INFLATION AFTER PLANCK: AND THE WINNERS ARE ...

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

We review the constraints that the recently released Cosmic Microwave Background (CMB) Planck data put on inflation and we argue that single field slow-roll inflationary scenarios (with minimal kinetic term) are favored. Then, within this class of models, by means of Bayesian inference, we show how one can rank the scenarios according to their performances, leading to the identification of “the best models of inflation”.

1 Introduction

The theory of inflation\textsuperscript{1,2,3,4,5} is currently the leading paradigm to describe the very early universe. The basic idea is quite simple: the problems of the pre-inflationary standard cosmological model are avoided if one postulates that a phase of accelerated expansion took place, at high energy, before the hot Big Bang era. If gravity is described by general relativity, then a negative pressure in the effective stress energy tensor sourcing the Einstein equations is all we need to produce this acceleration. Moreover, since, in the situation described above, field theory is the correct framework to describe matter and since a preferred direction (i.e. a spin or a vector) is not compatible with homogeneity and isotropy, a scalar field, the so-called inflaton field, appears to be the ideal candidate. Indeed, in that case, the pressure is given by the difference between the kinetic and the potential energy. Therefore, if the potential energy dominates over the kinetic energy, that is to say if the field slowly rolls down its potential, then one automatically produces a phase of inflation.

Inflation also naturally leads to a convincing mechanism for structure formation\textsuperscript{3,5} and this is probably the reason why this scenario is considered as very attractive. In brief, the quantum fluctuations of the coupled inflaton and gravitational fields are amplified and give rise to an almost scale invariant power spectrum in full agreement with the astrophysical observations. An attractive feature of this mechanism is that it is quite conservative: it is nothing but particle creation (the quantized cosmological perturbations) under the influence of a classical source (the background gravitational field). This is well-known in quantum field theory and is in fact the essence of the so-called Schwinger effect\textsuperscript{6}.

Although pretty straightforward regarding the physical principles, inflation turns out to be more complicated when it comes to concrete implementation. Indeed, there are literally hundreds of different models of inflation depending on whether there is one or several fields, with minimal or non-minimal kinetic terms, and/or with a featureless or not potential. In addition, all the possible combinations (for instance, several scalar fields with non-minimal kinetic terms) are also possible. How, then, can we identify to which version of inflation we are dealing with?

A priori, one could solely use theoretical considerations based on high energy physics to single out a unique consistent model. This seems to be unrealistic today since, at the energy
scales relevant for inflation, particle physics is not known and remains speculative. Moreover, the fact that we deal with so many models precisely originates from the fact that many possible versions of BSM (Beyond the Standard Model) physics exist leading to a plethora of different inflationary scenarios. For instance, models with a Dirac Born Infeld (DBI) kinetic term have been considered because this specific case can be motivated by string theory.

This leaves us with a “mixed approach” which consists, from the currently available scenarios and from the data, in inferring the correct model of inflation. In these proceedings, we explore this route and discuss the consequences for inflation of the recently released Planck data\(^{8,7}\). These data tell us that we live in a spatially flat universe, 100\(\Omega_K = -0.05^{+0.65\pm 0.66}\), which is of course very consistent with inflation and that the cosmological fluctuations are adiabatic (at 95\% CL) and Gaussian \(f_{\text{NL}}^{\text{loc}} = 2.7 \pm 5.8\), \(f_{\text{NL}}^{\text{eq}} = -42 \pm 75\) and \(f_{\text{NL}}^{\text{ortho}} = -25 \pm 39\). Another important message of the Planck data\(^8\) is the fact that a tilt in the power spectrum has now been detected at a significant statistical level, \(n_s = 0.9603 \pm 0.0073\), thus ruling out scale invariance at more than 5\(\sigma\). In addition, neither a significant running nor a significant running of the running have been detected since it was found that \(dn_s/d\ln k = -0.0134 \pm 0.009\) (Planck+WP) and \(d^2n_s/d\ln^2 k = 0.02 \pm 0.016\) (WMAP+WP), with a pivot scale chosen at \(k_* = 0.05\text{Mpc}^{-1}\).

Based on the above discussion, it is clear that single field slow roll models (with a minimal kinetic term) are favored from an observational point of view since this class of models precisely predicts no entropy perturbations and negligible non-Gaussianities. Of course, this does not mean that other inflationary scenarios are ruled out but simply that there are not needed to explain the data. Inflation therefore appears as a simple and non trivial, but non exotic, theory.

