, South Africa

⁹CSIR-SU Centre for AI Research, Computer Science Division, Stellenbosch University, P/Bag X1, 7602 Matieland, South Africa

¹⁰Department of Mathematics and Applied Mathematics, University of Cape Town, Rondebosch, Cape Town 7700, South Africa

(Received 11 April 2017; revised manuscript received 11 July 2017; published 7 September 2017)

We compute the Bayesian evidence for models considered in the main analysis of Planck cosmic microwave background data. By utilizing carefully defined nearest-neighbor distances in parameter space, we reuse the Monte Carlo Markov chains already produced for parameter inference to compute Bayes factors B for many different model-data set combinations. The standard 6-parameter flat cold dark matter model with a cosmological constant (Λ CDM) is favored over all other models considered, with curvature being mildly favored only when cosmic microwave background lensing is not included. Many alternative models are strongly disfavored by the data, including primordial correlated isocurvature models ($\ln B = -7.8$), nonzero scalar-to-tensor ratio ($\ln B = -4.3$), running of the spectral index ($\ln B = -4.7$), curvature ($\ln B = -3.6$), nonstandard numbers of neutrinos ($\ln B = -3.1$), nonstandard neutrino masses ($\ln B = -3.2$), nonstandard lensing potential ($\ln B = -4.6$), evolving dark energy ($\ln B = -3.2$), sterile neutrinos ($\ln B = -6.9$), and extra sterile neutrinos with a nonzero scalar-to-tensor ratio ($\ln B = -10.8$). Other models are less strongly disfavored with respect to flat Λ CDM. As with all analyses based on Bayesian evidence, the final numbers depend on the widths of the parameter priors. We adopt the priors used in the Planck analysis, while performing a prior sensitivity analysis. Our quantitative conclusion is that extensions beyond the standard cosmological model are disfavored by Planck data. Only when newer Hubble constant measurements are included does Λ CDM become disfavored, and only mildly, compared with a dynamical dark energy model ($\ln B \sim +2$).

DOI: 10.1103/PhysRevLett.119.101301

Introduction.—The standard cosmological model of a flat universe containing additional cold dark matter and a cosmological constant (Λ CDM) is a remarkably simple and successful description of the Universe. Based on CDM and a cosmological constant Λ , this flat model has six free parameters, which the Planck satellite has measured with very high precision. Extensions of the standard model have also been introduced, to relieve tensions with other data sets that have arisen with the standard model, or to probe for new physics, and in this respect model comparison is of more fundamental interest than parameter inference. Bayesian evidence (or marginal likelihood) is the Bayesian tool to address such questions, and it can be challenging to compute as it requires integration of the likelihood over the multidimensional parameter space. In a companion paper [1] we show how Monte Carlo Markov chains (MCMCs), produced for parameter inference, can be used to perform model comparison. In this Letter, we report

an analysis of all the main published Planck chains. Bayesian evidence has been computed for a small number of models, e.g., by [2,3] for curvaton models, [4] for holographic models, [5] for a comprehensive study of inflation models, and [6] for inflationary features, [7–9], which focused on neutrino extensions. However this is the first comprehensive study of the models and data sets considered by Planck.

Bayesian evidence.—The goal of parameter inference is to determine the posterior probability of model parameters θ , given a data set \mathbf{x} , any prior information, and a model M . Using Bayes theorem, this is

$$p(\theta|\mathbf{x}, M) = \frac{p(\mathbf{x}|\theta, M)\pi(\theta|M)}{p(\mathbf{x}|M)}, \quad (1)$$

where π is the prior and $p(\mathbf{x}|\theta, M)$ is the likelihood, which is regarded as a function of θ . The Bayesian evidence $p(\mathbf{x}|M)$ is used for model comparison,

and is the integral over the un-normalized posterior $\tilde{p}(\boldsymbol{\theta}|\mathbf{x}, M) \equiv p(\mathbf{x}|\boldsymbol{\theta}, M)\pi(\boldsymbol{\theta}|M)$. It trades the typically higher likelihood of the more complex model against the increased prior volume,

$$p(\mathbf{x}|M) = \int d\boldsymbol{\theta} p(\mathbf{x}|\boldsymbol{\theta}, M)\pi(\boldsymbol{\theta}|M). \quad (2)$$

