Pure Natural Inflation

Yasunori Nomura,1,2,3 Taizan Watari,3 and Masahito Yamazaki3

1Berkeley Center for Theoretical Physics, Department of Physics, University of California, Berkeley, CA 94720, USA
2Theoretical Physics Group, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA
3Kavli Institute for the Physics and Mathematics of the Universe (WPI), University of Tokyo, Kashiwa, Chiba 277-8583, Japan

We point out that a simple inflationary model in which the axionic inflaton couples to a pure Yang-Mills theory may give the values of the scalar spectral index ($n_s$) and tensor-to-scalar ratio ($r$) in complete agreement with the current observational data.

Cosmic inflation plays an important role in explaining observable features of our universe, including its extreme flatness, as well as the origin of primordial curvature perturbations. The detailed predictions of inflation, however, depend on the potential $V(\phi)$ of the inflaton field $\phi$. An important issue, therefore, is to understand what is the correct model of inflation and how it emerges from the underlying physics.

Recent observations by Planck [1] and BICEP2/Keck Array [2] have started constraining simple models of inflation. In particular, arguably the simplest model of inflation $V(\phi) = m^2 \phi^2/2$ [3]—which gives the correct value for the scalar spectral index $n_s \approx 0.96$—is now excluded at about the $3\sigma$ level because of the non-observation of tensor modes. This raises the following questions. Does the model of inflation need to be significantly complicated? Is the agreement of $n_s$ with the data purely accidental?

In this letter, we argue that the answers to these questions may both be no. In particular, we argue that a simple inflationary model in which the inflaton $\phi$ couples to the gauge field of a pure Yang-Mills theory

$$\mathcal{L} = \frac{1}{32\pi^2} \int e^{\mu\nu\rho\sigma} \text{Tr} F_{\mu\nu} F_{\rho\sigma},$$

may give the values of $n_s$ and the tensor-to-scalar ratio, $r$, in perfect agreement with the current observational data. Here, $\phi$ is a pseudo-Nambu-Goldstone boson—axion—of a shift symmetry $\phi \rightarrow \phi + \text{const.}$, and $f$ is the axion decay constant. For now we assume that the gauge group of the Yang-Mills theory is $SU(N)$, but the model also works for other gauge groups; see later.

Conventionally, the potential of the axion field as in Eq. (1) is assumed to take the form generated by non-perturbative instantons

$$V(\phi) = \Lambda^4 \left[1 - \cos \left(\frac{\phi}{f}\right)\right],$$

where $\Lambda$ is the dynamical scale of the Yang-Mills theory. The resulting inflation model is called natural inflation [4, 5], which has been extensively studied in the literature. The potential of Eq. (2), however, is not favored by the current data, and it would soon be excluded at a higher confidence level if the bound on $r$ improves with $n_s$ staying at the current value; see Fig. 1.

It is known since long ago, however, that the cosine potential in Eq. (2) is not correct in general, as argued by Witten [6, 7] in the large $N$ limit [8] with the 't Hooft coupling $\lambda \equiv g^2 N$ held fixed. In particular, while the physics is periodic in $\phi$ with the period of $2\pi f$ (because $\theta \equiv \phi/f$ is the $\theta$ angle of the Yang-Mills theory), the multi-valued nature of the potential allows for the potential of $\phi$ in a single branch

$$V(\phi) = N^2 \Lambda^4 V \left(\frac{\lambda \phi}{8\pi^2 N f}\right),$$

not to respect the periodicity under $\phi \rightarrow \phi + 2\pi f$. Here,

1 See, e.g., Refs. [9–13] for related discussion in the context of inflation.
the combination
\[ x \equiv \frac{\lambda \phi}{8\pi^2 N_f}, \quad (4) \]
appearing in the argument of \( V(x) \) is determined by analyzing the large \( N \) limit. This allows for building axionic models of inflation in which the range of the field excursion exceeds the decay constant \( f \) \cite{14-16}.

The potential of Eq. (3) has an expansion of the form
\[ V(\phi) = \sum_{n=1}^{\infty} b_{2n} \left( \frac{\phi}{F} \right)^{2n}, \quad (5) \]
where \( F \propto f \). The values of the coefficients \( b_{2n} \)—more precisely their signs and double ratios—are important for how the predictions for \( n_s \) and \( r \) change as \( F \) is varied. If the cosine potential in Eq. (2) were valid, then we would obtain
\[ \text{sgn}(b_{2n}) = (-1)^{n-1}, \quad (6) \]
and
\[ \frac{b_4}{b_2} = \frac{2}{5}, \quad \frac{b_8}{b_4} = \frac{15}{28}, \quad \ldots, \quad (7) \]
which lead to the curves labeled as “cosine” in Fig. 1. The correct values of the double ratios, however, are expected to be different from these values. In fact, \( b_{2n} \)'s obtained by lattice gauge theory disfavor the cosine form of Eq. (2) and are rather consistent with those expected from large \( N \) expansion \cite{17}.