It should however be clear that, even if we restrict our considerations to this simple class of models, it still remains a very large number of possible models\(^9\). Then comes the questions of how one can constrain these models, estimate their performances and rank them, in a statistically well-defined fashion in order to find “the best model(s) of inflation”. Once a well justified method has been designed, it can be applied to all inflationary models in order to actually identify which scenario is favored by the Planck data. Answering and discussing these questions is the main subject of the present paper.

This article is organized as follows. In the next section, Sec. 2, we briefly review slow-roll inflation. Then, in Sec. 3, we define and discuss what is meant by a model A is better than a model B. For this purpose, we review the Bayesian model comparison approach, we quickly recall how the Bayesian evidence of a slow-roll inflationary model can be estimated and we present the results of Ref.\(^{10}\) which give the model winners. Finally, in the conclusion, Sec. 4, we summarize our results.

## 2 Slow-Roll Inflation and CMB Measurements

Slow-roll inflation is a very simple system. It consists in one scalar field with a minimal kinetic term and a potential \(V(\phi)\) and its behavior is controlled by the Friedmann-Lemaitre and Klein-Gordon equations, namely

\[
H^2 = \frac{1}{3M_{\text{Pl}}^2} \left[ \ddot{\phi}^2 + V(\phi) \right], \quad \ddot{\phi} + 3H\dot{\phi} + V_{\phi} = 0, \tag{1}
\]

where \(H \equiv \dot{a}/a\) denotes the Hubble parameter, \(a(t)\) being the Friedmann-Lemaitre-Robertson Walker (FLRW) scale factor and \(\dot{a}\) its derivative with respect to cosmic time \(t\). \(M_{\text{Pl}} = 8\pi G\) denotes the reduced Planck mass. A subscript \(\phi\) means a derivative with respect to the inflaton field. Therefore, the only unknown function is the potential and, here, we try to constrain its shape using the Planck data.

When the potential is no longer flat enough (this usually happens when the system approaches its ground state, i.e. the minimum of the potential), inflation stops, the inflaton field decays\(^{11,12}\), the decay products thermalize\(^{13}\) and this is how inflation is smoothly connected
to the standard hot Big Bang phase. Let $\rho$ and $P$ be the energy density and pressure of the effective fluid dominating the Universe during reheating and $\rho_{\text{reh}} \equiv P/\rho$ the corresponding “instantaneous” equation of state. One can also define the mean equation of state parameter, $\overline{w}_{\text{reh}}$, by

$$\overline{w}_{\text{reh}} \equiv \frac{1}{\Delta N} \int_{N_{\text{end}}}^{N_{\text{reh}}} w_{\text{reh}}(n) dn,$$

where $\Delta N \equiv N_{\text{reh}} - N_{\text{end}}$ is the total number of e-folds during reheating, $N_{\text{end}}$ being the number of e-folds at the end of inflation and $N_{\text{reh}}$ being the number of e-folds at which reheating is completed and the radiation dominated era begins. Then, one introduces a new parameter $\overline{v}$

$$\ln R_{\text{rad}} \equiv \frac{\Delta N}{4} \left( -1 + 3 \overline{w}_{\text{reh}} \right).$$

As discussed in detail in Ref. [14], this parameter completely characterizes the reheating phase and its knowledge is necessary in order to work out the inflationary predictions for the CMB. In particular, it can be related to the so-called reheating temperature through

$$T_{\text{reh}} = \frac{30 \rho_{\text{end}}}{\pi^2 g_\ast} \frac{1}{f_{\text{rad}}^{12(1+\overline{w}_{\text{reh}})/(1-3\overline{w}_{\text{reh}})}},$$

where $\rho_{\text{end}}$ is the energy density at the end of inflation, which is known when $V(\phi)$ has been chosen, and $g_\ast$ is the number of degrees of freedom at that time.