The posterior probability of competing models is then given by the product of the ratio of the model priors and the ratio of evidences (the latter being known as the Bayes factor),

$$\frac{p(M_1|\mathbf{x})}{p(M_2|\mathbf{x})} = \frac{\pi(M_1)}{\pi(M_2)} \frac{p(\mathbf{x}|M_1)}{p(\mathbf{x}|M_2)}. \quad (3)$$

The evidence may be expensive to compute if the dimensionality of the parameter space is large. Also we typically do not know \tilde{p} , but have only samples of it obtained by MCMC techniques. Many such chains exist for the Planck data, for various data set-model combinations. In this application the standard model is a special case of the extended models, so the maximum likelihood of the extended model is at least as high as the standard model, so it is of limited use, whereas the Bayesian evidence, which includes an element of Occam's razor, quantifies whether the increase of likelihood throughout the parameter space warrants support for the more complex model.

Conventional wisdom is that MCMC chains are not good for computing the evidence, as it is claimed that they do not explore the tails of the distributions well; i.e., marginalizing over all parameters is thought not to be possible. However, it is common to marginalize over all but one or two parameters, to obtain marginal posteriors for parameters individually or in pairs, and the tails do not seem to be a concern in those cases. The real issue with Bayesian evidence is the normalization of the integral. The MCMC samples from the un-normalized posterior; i.e., the chain is a sample from a number density $n(\boldsymbol{\theta})$ that is proportional to the un-normalized posterior, $\tilde{p} = an$ but with an unknown constant of proportionality a . If this constant can be determined, then the evidence is readily computed, since, replacing n by the sample density (a sum of Dirac delta functions),

$$E = \int d\boldsymbol{\theta} an(\boldsymbol{\theta}) = a \int d\boldsymbol{\theta} \sum_{\alpha=1}^N \delta(\boldsymbol{\theta} - \boldsymbol{\theta}_\alpha) = aN, \quad (4)$$

where N is the length of the chain. Alternatively, since $n = Np(\boldsymbol{\theta}|\mathbf{x}, M) = N\tilde{p}/E$, $E = N\tilde{p}/n \equiv aN$.

We use the MCEvidence algorithm presented in [1], where the nearest-neighbor distances D , in an m -dimensional MCMC chain, are used in a Bayesian analysis to infer a . The chain is prewhitened such that the covariance matrix of the points is the identity, and a Euclidean distance measure then employed. This is equivalent to using the Mahalanobis distance, where the inverse covariance matrix defines the metric. The resulting posterior for the Bayesian evidence E

(assuming a $1/E$ prior) is given by a sum over the MCMC points α , each weighted by w_α ,

$$\ln p(E|\{D_\alpha\}) = \text{const} - (N+1) \ln E - \frac{W}{E} \sum_{\alpha=1}^N \frac{V_m(D_\alpha)\tilde{p}_\alpha}{w_\alpha}, \quad (5)$$

where $V_m(D) = \pi^{m/2}D^m/\Gamma(1+m/2)$ is the volume of an m ball of radius D and $W \equiv \sum_\alpha w_\alpha$. We assume that the MCMC algorithm produces independent distances, but we test this later. See [1] for more details.

The maximum *a posteriori* value of E is

$$E_{\text{MAP}} = \frac{W}{N+1} \sum_{\alpha=1}^N \frac{V_m(D_\alpha)\tilde{p}_\alpha}{w_\alpha}, \quad (6)$$

with a statistical variance in $\ln E$ of $1/(N+1)$. We marginalize over nuisance parameters, and run a nearest-neighbor distance algorithm to determine D_α , then compute the posterior for the evidence. Marginalizing over nuisance parameters adds scatter to \tilde{p} , which increases the statistical error on E . We have checked the effect of including some nuisance parameters and most $\ln B$ values change by <0.1 . Larger changes (up to about unity) occur only when $|\ln B|$ is itself very large.

Data and models.—The Planck chains use a variety of data sets, which are detailed in [10].

The models are the “base” flat Λ CDM model, with parameters $\omega_b = \Omega_b h^2$, θ_{MC} , $\omega_c = \Omega_c h^2$, τ , $\ln(10^{10}A_s)$ and n_s , representing baryon density, CDM density, angle of the first peak, optical depth, amplitude of primordial fluctuations, and scalar spectral index. h is the Hubble parameter $H_0/100 \text{ km s}^{-1} \text{ Mpc}^{-1}$. The other models consist of the base model with one to three extensions. Extensions and prior ranges are listed in Table I. It takes about 15 minutes on a laptop to analyze all the chains with MCEvidence.