While \( b_{2n} \)'s may in principle be determined by lattice calculations, their errors are still large. Instead, we may infer the form of the potential by the following arguments. First, invariance under the \( CP \) transformation \( \phi \to -\phi \) implies that \( V(\phi) \) is a function of \( x^2 \), where we have absorbed a possible bare \( \theta \) parameter in the definition of \( \phi \). Second, \( V(\phi) \) is expected to be flatter than the potential energy approaches the point of the deconfining phase transition with increasing \( |\phi| \) (since the dynamics generating the potential will become weaker). Assuming that the potential is given by a simple power law, we thus expect \( V(x) \sim 1/(x^2)^p \) \((p > 0)\). This potential is singular at \( x \to 0 \), and a simple way to regulate it is to replace \( x^2 \) with \( x^2 + \text{const} \). After setting the minimum of the potential to be zero, these considerations give
\[ V(x) = M^4 \left[ 1 - \frac{1}{(1 + (x^2)^p)} \right] \quad (p > 0), \quad (8) \]
where \( M \sim \sqrt{N\Lambda} \), and \( c > 0 \) is a parameter of order unity. Here, we have used the well-established fact that the coefficient of \( x^2 \) is positive when \( V(x) \) is expanded around \( x = 0 \). We call the model of inflation in which the axionic inflaton potential is generated by a pure Yang-Mills theory (whose potential we expect to take the form of Eq. (8)) pure natural inflation.

As in the cosine potential, the potential of Eq. (8) gives \( \text{sgn}(b_{2n}) = (-1)^{n-1} \). It, however, gives different values of the double ratios
\[ \frac{b_{4n}}{b_{2n}} = \frac{2(p + 2)}{3(p + 1)}, \ldots, \quad \frac{b_{2n+4}}{b_{2n+2}} = \frac{(n + 1)(p + n + 1)}{(n + 2)(p + n)}, \ldots. \quad (9) \]
Therefore, predictions of this model are different from those of conventional natural inflation. (For example, by equating \( (b_6/b_4)/(b_4/b_2) \) we obtain \( p = -7/2 < 0 \).) Here, we have assumed that the effect of a transition between different branches can be neglected, which we will argue to be the case.

The potential of Eq. (8) can be obtained by a holographic calculation \cite{10, 18}, which is applicable in the limit of large \( N \) and \( 't \) Hooft coupling. In this calculation, \( N \) D4-branes in type IIA string theory are considered, with the D4-branes wrapping a circle. Below the Kaluza-Klein scale \( M_{\text{KK}} \) for the circle, the theory reduces to a 4d (non-supersymmetric) pure \( SU(N) \) Yang-Mills theory, with the dynamical scale
\[ \Lambda = M_{\text{KK}} e^{-\frac{2n^2}{3N}}, \quad (10) \]
where \( \lambda \) is the \( 't \) Hooft coupling at \( M_{\text{KK}} \). Considering the backreaction to the geometry of the constant Wilson line of the Ramond-Ramond one-form, which represents the \( \theta \) angle of the gauge theory, the potential of the form of Eq. (8) is obtained with \( c = 1 \) and \( p = 3 \). Specifically, the potential of \( \phi \) for a single branch is given by
\[ V(\phi) = M^4 \left[ 1 - \frac{1}{\left( 1 + (\frac{\phi}{F})^2 \right)^3} \right], \quad (11) \]
where\(^{2}\)
\[ M^4 = \frac{\lambda N^2}{3\pi^2 M_{\text{KK}}^4}, \quad F = \frac{8\pi^2 N}{\lambda} f. \quad (12) \]
The potentials for the other branches are obtained by replacing \( \phi \) with \( \phi + 2\pi k f \) \((k \in \mathbb{Z})\); see Fig. 2.

To illustrate the parameter region we consider, let us choose
\[ \lambda \simeq 8\pi^2. \quad (13) \]
Strictly speaking, the holographic calculation is not quite valid with this value of the \( 't \) Hooft coupling—it requires a larger value of \( \lambda \). However, we may expect, e.g. based on the success of the AdS/QCD program \cite{19, 20}, that
\[ \text{The } 't \text{ Hooft coupling (and the gauge coupling squared) defined in Ref. [10] is a factor of 2 smaller than } \lambda \text{} (g^2) \text{ here.} \]
this reasonably approximates the true dynamics of the 4d Yang-Mills theory. With this choice of $\lambda$, we find

$$M \approx \sqrt{N} \Lambda, \quad F \approx N f,$$

as one naively expects from dimensional and $N$-scaling considerations. From the analysis of Ref. [10], we expect that for sufficiently large $N$ the effect of a transition between different branches is not important, unless $\phi$ becomes much larger than $F$ (which does not occur in our analysis below). To reproduce the observed amplitude of the scalar perturbation, we will need to take

$$M \sim 10^{16} \text{ GeV}. \quad (14)$$

The precise value depends on other parameters, e.g., $F$.