Let us now turn to the description of inflationary perturbations. Two types of fluctuations are relevant for inflation: density perturbations and primordial gravity waves. The density perturbations are described in terms of the Mukhanov-Sasaki variable $v(\eta, x)$. In the Schrödinger approach, the quantum state of the system is described by a wavefunctional, $\Psi[v(\eta, x)]$, which can be factorized into mode components as

$$\Psi[v(\eta, x)] = \prod_k \Psi_k(v_k^R, v_k^I) = \prod_k \Psi_k^R(v_k^R) \Psi_k^I(v_k^I),$$

where $v_k^R$ denotes the real part of $v$ and $v_k^I$ its imaginary part. Each wavefunction obeys a Schrödinger equation with an Hamiltonian that can be deduced from a second order expansion of the action “gravity + inflaton field”. Then, one can show that the solution is explicitly time-dependent and given by a Gaussian ($\eta$ being the conformal time)

$$\Psi_{k}^{R,I}(\eta, v_k^{R,I}) = N_k(\eta) e^{-\Omega_k(\eta)(v_k^{R,I})^2}.$$

where the functions $N_k(\eta)$ and $\Omega_k(\eta)$ can be expressed as

$$|N_k| = \left( \frac{2 \Re \Omega_k}{\pi} \right)^{1/4}, \quad \Omega_k = -\frac{i f'_k}{2 f_k}.$$

The function $f_k$ obeys the equation of motion of a parametric oscillator, namely $f''_k + \omega^2 f_k = 0$, where the time dependent frequency of this oscillator is given by $\omega^2(\eta, k) = k^2 - (a \sqrt{\epsilon_1})'' / (a \sqrt{\epsilon_1})$, $k$ being the wavenumber of the mode under consideration and $\epsilon_1 \equiv -H^2$ the first slow-roll parameter characterizing the cosmological expansion during inflation. For gravitational waves, one also obtains a Gaussian wave-function except that the fundamental frequency of the oscillator $f_k$ is now given by $\omega^2 = k^2 - a''/a$.

One of the great advantage of inflation is that it is possible to choose well justified initial conditions. In brief, this is because, at the beginning of inflation, the physical wavelengths of Fourier modes of cosmological relevance today are much smaller than the Hubble radius. These modes do not feel spacetime expansion and, as a consequence, it is natural to choose the vacuum
state as their initial state. Technically, this amounts to take \( \Omega_k = k/2 \) at initial time in Eq. (7) which indeed corresponds to the ground state wavefunction of an harmonic oscillator.

We have just seen that the effective frequency of density perturbations depends on the first slow-roll parameter and its derivatives. For this reason, it is interesting to define a hierarchy of slow-roll parameters by means of the following formula\(^\text{16}\)

\[
\epsilon_{n+1} \equiv \frac{d \ln |\epsilon_n|}{d N}, \quad n \geq 0, \tag{8}
\]

where \( \epsilon_0 \equiv H_{\text{ini}}/H \). The slow-roll conditions refer to a situation where all the \( \epsilon_n \)'s satisfy \( \epsilon_n \ll 1 \). From this definition, we see that \( \omega(k, \eta) \) for density perturbations depends on \( \epsilon_1 \), \( \epsilon_2 \) and \( \epsilon_3 \) while, for gravity waves, it only depends on \( \epsilon_1 \). Notice that, since \( H(\phi) \) and \( V(\phi) \) are related through the Einstein equations, the parameters \( \epsilon_n \) can also be expressed in terms of the successive derivatives of the potential, namely

\[
\epsilon_1 \approx \frac{M_{\text{pl}}^2}{2} \left( \frac{V_\phi}{V} \right)^2, \tag{9}
\]

\[
\epsilon_2 \approx 2 M_{\text{pl}}^2 \left[ \left( \frac{V_\phi}{V} \right)^2 - \frac{V_{\phi\phi}}{V} \right], \tag{10}
\]

\[
\epsilon_2 \epsilon_3 \approx 2 M_{\text{pl}}^4 \left[ \frac{V_{\phi\phi\phi\phi}}{V^2} - \frac{3}{2} V_{\phi\phi} \left( \frac{V_\phi}{V} \right)^2 + 2 \left( \frac{V_\phi}{V} \right)^4 \right]. \tag{11}
\]

The slow-roll approximation also allows us to solve the equation that controls the evolution of the function \( f_k \) and, therefore, of the wavefunction. Since the initial conditions are also completely specified (see the above discussion), the function \( f_k \) and, hence, the wavefunction, is completely known. One can then calculate the two-point correlation function of the Mukhanov-Sasaki variable or, in Fourier space, of the power spectrum\(^a\). This involves a double expansion. The power spectrum is first expanded around a chosen pivot scale \( k_* \) such that

\[
\frac{\mathcal{P}(k)}{\mathcal{P}_0} = a_0 + a_1 \ln \left( \frac{k}{k_*} \right) + \frac{a_2}{2} \ln^2 \left( \frac{k}{k_*} \right) + \ldots, \tag{13}
\]