Results.—Table II shows Bayes factors with respect to the favored model for selected data sets. Bayes factors for other *Planck* data sets are available online. Internal analysis indicates a typical statistical error of ~ 0.02 . The scatter in the mean of the MAP estimates from individual chains is typically 0.02–0.1. In the revision [11] of the Jeffreys scale, $|\ln B| > 3$ is regarded as strong evidence (relative probability >20), and $|\ln B| > 5$ as very strong (relative probability >150). As is always the case, the Bayesian evidence depends on the priors chosen. Here we have used uniform box priors (in addition, cosmoMC excludes some physically impossible subregions) based on the 2015 Planck analysis (Antony Lewis), and these are listed in Table I. For parameter inference these are not important if the data are informative, and some are set to wide uniform priors. In high dimensions the evidence is very prior dependent, but Bayes factors depend only on the width of the prior of the (usually one) additional parameter(s). $\ln B$ values can be adjusted to a

TABLE I. Cosmological parameters (adapted from [10]), their prior range in square brackets, the baseline values assumed, the nomenclature used in the model extensions in Table II, and a summary definition (see the text for details). The flat Λ CDM base model is parametrized by the parameters above the horizontal line. For completeness priors are given for all relevant extensions, even if the models are not discussed in this Letter.

Parameter	Prior range	Baseline	Nomenclature	Definition
$\omega_b \equiv \Omega_b h^2$	[0.005, 0.1]	...		Baryon density today
$\omega_c \equiv \Omega_c h^2$	[0.001, 0.99]	...		Cold dark matter density today
$100\theta_{\text{MC}}$	[0.5, 10]	...		100 \times approximation to r_*/D_A (CosmoMC)
τ	[0.01, 0.8]	...		Thomson scattering optical depth due to reionization
$\ln(10^{10}A_s)$	[2, 4]	...		Log power of the primordial curvature perturbations ($k_0 = 0.05 \text{ Mpc}^{-1}$)
n_s	[0.8, 1.2]	...		Scalar spectrum power-law index ($k_0 = 0.05 \text{ Mpc}^{-1}$)
Ω_K	[-0.3, 0.3]	0	omegak	Curvature parameter today with $\Omega_{\text{tot}} = 1 - \Omega_K$
$\sum m_\nu$	[0, 5]	0.06	mnu	The sum of neutrino masses in eV
$m_{\nu, \text{sterile}}^{\text{eff}}$	[0, 3]	0	meffsterile	Effective mass of sterile neutrino in eV
w_0	[-3, 1]	-1	w	Dark energy equation of state, $w(a) = w_0 + (1-a)w_a$
w_a	[-3, 2]	0	w_wa	As above (perturbations modeled using PPF)
N_{eff}	[0.05, 10]	3.046	nnu	Effective number of neutrino-like relativistic degrees of freedom
Y_P	[0.1, 0.5]	BBN	yhe	Fraction of baryonic mass in helium
α_{-1}	[-1, 1]	0	alpha1	Correlated isocurvature parameter
A_L	[0, 10]	1	Alens	Amplitude of the lensing power relative to the physical value
A_L^{4pt}	[0, 10]	1	Aphphi	Amplitude of the lensing power from the four-point function relative to the physical value
$dn_s/d \ln k$	[-1, 1]	0	nrun	Running of the spectral index
$r_{0.05}$	[0, 3]	0	r	Ratio of tensor primordial power to curvature power at $k_0 = 0.05 \text{ Mpc}^{-1}$

new prior range $\Delta\theta$ (assuming it is larger than the extent of the chain), by adding $\ln(\Delta\theta_{\text{Table I}}/\Delta\theta)$ to the table values. A narrower prior range increases the relative probability of the extended model. Some models are not excluded with high probability, and one should be cautious of these, given the prior uncertainty. However, we see that Planck data provide very strong evidence against some models (note that in the text we discuss Bayes factors compared with flat Λ CDM. In Table II the numbers quoted are with respect to the highest evidence model, which for a few data sets is not the base model): e.g., correlated isocurvature models are disfavored with $\ln B = -7.8$ by Planck alone, using polarization; running of the spectral index is disfavored (-4.7); and the evidence is against nonstandard neutrino masses and number (-6.5). The evidence against nonzero $r_{0.05}$ is strong compared with the base model (-4.3).

