Information gain on reheating: The one bit milestone

Jérôme Martin∗
Institut d’Astrophysique de Paris, UMR 7095-CNRS,
Université Pierre et Marie Curie, 98 bis boulevard Arago, 75014 Paris, France

Christophe Ringeval†
Centre for Cosmology, Particle Physics and Phenomenology,
Institute of Mathematics and Physics, Louvain University,
2 Chemin du Cyclotron, 1348 Louvain-la-Neuve, Belgium

Vincent Vennin‡
Institute of Cosmology & Gravitation, University of Portsmouth,
Dennis Sciama Building, Burnaby Road, Portsmouth, PO1 3FX, United Kingdom
(Dated: September 16, 2016)

We show that the Planck 2015 and BICEP2/KECK measurements of the cosmic microwave background (CMB) anisotropies provide together an information gain of 0.82±0.13 bits on the reheating history over all slow-roll single-field models of inflation. This corresponds to a 40% improvement compared to the Planck 2013 constraints on the reheating. Our method relies on an exhaustive CMB data analysis performed over nearly 200 models of inflation to derive the Kullback-Leibler entropy between the prior and the fully marginalized posterior of the reheating parameter. This number is a weighted average by the Bayesian evidence of each model to explain the data thereby ensuring its fairness and robustness.

PACS numbers: 98.80.Cq, 98.70.Vc

I. INTRODUCTION

The release of the Planck 2015 data and its awaited CMB polarization measurements have provided an unprecedented understanding of the Universe content from the time of last scattering to today [1]. In addition to having solved the puzzle concerning the origin of the B-modes detection by the BICEP2 telescope [2–7], the Planck 2015 data have increased the effective number of modes, i.e., of measured $a_{\ell m}$, by 55% compared to the Planck 2013 release [8]. As shown in Ref. [9], this translates into a reduction of uncertainties by typically one-sigma on essentially all the cosmological parameters and, in particular, on the scalar spectral index and tensor-to-scalar ratio. Meanwhile, all tests performed so far by the Planck Collaboration to search for any deviations with respect to the predictions of slow-roll single-field inflation, such as presence of isocurvature modes or features, remain null [10]. Although any measurements of primordial non-Gaussianities would provide an overwhelming amount of information onto nonlinear physics during cosmic Inflation, their nondetection does not mean that the microphysics of the early Universe remains invisible.

As discussed in Refs. [11, 12] (see also subsequent works [13–16]), the epoch of reheating, the era during which the vacuum energy of the inflaton is converted into radiation, impacts the predicted values of the scalar spectral index and tensor-to-scalar ratio for all inflationary models. As a result, reducing uncertainties on the measured values of these two parameters provides some non-trivial information on how reheating proceeds, even if inflation is as simple as a slow-roll single-field scenario [17]. All kinematic effects from the reheating modify the CMB predictions through one, and only one, parameter, $R_{\text{reh}}$ defined by [12]

$$\ln R_{\text{reh}} = \ln R_{\text{rad}} + \frac{1}{4} \ln \left( \frac{\rho_{\text{end}}}{M_{\text{Pl}}^2} \right),$$

where $R_{\text{rad}}$ is a peculiar combination of reheating quantities

$$\ln R_{\text{rad}} = \frac{\Delta N_{\text{reh}}}{4} \left( 3\overline{w}_{\text{reh}} - 1 \right) = \frac{1 - 3\overline{w}_{\text{reh}}}{12(1 + \overline{w}_{\text{reh}})} \ln \left( \frac{\rho_{\text{reh}}}{\rho_{\text{end}}} \right).$$

Here, $\Delta N_{\text{reh}} = N_{\text{reh}} - N_{\text{end}}$ is the duration of reheating in number of $e$-folds, $\overline{w}_{\text{reh}}$ is the mean (i.e., averaged over number of $e$-folds) equation of state parameter of the Universe during this epoch, $\rho_{\text{reh}}$ is the energy density at the end of reheating, defined to be the beginning of the radiation era, and $\rho_{\text{end}}$ is the energy density at the end of inflation.