where \( \mathcal{P}_0 = H^2 / (8 \pi^2 \epsilon_1 M_{\text{pl}}^2) \) and, then, the coefficients \( a_n \) are expanded in terms of the slow-roll parameters. Concretely, for scalar perturbations, at second order in the slow-roll approximation, one obtains\(^a, 17\)

\[
a_0 = 1 - 2 (C + 1) \epsilon_{1*} - C \epsilon_{2*} + \left( 2 C^2 + 2 C + \frac{\pi^2}{2} - 5 \right) \epsilon_{1*}^2
\]

\[
+ \left( C^2 - C + \frac{7 \pi^2}{12} - 7 \right) \epsilon_{1*} \epsilon_{2*} + \left( \frac{1}{2} C^2 + \frac{\pi^2}{8} - 1 \right) \epsilon_{2*}^2
\]

\[
+ \left( - \frac{1}{2} C^2 + \frac{\pi^2}{24} \right) \epsilon_{2*} \epsilon_{3*} \], \tag{14}
\]

\[
a_1 = -2 \epsilon_{1*} - \epsilon_{2*} + 2 (2 C + 1) \epsilon_{1*}^2 + (2 C - 1) \epsilon_{1*} \epsilon_{2*} + C \epsilon_{2*}^2 - C \epsilon_{2*} \epsilon_{3*}, \tag{15}
\]

\[
a_2 = 4 \epsilon_{1*}^2 + 2 \epsilon_{1*} \epsilon_{2*} + \epsilon_{2*}^2 - \epsilon_{2*} \epsilon_{3*}, \tag{16}
\]

where \( C \equiv \gamma_E + \ln 2 - 2 \approx -0.7296 \), \( \gamma_E \) being the Euler constant. \( \epsilon_{n*} \) denotes the value of the function \( \epsilon_n \) at Hubble radius crossing during inflation. For gravitational waves, the power spectrum has the same structure but the expressions of the coefficients \( a_n \) differ.

\( ^a \) For density perturbations, the definition of the power spectrum reads

\[
\mathcal{P}_\zeta(k) \equiv \frac{k^3}{4 \pi^2 M_{\text{pl}}^2} \left| \frac{v_k}{a H} \right|^2. \tag{12}
\]
In order to make concrete predictions, we must calculate the numerical values of the quantities $\epsilon_{n_\star}$. In order to do so, one needs to know the slow-roll trajectory and we need to calculate accurately when inflation stops. As a result, $\epsilon_{n_\star}$ usually depends on $\theta_{\text{inf}}$, the parameters of the potential $V(\phi)$, and on the reheating temperature: $\epsilon_{n_\star} = \epsilon_{n_\star}(\theta_{\text{inf}}, T_{\text{reh}})$.

The above considerations explain how the CMB can tell us something about inflation. Indeed, CMB measurements constrain the power spectrum, that is the say, given the form of $P(k)$ above, the values of the parameters $\epsilon_{n_\star}(\theta_{\text{inf}}, T_{\text{reh}})$. These parameters carry information about the shape of the potential (recall the expression of the slow-roll parameters in terms of the derivative of the potential) and on the reheating temperature. As a consequence, one can infer what are the properties of the inflaton potential $V(\phi)$ and learn about the physical conditions that prevailed in the early universe.

3 Ranking the Inflationary Models

3.1 Bayesian Analysis in Brief

In the previous section, we have described how one can calculate the predictions of a given inflationary model. However, we also would like to compare the performances of the different inflationary scenarios and one way to achieve this program is to compare the quality of the fits provided by the different models.

Let us now briefly describe how this can be achieved\textsuperscript{18,19,20}. Let us call $\mathcal{M}_1$ and $\mathcal{M}_2$ two competing models, aiming at explaining some data $D$ (here, of course, we have in mind the Cosmic Microwave Background - CMB - measurements), the model one depending on one parameter, $\theta$, and the model two depending on two parameters, $\alpha$ and $\beta$. Their likelihood function can be written as

$$
L_1(D|\theta) = L_{1,\max}e^{-\chi^2(\theta)/2}, \quad L_2(D|\alpha, \beta) = L_{2,\max}e^{-\chi^2(\alpha, \beta)/2},
$$