At a weaker level, we find that a nonflat universe is marginally preferred over the standard model by Planck data until cosmic microwave background (CMB) lensing is included, after which there is strong evidence in favor of the flat standard model (-3.6). There is no evidence for nonstandard lensing (-4.6), or for varying the number of neutrino species (-3.1) or masses (-3.2). A model including massive extra neutrinos, which was introduced to alleviate tension with direct Hubble constant measurements (e.g., [12]) and low-amplitude cosmic shear (e.g., [13–16]) has a very low probability from Planck data alone ($\ln B = -6.9$ with respect to base), and a model that also

has nonzero tensor-to-scalar ratio r has exceptionally strong evidence against (-10.8). The results are shown in Fig. 1.

These results are understood when compared with the marginal distributions of parameters from the Planck chains. For example, in the absence of CMB lensing the posterior for Ω_m and Ω_Λ follows the geometric degeneracy line (Fig. 26 of [17]), with most of the probability lying away from the flat line in this plane. When CMB lensing is included, the posterior is concentrated close to the intersection of the lines, and the Bayesian evidence favors the flat model. Correlated isocurvature modes are similarly constrained to be very close to zero amplitude (Fig. 24 of [17]), when the temperature-E mode cross-power (TE) and E-mode power spectrum (EE) are included. Similar observations can be made for other parameters. Note that naive interpretations of credible intervals may not be supported by the evidence. One example of this is the lensing power amplitude, for which $A_L = 1$ is in the tails of the distribution for some data sets (see Fig. 12 of [17]), but for which the evidence favors the standard model. See Fig. 3 of [18] for illustration of this general point.

Robustness tests.—MCEvidence assumes that the points in the chain are independent, and this is not strictly the case. To test the effect of this we have computed the correlations to find that they are generally small, but we have also aggressively thinned the chains. Thinning by a factor of 10 makes little difference to the results, with the vast majority

TABLE II. Bayes factors with respect to the model favored by each data set (flat Λ CDM except for Planck only without lensing). For full description of data sets, see the Planck descriptions, where names should be preceded by *base_plikHM_*; in short BAO = baryon acoustic oscillations, lensing = CMB lensing, JLA = supernovae, H070p6 = Hubble constant prior centred on 70.6, and zre6p5 = recombination at $z = 6.5$. For model nomenclature, see Table I. Bayes factors in bold are referred to in the text, adjusted where necessary if the Λ CDM evidence is not the highest in the column (designated by italics). This is a subset of the model or data set combinations considered in this analysis.