Within a given model of inflation, predicting the scalar spectral index and tensor-to-scalar ratio requires us to specify a value for the so-called rescaled reheating parameter $R_{\text{reh}}$ [11]. More specifically, once $R_{\text{reh}}$ is given, one can uniquely determine the number of $e$-folds before the end of inflation at which an observable mode

∗Electronic address: jmartin@iap.fr
†Electronic address: christophe.ringeval@uclouvain.be
‡Electronic address: vincent.vennin@port.ac.uk
$k_*$ crosses the Hubble radius [18], a quantity which is required to get the actual values of the observed slow-roll parameters [19, 20]. As discussed in Ref. [21], in the absence of any information on how reheating proceeds, one can use a set of minimal assumptions. Reheating should occur after inflation and before big-bang nucleosynthesis such that $\rho_{\text{nuc}} < \rho_{\text{reh}} < \rho_{\text{end}}$. Moreover, from the energy positivity conditions in general relativity, and the definition of reheating which is not inflation, one has $-1/3 < \bar{m}_{\text{reh}} < 1$. If one makes the conservative choice $\frac{1}{4}$ $\rho_{\text{nuc}} = 10$ MeV, one obtains

$$-46 < \ln R_{\text{reh}} < 15 + \frac{1}{3} \ln \left( \frac{\rho_{\text{end}}}{M_{\text{Pl}}} \right) . \quad (3)$$

These bounds define a flat prior probability distribution $\pi(\ln R_{\text{reh}})$ on the parameter $\ln R_{\text{reh}}$, i.e., a Jeffreys’ prior on the rescaled reheating parameter $R_{\text{reh}}$. Let us notice that, within a given model of inflation, $\rho_{\text{end}}$ is a theoretical output and not an additional parameter. This however implies that the prior on the rescaled reheating parameter has an upper bound which is model dependent. Along these lines, performing a CMB data analysis within one model of inflation allows one to infer, among others, the marginalized posterior probability distribution $P(\ln R_{\text{reh}}|D)$ for the parameter $\ln R_{\text{reh}}$ under the data set $D$. The data are constraining the reheating epoch as soon as the posterior $P$ “is peaked” compared to the prior $\pi$. In Ref. [21], we have followed this route using the Planck 2013 data for almost 200 models of inflation taken from the Encyclopaedia Inflationaris collection [22]. Over all these models, the Planck 2013 data have been shown to give a reduction factor of 40% in the ratio of standard deviations of $\ln R_{\text{reh}}$ between the prior $\pi$ and the inferred posterior $P$.

Although such a result shows that the Planck data yield significant constraints on reheating, the ratio of standard deviations is wasting some amount of information. For instance, if the posterior $P(\ln R_{\text{reh}}|D)$ is multi-valued, i.e., has more than one maximum, the standard deviation within the posterior could remain as large as within a flat prior. In such a situation, the data would be disfavoring some intermediate ranges of values and, thus, would contain some information not accounted for in the standard deviation. This is why in the present work we prefer to use the Kullback-Leibler divergence [23] between the prior distribution $\pi$ and the posterior $P$,

$$D_{\text{KL}} = \int P(\ln R_{\text{reh}}|D) \ln \left[ \frac{P(\ln R_{\text{reh}}|D)}{\pi(\ln R_{\text{reh}})} \right] \, d\ln R_{\text{reh}}, \quad (4)$$

which precisely is a measure of the amount of information provided by the data $D$ about $\ln R_{\text{reh}}$ [24, 25]. This quantity is also the discrepancy measure between the posterior $P$ and the prior $\pi$ when the prior is viewed as an approximation of the posterior. Because the Kullback-Leibler divergence is invariant under any reparametrizations $x = f(\ln R_{\text{reh}})$ and uses a logarithmic score function as in the Shannon’s entropy, it is a well-behaved measure of information [26].

In the following, after having presented our data analysis method in Sec. II, we use Eq. (4) over the nearly 200 models of Encyclopaedia Inflationaris to extract the amount of information gained on the reheating parameter using the Planck 2015 data [27] complemented by the BICEP2/KECK measurements of the B-modes polarization [7]. By using a base 2 logarithmic function instead of the natural logarithm in Eq. (4), the information gain unit is the “bit” and this is our convention in the rest of the paper. In Sec. III, we show that these two data sets combined give an information gain on reheating equals to $\langle D_{\text{KL}} \rangle = 0.83 \pm 0.13$. We have also performed a new analysis of the Planck 2013 data, as in Ref. [21] but in terms of the Kullback-Leibler entropy, to get $\langle D_{\text{KL},13} \rangle = 0.55 \pm 0.14$. As a result, Planck 2015 and BICEP2/KECK achieve an increase of 40% more information gain on reheating. The relevance of this number and its implication for the Bayesian optimal design of
future experiments are discussed in the conclusion.