(17)

where $\chi^2$ is the effective chi-squared of the corresponding model that we do not need to specify at this stage. The quality of the fits can be estimated by computing the ratio of the maximums of the two likelihoods. However, this does not give us information regarding the complexity of the two models\textsuperscript{6}. If, for instance, model $\mathcal{M}_2$ achieves a very good fit only at the price of a fine-tuning, while $\mathcal{M}_1$ “naturally” performs well, one may wish to penalize $\mathcal{M}_2$ for its complexity. This “Occam’s razor” criterion is automatically included if one characterizes a model by its Bayesian evidence\textsuperscript{19}. The Bayesian evidence is the integral of the likelihood function over the prior space. Concretely, for $\mathcal{M}_1$ and $\mathcal{M}_2$, this leads to

$$
E_1 = \int L_1(D|\theta)\pi(\theta)d\theta, \quad E_2 = \int L_2(D|\alpha, \beta)\pi(\alpha, \beta)d\alpha d\beta.
$$

(18)

The prior distributions $\pi(\theta)$ and $\pi(\alpha, \beta)$, satisfying $\int \pi(\theta)d\theta = 1$ [and a similar expression for $\pi(\alpha, \beta)$], encodes what we know about the parameter $\theta$ before our information is updated when we learn about the data $D$. Let us notice that the likelihood functions are not normalized in the sense that $\int L_1(D|\theta)d\theta \neq 1$. For simplicity, let us now assume that the prior $\pi(\theta)$ is flat in the range $[\theta_{\text{min}}, \theta_{\text{max}}]$ and vanishes elsewhere. Because the distribution is normalized, one has $\pi(\theta) = 1/\Delta\theta$ with $\Delta\theta = \theta_{\text{max}} - \theta_{\text{min}}$. Let us also assume that the likelihood function has a bell shape (for instance, but necessarily, is a Gaussian function) characterized by the width $\delta\theta$. Let us finally suppose that the data give more information than the prior, in other words that the likelihood is more peaked than the prior. In that case, the Bayesian evidence of model $\mathcal{M}_1$ can be approximated by

$$
E_1 \simeq L_{1,\max}\frac{\delta\theta}{\Delta\theta}.
$$

(19)

\textsuperscript{6}In the following, we will introduce a quantity called the “Bayesian complexity”. Here, we use the word “complexity” in the standard sense, i.e. a model is more complicated than another if, for instance, it has more parameters or more fine-tuning. At this stage, it should not be confused with the Bayesian complexity.
In the same fashion, with the same assumptions (and obvious notations), the evidence of model \( M_2 \) can be expressed as

\[
\mathcal{E}_2 \simeq \mathcal{L}_{2,\text{max}} \frac{\delta \alpha}{\Delta \alpha} \frac{\delta \beta}{\Delta \beta}.
\]

Then, applying Bayes’ theorem, the probability of model \( M_1 \) is given by \( p(M_1|D) = \mathcal{E}_1 \pi(M_1)/p(D) \) and a similar formula for \( p(M_2|D) \). In this expression, \( \pi(M_1) \) represents the prior of model \( M_1 \) and the quantity \( p(D) \) is a normalization factor. If we say that, initially, the two models are equally probable, that is to say \( \pi(M_1) = \pi(M_2) \), then the ratio of their posterior probabilities, the so-called Bayes factor, can be expressed as

\[
B_{21} = \frac{p(M_2|D)}{p(M_1|D)} = \frac{\mathcal{E}_2}{\mathcal{E}_1} \frac{\mathcal{L}_{2,\text{max}}}{\mathcal{L}_{1,\text{max}}} \frac{\delta \alpha}{\Delta \alpha} \frac{\delta \beta}{\Delta \beta} \frac{\Delta \theta}{\delta \theta}.
\]

We see that the Bayes factor is controlled by the ratio \( \mathcal{L}_{2,\text{max}}/\mathcal{L}_{1,\text{max}} \) but now weighted by a factor, the so-called Occam factor, which penalizes the more complicated model, \( M_2 \), for any wasted parameter space. If, for instance, we take \( \delta \alpha/\Delta \alpha = \delta \beta/\Delta \beta = \delta \theta/\Delta \theta = 0.01 \), then \( B_{21} = 0.01 \mathcal{L}_{2,\text{max}}/\mathcal{L}_{1,\text{max}} \) and the more complicated model can win only if its likelihood at the “best fit point” is two orders of magnitude larger than that of \( M_1 \). So the best model is the model which can achieve the best compromise between simplicity and quality of the fit.