The second expression in Eq. (14) implies that for $N \gg 1$ the characteristic scale for the field excursion, $F$, can be much larger than the axion decay constant $f$. This, however, does not mean that $F$ can be much larger than the Planck scale $M_{\text{Pl}} \approx 1.22 \times 10^{19} \text{ GeV}$. In general, the decay constant $f$ is expected to be smaller than the field theoretic cutoff (string) scale $M_s$,

$$f \lesssim M_s. \quad (16)$$

On the other hand, the Planck scale is related with $M_s$ as $M_{\text{Pl}}^2 \sim N^2 M_s^2$ (see, e.g., Ref. [21]), so that $N$ drops from the relation between $F$ and $M_{\text{Pl}}$:

$$F \lesssim O(M_{\text{Pl}}). \quad (17)$$

We argue that this is a desired feature. If $F \gg M_{\text{Pl}}$, the inflaton potential would be well approximated by the first, quadratic term in Eq. (5), which is excluded by the data. Because of Eq. (17), however, we expect that higher terms in Eq. (5) are important. This makes the predictions of the model deviate from those of the quadratic potential $V(\phi) = m^2 \phi^2/2$.

In Fig. 1, we plot the values of $n_s$ and $r$ predicted by the potential of Eq. (11). For $F \gg M_{\text{Pl}}$, the predictions approach those of the quadratic potential, denoted by the black dots for $e$-folding $N_e = 50$ (left) and 60 (right).

As $F$ decreases, however, they deviate from these values. In particular, the tensor-to-scalar ratio $r$ decreases while $n_s$ being (almost) kept, as represented by the lines indicated as “pure natural.” This is because higher terms in Eq. (5) start contributing. In the figure, we have varied $F$ from $100M_{\text{Pl}}$ to 0.1$M_{\text{Pl}}$, with the predictions for $F/M_{\text{Pl}} = 10, 5, 1$ indicated by the red dots (from top to bottom). We find that the model gives the values of $n_s$ and $r$ consistent with the data at the 95% (68%) CL for $F/M_{\text{Pl}} \lesssim 3.3 (0.7)$ for $N_e = 50$ and $F/M_{\text{Pl}} \lesssim 6.8 (4.4)$ for $N_e = 60$. Given Eq. (17), this is quite satisfactory.

In Fig. 3, we plot the predictions arising from the potential in which $p$ in Eq. (8) takes more general values $p = 1, 2, 3, 4, 10$. We have parameterized the potentials as

$$V(\phi) = M^4 \left[ 1 - \frac{1}{\left(1 + (\phi/F)^2\right)^p} \right], \quad (18)$$

and, as in Fig. 1, varied $F/M_{\text{Pl}}$ in the range between 100 and 0.1, with the dots representing $F/M_{\text{Pl}} = 10, 5, 1$ (from top to bottom). We find that the success of the model is robust for a wide range of $p \approx 1 - O(10)$: the predicted values of $n_s$ and $r$ agree well with the current data for $F/M_{\text{Pl}} \lesssim O(1)$. We thus conclude that pure natural inflation is consistent with the data even if the true potential does not take exactly the form of Eq. (11) as suggested by the holographic analysis.

So far, we have focused on the case that the gauge group of the Yang-Mills theory is $SU(N)$. However, our basic arguments, e.g. those around Eq. (8), do not depend on this specific choice. We thus expect that similar predictions also result for other gauge groups, with $N$ replaced by the dual Coxeter number of the group.

3 A special case of $V(\phi) \approx A - B/\phi^6$ leading to $n_s \simeq 0.965$ and $r \simeq 8 \times 10^{-4}$ was discussed in Ref. [10].
Finally, we mention that a large value of $N$—despite the fact that it does not help to make $F$ larger than $M_{Pl}$—can make the decay constant $f$ smaller than $F$; see Eq. (14). This allows for enhancing couplings of the inflaton $\phi$ to the standard model gauge fields for fixed $F$, i.e. for a fixed inflation potential. This in turn allows for raising reheating temperature $T_R$. The reheating temperature is given by $T_R \simeq 0.2 \sqrt{\Gamma \phi M_{Pl}}$, where

$$\Gamma \phi \approx \frac{nm^3_\phi}{4096\pi^5f^2}, \quad (19)$$

is the inflaton decay width. Here,

$$m_\phi = \sqrt{2}p M^2 \frac{F}{F/M_{Pl}} \approx 10^{13} \text{ GeV}, \quad (20)$$

is the inflaton mass, and $n$ is the final state multiplicity. ($n = 12$ if $\phi$ decays only to the standard model gauge fields.) This yields

$$T_R \sim 10^9 \text{ GeV} \left(\frac{N}{10}\right) \left(\frac{0.5}{F/M_{Pl}}\right)^{5/2}, \quad (21)$$

The model can, therefore, be made consistent with thermal leptogenesis, which requires $T_R \gtrsim 2 \times 10^9 \text{ GeV}$ [22], even for $F/M_{Pl} \approx O(1)$.

This work was supported in part by the WPI Research Center Initiative (MEXT, Japan). The work of Y.N. was also supported in part by the National Science Foundation under grants PHY-1521446, by MEXT KAKENHI Grant Number 15H05895, and by the Department of Energy (DOE), Office of Science, Office of High Energy Physics, under contract No. DE-AC02-05CH11231. The work of M.Y. was supported by JSPS KAKENHI (15K17634) and JSPS-NRF Joint Research Project.