II. METHOD

In order to perform the CMB data analysis of the hundreds of slow-roll inflationary models at hand, we have followed the method described in Ref. [28] and applied in Refs. [21, 29]. It consists first in the determination of a machine learned effective likelihood $\mathcal{L}_{\text{eff}}$, depending only on the slow-roll parameters $\{P_*, \epsilon_*\}$. The effective likelihood has been obtained by marginalization of the joint Planck 2015 and BICEP2/KECK likelihood over all the other parameters $\theta_{\text{mic}}$. These ones correspond to the instrumental, astrophysical and cosmological parameters. One has

$$
\mathcal{L}_{\text{eff}}(P_*, \epsilon_*) \equiv \int P(D|\theta_{\text{mic}}, P_*, \epsilon_*)\pi(\theta_{\text{mic}})d\theta_{\text{mic}}.
$$

Within a given slow-roll model of inflation $\mathcal{M}$, with theoretical parameters $\theta_{\text{inf}}$, the quantities $P_*$ and $\epsilon_*$ are explicit (even if potentially complicated) functions of $\theta_{\text{inf}}$ and, most importantly, of $\ln R_{\text{reh}}$. As a result, from Bayes’ theorem, the posterior on $\ln R_{\text{reh}}$ is given by [28]

$$
P(\ln R_{\text{reh}}|D) = \frac{\pi(\ln R_{\text{reh}})}{P(D|\mathcal{M})} \times \int \mathcal{L}_{\text{eff}}[P_*(\theta_{\text{inf}}, \ln R_{\text{reh}}), \epsilon_*(\theta_{\text{inf}}, \ln R_{\text{reh}})]\pi(\theta_{\text{inf}})d\theta_{\text{inf}},
$$

where $P(D|\mathcal{M})$ is the global likelihood, which is proportional to the Bayesian evidence $P(\mathcal{M}|D) = P(D|\mathcal{M})\pi(\mathcal{M})$ of the model $\mathcal{M}$ to explain the data $D$.

In practice, we have used a modified version of the CAMB and COSMOMC codes implementing the slow-roll primordial power spectra at second order to perform a Markov-Chain-Monte-Carlo (MCMC) analysis of the Planck 2015 and BICEP2/KECK data [30, 31]. More precisely, we have used the public {\tt plik} likelihood provided by the Planck Collaboration, including the polarization cross spectra $TE$ and $EE$ at large multipoles, together with the BICEP2/KECK likelihood on $B$-modes based on the 217 and 353 GHz maps [6]. The parameter space associated with $\theta_{\text{mic}}$ contains the 4 $\Lambda$CDM parameters for a flat Universe: $\Omega_b h^2$, the density parameter of baryons (times the squared reduced Hubble parameter $h$), $\Omega_c h^2$, of cold dark matter, $\theta_{\text{MC}}$ related to the angular size of the sound horizon at last scattering and $\tau$, the reionization optical depth. There are 28 additional parameters associated with astrophysical signals and instrumental nuisances for both Planck 2015 and BICEP2/KECK. In total $\theta_{\text{mic}}$ belongs to a 32-dimensional parameter space and reads

$$
\theta_{\text{mic}} = \{\Omega_b h^2, \Omega_c h^2, 100\theta_{\text{MC}}, \tau, y_{\text{kal}}, a_{B, \text{dust}}, b_{B, \text{dust}}, A_{\text{CIB}}^{217}, \xi_{\text{SZ}, \text{CIB}}, A_{\text{tSZ}}^{143},
$$

$$
A_{\text{PS}}^{100, A^{143}, A^{143\times217}, A^{217}}, A_{\text{PS}}^{k_{\text{SZ}}}, A_{\text{dust TT}}, A_{\text{dust EE}}, A_{\text{dust EE}, EE}^{100, A^{143}}, A_{\text{dust EE}}^{100, A^{143}}, A_{\text{dust EE}}^{100, A^{143}, A^{143\times217}}, A_{\text{dust EE}}^{100, A^{143}}, A_{\text{dust EE}}^{100, A^{143}, A^{143\times217}}, A_{\text{dust EE}}^{100, A^{143}}, A_{\text{dust EE}}^{100, A^{143}, A^{143\times217}}, C_{100}^{A^{217}}, C_{22}^{A^{217}}\}.
$$