From the previous considerations, we see that the Bayesian evidence is an ideal tool to rank models and to find the best model. Nevertheless, it has the following property that could be considered as shortcomings. Suppose we define a model \( M_3 \) such that it is in fact model \( M_2 \) but with a third parameter, say \( \gamma \), such that this new parameter does not affect in any way the fit to the data; in other words, such that the likelihood is flat along \( \gamma \). In that case, the evidence of model \( M_3 \) is given by

\[
\mathcal{E}_3 = \int \mathcal{L}_3(D|\alpha, \beta, \gamma) \pi(\alpha) \pi(\beta) \pi(\gamma) d\alpha d\beta d\gamma = \int \mathcal{L}_3(D|\alpha, \beta) \pi(\alpha) \pi(\beta) d\alpha d\beta d\gamma = \mathcal{E}_2.
\]

Therefore, the two models have the same evidence despite the fact that \( M_2 \) is obviously simpler than \( M_3 \). In order to break this degeneracy, one has to introduce another quantity, the Bayesian complexity\(^\text{18}\), which allows us to distinguish \( M_2 \) and \( M_3 \).

In order to discuss the definition of the complexity, we work with a one parameter model only, i.e. \( M_1 \), (the generalization to an arbitrary number of parameters is straightforward) and we explicitly assume that the likelihood of the model is a Gaussian, namely

\[
\mathcal{L}_1(D|\theta) = \mathcal{L}_{1,\text{max}} e^{-(\theta - d)^2/(2\sigma^2)},
\]

where \( d \) represents a measurement of the parameter \( \theta \). Regarding the prior, instead of considering a flat distribution as before, we also assume it is given by a Gaussian centered at \( \theta = \mu \),

\[
\pi(\theta) = \frac{1}{\Sigma \sqrt{2\pi}} e^{-(\theta - \mu)^2/(2\Sigma^2)}.
\]

We can check that this distribution is properly normalized. These new assumptions are made for convenience only and do not change the above discussion (in fact, not quite exactly, see below). In particular, now, \( \delta \theta \) is clearly given by \( \sigma \) and the \( \Delta \theta \) by \( \Sigma \) so that the condition that the data are more informative than the prior, \( \delta \theta \ll \Delta \theta \), corresponds to \( \sigma \ll \Sigma \). Then one can calculate the posterior distribution of the parameter \( \theta \),

\[
p(\theta|D) = \frac{1}{\mathcal{E}_1} \mathcal{L}_1(D|\theta) \pi(\theta) = \frac{1}{\sqrt{2\pi}} \frac{1}{\Sigma^2 + 1/\sigma^2} \exp \left[ -\frac{1}{2} \left( \frac{1}{\Sigma^2} + \frac{1}{\sigma^2} \right) \left( \theta - \frac{d + \mu \sigma^2/\Sigma^2}{1 + \sigma^2/\Sigma^2} \right)^2 \right],
\]
which is a properly normalized Gaussian with mean and variance respectively given by
\[
\frac{d + \mu \sigma^2 / \Sigma^2}{1 + \sigma^2 / \Sigma^2}, \quad \frac{1}{\sqrt{1 + \Sigma^2 / \sigma^2}}.
\] (28)

On the other hand, the evidence of the model can be expressed as
\[
\mathcal{E}_1 = \frac{\mathcal{L}_{1,\text{max}}}{\sqrt{1 + \Sigma^2 / \sigma^2}} e^{-\langle \mu - d \rangle^2 / [2(\sigma^2 + \Sigma^2)]}.
\] (29)

This result is compatible with the previous discussion. Indeed, if the likelihood is more informative than the prior, then $\Sigma / \sigma \gg 1$ and the factor in front of the exponential reads $\sim \mathcal{L}_{1,\text{max}} \sigma / \Sigma$ which is equivalent to $\mathcal{L}_{1,\text{max}} \delta \theta / \Delta \theta$ and shows that the Occam’s factor is simply $\sigma / \Sigma$.