$\ln B$	TT_lowTEB	TT_lowTEB_post_BAO	TT_lowTEB_post_lensing	TT_lowTEB_post_H070p6	TT_lowTEB_post_JLA	TT_lowTEB_post_zre6p5	TT_lowTEB_BAO	TT_lowTEB_BAO_post_lensing	TT_lowTEB_lensing	TT_lowTEB_lensing_post_BAO	TT_tau07	TT_lowTEB_lensing_BAO	TTTTTEE_lowTEB	TTTTTEE_lowTEB_post_BAO	TTTTTEE_lowTEB_post_lensing	TTTTTEE_lowTEB_post_H070p6	TTTTTEE_lowTEB_post_JLA	TTTTTEE_lowTEB_post_zre6p5	TTTTTEE_lowTEB_lensing
1 base	-0.7	0.0	0.0	0.0	0.0	0.0			0.0	0.0	0.0		-0.2	0.0	0.0	0.0	0.0	0.0	0.0
2 omegak	0.0						-1.7	-1.5	-3.6				0.0						-3.6
3 Alens	-1.7	-0.9		-0.3	-0.5	-0.8			-4.2	-4.1	-2.5		-2.1	-1.5		-1.5	-1.5	-1.8	-4.6
4 nnu	-3.4		-1.7	-2.5	-2.5	-2.5		0.0			-2.2		-3.3		-1.7	-3.0	-2.9	-2.9	
5 mnu	-3.7			-3.2	-3.2	-2.8		-0.5		-2.6		0.0	-3.4					-3.0	-2.9
6 nrun	-5.1	-4.1	-3.8	-4.0	-4.0	-4.1					-4.4		-4.9	-4.5	-3.8	-4.5	-4.5	-4.5	
7 r	-4.9	-4.0		-4.0	-4.0	-4.0			-4.0	-3.9			-4.5	-4.1		-4.1	-4.1	-4.1	-4.1
8 w	-1.2		0.0	-1.7	-3.2	-0.2	-0.3	-0.2					-0.7		0.0	-1.9		-0.3	
9 alpha1	-6.4	-5.6	-5.1	-5.5	-5.5	-5.4							-8.0	-7.6	-6.2	-7.5	-7.6	-7.6	
10 Ahiphi									-3.9										-3.8
11 yhe	-2.9	-2.0	-1.1	-1.8	-1.9	-1.9							-2.9	-2.5	-1.5	-2.5	-2.5	-2.5	
12 mnu_Alens													-3.6						
13 mnu_omegak													-4.9						
14 mnu_w													-3.1						
15 nnu_mnu	-6.6	-6.1		-5.7					-5.4	-6.0			-6.6	-6.5		-6.3			-6.0
16 nnu_yhe	-5.2	-4.4	-3.4	-4.2									-5.1	-4.6	-3.3	-4.8			
17 w_wa							-0.1	-0.3											
18 nnu_meffsterile	-6.5						-3.2		-5.7			-2.4	-7.1						-6.6
19 nnu_meffsterile_r									-9.9										-10.8

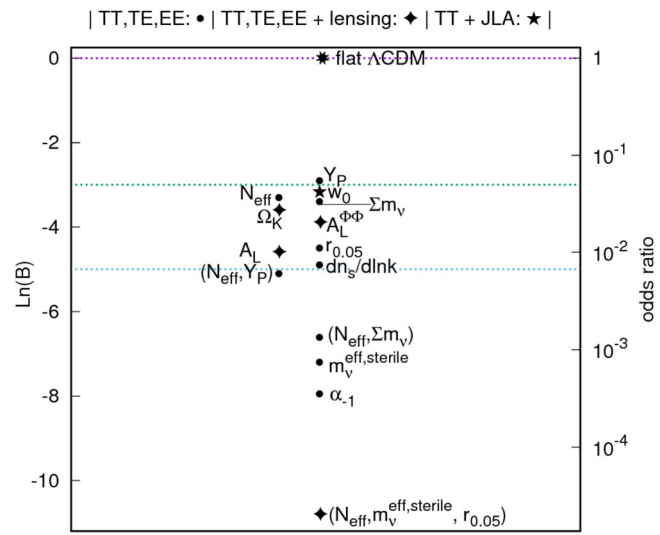


FIG. 1. Bayes factors $\ln B$ with respect to the highest evidence model (base: flat Λ CDM). The most constraining data set is indicated by the symbol; see the legend for details. Horizontal lines mark the boundaries corresponding to strong ($\ln B < -3$) and very strong (< -5) evidence in the Kass and Raftery [11] scale.

of $\ln B$ values changing by < 0.2 , with very few over 0.5. A few heavily disfavored models change by more, up to 0.7, so the conclusions are robust. We also note that if weights are set to unity in Planck chains the same conclusions are reached.

Since the Bayes factors depend on the width of the prior for the additional parameter(s), we can ask by how much they have to be changed for the models to be preferred over the standard model. The most competitive models are disfavored with $\ln B \approx -3$, which requires the prior range to be reduced by a factor 20 for them to be more probable than flat Λ CDM. For example, w would need to be restricted to a prior range < 0.2 , Y_p to 0.02, and $r_{0.05}$ to 0.03, within the current credible region. The neutrino mass conclusion is least secure, as a prior range less than 0.2 eV would favor nonstandard masses.