The parameters labeled by “$A$” refer to amplitude measurements at various frequencies, 100, 143 and 217 GHz for the temperature $T$ and polarization channels $E$, $B$, and all their relevant cross-correlations. The astrophysical signals are associated with unresolved point sources (PS), cosmic infrared background (CIB), dust emission (dust) and kinetic (kSZ) and thermal (tSZ) Sunyaev-Zeldowitch’s effects. The parameter $b_{B, \text{dust}}$ refers to the spectral index of galactic dust emission in the $B$-mode polarization channel and is required to correctly analyse the BICEP2/KECK data. The other parameters encode calibration uncertainties. More details on the meaning of these parameters can be found in Ref. [27].

For our purpose, the prior distributions on the $\theta_{\text{mic}}$ parameters have been chosen as specified in the Planck Collaboration’s paper [9] while the priors on the slow-roll parameters are as in Ref. [29]. Our MCMC exploration contains more than $10^6$ samples which allows us to evaluate $\mathcal{L}_{\text{eff}}$ by marginalization over $\theta_{\text{mic}}$. The machine learning algorithm used to fit $\mathcal{L}_{\text{eff}}(P_*, \epsilon_*)$ is a modified quadratic Shepard’s method [32] as described in Ref. [28]. The two-dimensional marginalized posteriors in the slow-roll parameter space $(P_*, \epsilon_1, \epsilon_2, \epsilon_3)$ are represented in Fig. 1.

For each models $\mathcal{M}_i$ of the Encyclopaedia Inflationaris collection, we have obtained the posterior $P(\ln R_{\text{reh}}|D)$
III. RESULTS

Let us now describe our main results. In Fig. 2, for explanatory purposes, we have represented the posteriors of ln $R_{\text{reh}}$ for two (favored) models, named loop inflation (LI) and supergravity brane inflation (SBI), both from the Planck 2013 data and from the Planck 2015 data with BICEP2/KECK. This figure illustrates the gain of information between these two data sets as well as the overall constraining power of CMB data on reheating. Vanishing values of ln $R_{\text{reh}}$ correspond to a radiation-like, or instantaneous, reheating scenario such that the current data are actually ruling out such a scenario for LI but favoring it for SBI [22]. Of course, for other models $M_i$, the posteriors on ln $R_{\text{reh}}$ are different and may be peaked over large or small values, or not constrained at all. But for all of them, $D_{\text{KL}}$ can be calculated.

In Fig. 3, we have represented by a circle each model $M_i$ in the plane ($B, D_{\text{KL}}$) where $B$ is the Bayes’ factor normalized to the best model. For a model $M_i$, assuming noncommittal priors $\pi(M_i) = \pi(M_j)$, it is obtained from the global likelihoods by

$$B_i = \frac{P(M_i | D)}{\sup_j [P(M_j | D)]} = \frac{P(D | M_i)}{\sup_j [P(D | M_j)]}. \quad (8)$$

This figure shows that most of the models having large Bayes factors are concentrated around values $D_{\text{KL}} \lesssim 1$ whereas disfavored models may have $D_{\text{KL}} > 2.5$. Such a correlation between information gain and Bayes factors for disfavored models is not surprising. If a model genuinely provides a bad fit to the data, the posterior of its free parameters, including the reheating parameter, can be pushed to the boundaries of their prior to be as good as it gets in improving the fit. As a result, the model parameter space may end up being very constrained while the model does not fit the data well compared to others. Figure 3 also shows some strongly disfavored models with $D_{\text{KL}} = 0$. These models are so far from the favored region that even changing the reheating history does not help to improve the fit to the data. For these reasons, a fair and robust estimation of the information gain on reheating is given by the average value of $D_{\text{KL}}$ in the space of all models

$$\langle D_{\text{KL}} \rangle = \sum_i P(M_i | D) D_{\text{KL}}(M_i) \simeq 0.82. \quad (9)$$