We now come to the definition of the Bayesian complexity denoted by $C_b$ in what follows. It reads
\[
C_b = \langle \chi^2 (\theta) \rangle - \chi^2 ((\theta)),
\] (30)
where the symbol $\langle \cdots \rangle$ means an average of the quantity $\cdots$ with a weight given by the posterior $p(\theta|D)$. In the above expression, the effective $\chi^2$ is defined by $-2 \ln \mathcal{L}$, which in the present case, reads
\[
\chi^2 (\theta) = \frac{1}{\sigma^2} (\theta - d)^2 - 2 \ln \mathcal{L}_{1,\text{max}}.
\] (31)

Then, using the explicit expression for the posterior distribution, see Eq. (26), and the previous expression for the $\chi^2$, one obtains the following formula for the Bayesian complexity
\[
C_b = \int p(\theta|D) \chi^2 (\theta) d\theta - \chi^2 \left[ \int p(\theta|D) \theta d\theta \right] = \frac{1}{1 + \sigma^2 / \Sigma^2}.
\] (32)

Therefore, if $\sigma \ll \Sigma$, one has $C_b \simeq 1$. In other words, since the likelihood function is much more peaked than the prior, the parameter $\theta$ is well-measured and the complexity is one. If, on the contrary, $\sigma \gg \Sigma$, then $C_b \simeq 0$ and the data are not accurate enough to constraint $\theta$. In the multidimensional case (i.e. a model with $n$ parameters), one has $C_b = \sum_{i=1}^n \frac{1}{1 + \sigma_i^2 / \Sigma_i^2}$, and the complexity gives the number of parameters that have been measured with the data $D$ or, in other words, the number of eigendirections in which the likelihood is more informative than the prior.

Finally, to conclude this section, let us try to derive the complexity for another very simple one parameter model, similar to the example we treated at the beginning of this article. This will help us to understand the meaning of complexity in another context. We assume that the likelihood is flat, centered at $\theta = 0$ with a width given by $\delta \theta$ and a height $\mathcal{L}_{\text{max}}$. We also assume that the prior is flat in the range $[-\Delta \theta / 2, \Delta \theta / 2]$ and has height $1 / \Delta \theta$ (and is less informative than the likelihood). In that case, it is straightforward to estimate the evidence of the model which is $\mathcal{E} = \mathcal{L}_{\text{max}} \delta \theta / \Delta \theta$. On the other hand, the posterior on the parameter $\theta$ can be expressed as
\[
p(\theta|D) = \frac{\mathcal{L}_{\text{max}}}{\Delta \theta} = \frac{1}{\delta \theta}, \quad \text{for } -\frac{\delta \theta}{2} < \theta < \frac{\delta \theta}{2},
\] (33)
and vanishes otherwise. As a consequence, one finds that the complexity can be written as
\[
C_b = -2 \int_{-\delta \theta / 2}^{\delta \theta / 2} \frac{1}{\delta \theta} \ln \mathcal{L} \, d\theta + 2 \ln \mathcal{L}_{\text{max}} = 0.
\] (34)

We see that one can no longer interpret the complexity as we did before. The reason is that the model we have used is too far from a Gaussian model and the concept of complexity cannot be really defined in that case. This illustrates the limitation of this statistical tool which is efficient only if the underlying statistics is not too far from a Gaussian. This is a warning that should be kept in mind in the following.
| $\ln B_{i,\text{REF}}^i$ | Odds     | Strength of evidence   |
|----------------------|----------|-------------------------|
| $< 1.0$              | $< 3 : 1$| Inconclusive            |
| $1.0$                | $\sim 3 : 1$ | Weak evidence           |
| $2.5$                | $\sim 12 : 1$ | Moderate evidence       |
| $5.0$                | $\sim 150 : 1$ | Strong evidence         |

Table 1: Jeffreys scale for evaluating the strength of evidence when comparing two models, $M_i$, versus a reference model $M_{\text{REF}}$.

3.2 Inflationary Bayesian Inference

Following the above considerations, it should now be clear that one way to estimate the performances of inflationary models (in explaining the recently released Planck data) is to calculate their evidence and their complexity. Then, one can rank them in a statistically consistent way and find the best scenarios. The predictions of all single field scenarios have been worked out and compared to Planck data in Encyclopædia Inflationaris and the calculation of the evidences and complexity for those models was performed in Ref. using a method recently developed in Ref. From these results, one can determine the Bayes factor defined by

$$B_{i,\text{REF}}^i = \frac{\mathcal{E}(D|M_i)}{\mathcal{E}(D|M_{\text{REF}})},$$