Discussion.—The main aim of this paper is to compute Bayesian evidence values for the many models and data sets produced in the primary Planck analysis, where we find that the 6-parameter flat Λ CDM model is preferred, with no evidence in favor of extensions. As is usual with evidence calculations, the results sometimes favor simpler models

even when naive inspection of credible intervals suggests otherwise. We agree with the conclusions of [7,8] but not [19] in disfavoring adding extra massive neutrino components to the base model, but our conclusions are far more wide ranging. We also complement the analysis of [20] that shows a model-independent lack of evidence for deviations from standard physical parameters. The inclusion of recent Hubble constant measurements [12,21] (the latter with or without outliers) favors Λ CDM over w CDM, but only with modest odds ($\ln B = 2.2, 2.0, 1.5$), respectively, from chains that allow all parameters to vary. We do not include strong lensing constraints on H_0 (e.g., [22]) as the constraints are model dependent so are not straightforward to add.

MCEvidence is written in PYTHON and is freely available following the link in Ref. [23]. The full table of evidence results is linked from [24].

We thank Andrew Jaffe, Bruce Bassett, and Raul Jimenez for helpful discussions, and Antony Lewis for 2015 Planck priors. Y.F. is supported by the Robert Bosch Stiftung under the AIMS ARETÉ Junior Research Chair Program. Z.H. is funded by a SKA SA masters bursary and NWU international bursary. A.M. is funded by a SKA SA masters bursary. E.S. is funded by a DAAD research fellowship of the German Academic Exchange Service. We thank referees for helpful comments.

* a.heavens@imperial.ac.uk

- [1] A. F. Heavens, Y. Fantaye, A. Mootooyaloo, H. Eggers, Z. Hosenie, S. Kroon, and E. Sellentin, [arXiv:1704.03472](https://arxiv.org/abs/1704.03472).
- [2] R. Hardwick and C. Byrnes, *J. Cosmol. Astropart. Phys.* **8** (2015) 010.
- [3] V. Vennin, K. Koyama, and D. Wands, *J. Cosmol. Astropart. Phys.* **11** (2015) 008.

- [4] N. Afshordi, C. Corianò, L. Delle Rose, E. Gould, and K. Skenderis, *Phys. Rev. Lett.* **118**, 041301 (2017).
- [5] J. Martin, C. Ringeval, and V. Vennin, *Phys. Dark Universe* **5–6**, 75 (2014).
- [6] M. Benetti and J. Alcaniz, *Phys. Rev. D* **94**, 023526 (2016).
- [7] S. Feeney, H. Peiris, and L. Verde, *J. Cosmol. Astropart. Phys.* **4** (2013) 036.
- [8] B. Leistedt, H. Peiris, and L. Verde, *Phys. Rev. Lett.* **113**, 041301 (2014).
- [9] F. Simpson, R. Jimenez, C. Pena-Garay, and L. Verde, *J. Cosmol. Astropart. Phys.* **06** (2017) 029.
- [10] Planck Collaboration XVI, *Astron. Astrophys.* **571**, A16 (2014).
- [11] R. E. Kass and A. Raftery, *J. Am. Stat. Assoc.* **90**, 773 (1995).
- [12] A. Riess *et al.*, *Astrophys. J.* **826**, 56 (2016).
- [13] J. Hamann and J. Hasenkamp, *J. Cosmol. Astropart. Phys.* **10** (2013) 044.
- [14] H. Hildebrandt *et al.*, *Mon. Not. R. Astron. Soc.* **465**, 1454 (2017).
- [15] L. Verde, P. Protopapas, and R. Jimenez, *Phys. Dark Universe* **2**, 166 (2013).
- [16] M. Wyman, D. Rudd, R. A. Vanderveld, and W. Hu, *Phys. Rev. Lett.* **112**, 051302 (2014).
- [17] Planck Collaboration XIII, *Astron. Astrophys.* **594**, A13 (2016).
- [18] R. Trotta, *Contemp. Phys.* **49**, 71 (2008).
- [19] R. A. Battye and A. Moss, *Phys. Rev. Lett.* **112**, 051303 (2014).
- [20] A. F. Heavens, R. Jimenez, and L. Verde, *Phys. Rev. Lett.* **113**, 241302 (2014).
- [21] S. Feeney, D. Mortlock, and N. Dalmaso, [arXiv:1707.00007](https://arxiv.org/abs/1707.00007).
- [22] V. Bonvin *et al.*, *Mon. Not. R. Astron. Soc.* **465**, 4914 (2017).
- [23] See <https://github.com/yabebalFantaye/MCEvidence>.
- [24] <http://astro.ic.ac.uk/aheavens/home>.