As expected, it is weighted by the Bayesian evidence, namely the probability of a model to explain the data: Disfavored models weigh less than favored models. Similarly, we find for the standard deviation

$$\sqrt{\langle D_{\text{KL}}^2 \rangle - \langle D_{\text{KL}} \rangle^2} \simeq 0.26. \quad (10)$$

As mentioned in the introduction, we have performed the same analysis for the Planck 2013 data and one gets $\langle D_{\text{KL,13}} \rangle = 0.55 \pm 0.14$. Therefore, Planck 2015 and BICEP2/KECK provide a 40% improvement in information gain compared to Planck 2013.
IV. CONCLUSION

Because $\langle D_{KL} \rangle$ quoted in Eq. (9) has been derived over a significant number of models, it should be representative of the whole information content of the current CMB data about the reheating epoch within inflation. It is almost 1 bit. Although this is a very modest number, 1 bit is the amount of information contained in answering “yes” or “no” to a given question. As illustrated in Fig. 2, the question is about the values of $ln R_{reh}$, and the current CMB data answer, on average, whether $ln R_{reh}$ is large or small. According to Shannon [36], 1 bit is also the typical amount of information carried by one letter within the English language. As such, one might argue that if the reheating scenario could be spelled, the Planck 2015 and BICEP2/KECK data would allow us to know one letter.

Within the deployment of the new generation of ground-based CMB polarization telescopes, the soon to be operational Euclid satellite [37] and the much needed next generation of CMB satellites [38, 39], one can only expect $\langle D_{KL} \rangle$ to dramatically increase in the future. Measuring $R_{reh}$ from the CMB would be a direct window onto the microphysics after inflation, i.e., at energy scales which could be as large as the grand unified theory energy scale. Moreover, it would allow us to disambiguate two inflationary models having exactly the same potential but not the same reheating history, as it is the case for Starobinsky inflation and Higgs inflation [40–42], or for some curvaton scenarios [43, 44].

Another application of our result concerns the Bayesian optimal design of future CMB measurements in which $D_{KL}$ could be used as a figure or merit. In this situation, the best experimental setup is the one maximizing the information gain $D_{KL}$. Because, for the reheating parameter, $D_{KL}$ can be improved by increasing sensitivity in all the slow-roll parameters, this suggests that it is certainly important to design future experiments to be as sensitive in the scalar spectral index and scalar running as in the tensor-to-scalar ratio.

Acknowledgments

V.V.’s work is supported by STFC Grants No. ST/K00090X/1 and No. ST/L005573/1. C.R.’s work is supported by the Belgian Federal Office for Science, Technical and Cultural Affairs.
[34] F. Feroz, M. P. Hobson, and M. Bridges, Mon. Not. R. Astron. Soc. 398, 1601 (2009), 0809.3437.
[35] F. Feroz, M. Hobson, E. Cameron, and A. Pettitt (2013), 1306.2144.
[36] C. E. Shannon, Bell System Technical Journal 30, 50 (1951), ISSN 1538-7305.
[37] R. Scaramella et al. (Euclid), IAU Symp. 306, 375 (2015), 1501.04908.
[38] T. Matsumura et al. (2013), [J. Low. Temp. Phys.176,733(2014)], 1311.2847.
[39] CORe, A satellite mission for probing cosmic origins, neutrinos masses and the origin of stars and magnetic fields, URL http://www.core-mission.org/science.php.
[40] J. Garcia-Bellido, D. G. Figueroa, and J. Rubio, Phys. Rev. D79, 063531 (2009), 0812.4624.
[41] T. Terada, Y. Watanabe, Y. Yamada, and J. Yokoyama, JHEP 02, 105 (2015), 1411.6746.
[42] D. G. Figueroa, J. Garcia-Bellido, and F. Torrenti, Phys. Rev. D92, 083511 (2015), 1504.04600.
[43] V. Vennin, K. Koyama, and D. Wands, JCAP 1511, 008 (2015), 1507.07575.
[44] V. Vennin, K. Koyama, and D. Wands, JCAP 1603, 024 (2016), 1512.03403.