where the reference model was taken to be the Starobinsky model. The “Jeffreys scale”, see Table 1, gives an empirical prescription for translating the values of $B_{i,\text{REF}}^i$ into strengths of belief. One can summarize our results as follows. Firstly, for convenience, one can change the reference point of the Bayes factor and estimate the quantity $B_{i,\text{BEST}}^i = \mathcal{E}(D|M_i)/\mathcal{E}(D|M_{\text{BEST}})$ (rather than $B_{i,\text{REF}}^i$ before) with non-committal model priors. Then, one uses the Jeffreys scale with $B_{i,\text{BEST}}^i$, instead of $B_{i,\text{REF}}^i$, and count the number of models in the “inconclusive”, “weak evidence”, “moderate evidence” and “strong evidence” zones. The models in the “inconclusive” category can be viewed as the best models. We have found that this is the case for 52 models for a total of 193 models, that is to say 26% of the models. Therefore, this means that $\sim 73\%$ of the inflationary scenarios can now be considered as disfavored and/or ruled out by the Planck data.

Secondly, one determines the number of unconstrained parameters, $N_{\text{uc}}^i$, which is the number of parameters of model $M_i$, $N_{\text{param}}^i$, minus its complexity $C_{i,\text{b}}$

$$N_{\text{uc}}^i = N_{\text{param}}^i - C_{i,\text{b}}.$$  

(36)

Then, among the models in the “inconclusive” region, one should prefers models for which $N_{\text{uc}}^i \simeq 0$. If one retains the criterion $0 < N_{\text{uc}}^i < 1$, then one reduces the number of “good models” to 17, that is to say to $\sim 9\%$ of the Encyclopædia Inflationaris scenarios.

These results are summarized in Fig. 1 which shows the histogram corresponding to the number of models in each Jeffreys category with a given value of $N_{\text{uc}}^i$. A complete analysis and the list of the best models can be found in Ref. 10.

4 Conclusions

In these proceedings, we have analyzed the implications of the recently released Planck data for inflation. We have argued that single field slow-roll scenarios with minimal kinetic term are favored by Planck 2013. Then, we have designed specific Bayesian tools to further constrain the models within the class of favored scenarios. We have shown that Planck2013 can then single out about $\sim 10\%$ of the models, thus strongly reducing the inflationary landscape compatible with the astrophysical observations. Our results demonstrate concretely that CMB data can
Figure 1 – Histogram representing the number of inflationary models after Planck2013 according to the Jeffrey category and the number of unconstrained parameters.

The next release of Planck measurements will allow us to learn even more about inflation.

Acknowledgments

I would like to thank C. Ringeval and V. Vennin for careful reading of the manuscript.

References

1. A. Starobinsky, *Phys. Lett.* B 91, 99 (1980).
2. A. Guth, *Phys. Rev.* D 23, 347 (1981).
3. V. Mukhanov and G. Chibisov, *JETP Lett.* 33, 532 (1981).
4. A. Linde, *Phys. Lett.* B 108, 389 (1982).
5. A. Starobinsky, *Phys. Lett.* B 117, 175 (1982).
6. J. Martin, *Lect. Notes Phys.* 738, 193 (2008).
7. P. Ade et al., arXiv:1303.5084
8. P. Ade et al., arXiv:1303.5076.
9. J. Martin, C. Ringeval, and V. Vennin, arXiv:1303.3787.
10. J. Martin, C. Ringeval, R. Trotta and V. Vennin, arXiv:1312.3529.
11. M. Turner, *Phys. Rev.* D 28, 1243 (1983).
12. L. Kofman, A. Linde and A. Starobinsky, *Phys. Rev.* D 56, 3258 (1997).
13. D. Podolsky, G. Felder, L. Kofman and M. Peloso *Phys. Rev.* D 73, 023501 (2006).
14. J. Martin and C. Ringeval, *Phys. Rev.* D 82, 023511 (2010).
15. J. Martin, V. Vennin and P. Peter, *Phys. Rev.* D 86, 103524 (2012).
16. D. Schwarz and C. Terrero-Escalante, *JCAP* 0408, 003 (2004).
17. J. Martin, C. Ringeval and V. Vennin, *JCAP* 1306, 021 (2013).
18. M. Kunz, R. Trotta and D. Parkinson, *Phys. Rev.* D 74, 023503 (2006).
19. R. Trotta, *Contemp. Phys.* **49**, 71 (2008).
20. J. Martin, C. Ringeval and R. Trotta, *Phys. Rev. D* **83**, 063524 (2011).
21. F. Feroz, K. Cranmer, M. Hobson, R. de Austri and R. Trotta, *JHEP* **1106**, 042 (2011).
22. C. Ringeval, arXiv:1312.2347.