PHYSICAL REVIEW D **89,** 043527 (2014)

Remarks on the Starobinsky model of inflation and its descendants

Alexandros Kehagias,[1,2] Azadeh Moradinezhad Dizgah,[1] and Antonio Riotto[1]

[1]*Department of Theoretical Physics and Center for Astroparticle Physics (CAP), 24 quai E. Ansermet, CH-1211 Geneva 4, Switzerland*

[2]*Physics Division, National Technical University of Athens, 15780 Zografou Campus, Athens, Greece*

(Received 30 December 2013; published 28 February 2014)

We point out that the ability of some models of inflation, such as the Higgs inflation and the universal attractor models at large values of the coupling ξ, in reproducing the available data is due to their relation to the Starobinsky model of inflation. For large field values, where the inflationary phase takes place, all of these classes of models are indeed identical to the Starobinsky model. Nevertheless, the inflation is just an auxiliary field in the Jordan frame of the Starobinsky model, and this leads to two important consequences: first, the inflationary predictions of the Starobinsky model and its descendants are slightly different (albeit not measurably); second, the theories have different small-field behavior, leading to different ultraviolet cutoff scales. In particular, one interesting descendant of the Starobinsky model is the nonminimally coupled quadratic chaotic inflation. Although the standard quadratic chaotic inflation is ruled out by the recent Planck data, its nonminimally coupled version is in agreement with observational data and valid up to Planckian scales.

DOI: 10.1103/PhysRevD.89.043527 PACS numbers: 98.80.-k, 04.50.-h, 98.80.Cq

I. INTRODUCTION

The recent Planck results [1] have indicated that the cosmological perturbations in the cosmic microwave background (CMB) radiation are nearly Gaussian and of the adiabatic type. If one insists in assuming that these perturbations are to be ascribed to single-field models of inflation [2], the data put severe restriction on the inflationary parameters. In particular, the Planck results have strengthened the upper limits on the tensor-to-scalar ratio, $r \lesssim 0.12$ at 95% C.L., disfavoring many inflationary models [1]. For instance, the chaotic models with potential ϕ^n with $n \geq 2$ are not in good shape; in particular, the simplest quadratic chaotic model $m^2\phi^2$ has been excluded at about 95% C.L.

Among the inflationary models discussed by the Planck collaboration is the Starobinsky $(R + R^2)$ theory proposed in Ref. [3], whose predictions for the perturbations were originally discussed in Ref. [4]. Although this model looks quite ad hoc at the theoretical level, its perfect agreement with the Planck data is basically due to an additional $1/N$ suppression (N being the number of e-folds till the end of inflation) of r with respect to the prediction for the scalar spectral index n_s. As expected, this has renewed interest in this model. Particular recent efforts have been in the direction of the supersymmetric version of it [5–11], along the lines originated in Refs. [12,13].

Of course there are also other models which are in agreement with the Planck data. For example, the so-called Higgs inflation [14–16] and the so-called universal attractor models at large values of the coupling ξ [17,18] give exactly the same inflationary predictions to leading order as the Starobinsky theory. In this paper we stress that there is a simple reason why this apparent coincidence takes place: all of these models are the Starobinsky model during inflation. While this might be known to some (see for instance Ref. [19] for the Higgs model of inflation), it seems to be mysterious to others [20]. In the Planck paper [1], for instance, the Starobinsky and the Higgs inflation models are treated as different. The reason that these models may be considered descendants of the Starobinsky model is that during inflation the kinetic terms are subleading with respect to the potential terms, and therefore, they can be neglected in first approximation. If so, the scalar field present in the Higgs model and in the universal attractor models is just an auxiliary field which can be integrated out, giving rise to the Starobinsky model. During the inflationary phase, where kinetic energies are negligible, apparent unrelated models are described effectively by the same dynamics.

The next natural question is therefore if one can distinguish these descendants from the Starobinsky model. An obvious way is to compare the inflationary parameters in these models beyond the leading order. As we show, the slow-roll parameters are the same up to $\sim 10^{-5}$ corrections, which are quite small to be measured in the upcoming measurements. Another difference relies on the different way reheating after inflation proceeds in the different models [19], but again, differences are of the order of 10^{-3} in the spectral index, hardly detectable by Planck (the often-quoted Planck result $n_s = 0.960 \pm 0.007$ is based on assumptions on the reionization, the primordial Helium abundance and the effective number of neutrino).

The fact that the Starobinsky model and its descendants differ by the kinetic term is also interesting from another point of view. While the kinetic terms play a subleading role during inflation, they play a fundamental role in determining the UV behavior of the theories and its cutoff Λ. In particular, there is an ongoing discussion about the validity of the Higgs inflation as it seems that the cutoff of this theory is lower than the inflationary scale [21–23] (see Ref. [24] for a criticism of these results). On the other side, the cutoff of the Starobinsky theory is the Planck scale $M_{\rm p}$ [23] so that inflation can be trusted in this framework. The difference relies exactly in the role played by the kinetic energy. We extend the discussion of the cutoff for the universal attractor models. We find that when the potential in the Jordan frame is of the power-law type $\sim\phi^{2n}$, the cutoff is always above the inflationary scale only for $n > 7/2$. Therefore, for any value of $n < 7/2$ (like for example the Higgs inflation case for which $n = 2$), the cutoff satisfies the relation $\Lambda < V^{1/4}$, where V is the vacuum energy driving inflation, thus making the inflationary predictions questionable. The case $n = 1$ is particular as it corresponds to a nonminimally coupled simple quadratic chaotic inflation. We find in this case that the cutoff of this theory is at the Planck scale as in the Starobinsky theory. Therefore, inflation can be trusted for the nonminimally coupled version of the simple quadratic chaotic inflation.

The structure of this work is as follows. In Sec. II, we briefly describe the Starobinsky model and show why the Higgs inflation model, the universal attractor models as well as a higher dimensional Starobinsky-like model, which is related to the T model of Ref. [20], may be considered descendants of the Starobinsky model during inflation. In Sec. III, we discuss the differences between these models in their predictions for inflationary parameters, deferring the discussion of their cutoffs, if viewed as effective field theories, until Sec. IV. Finally, we conclude in Sec. V.

II. THE STAROBINSKY MODEL AND ITS DESCENDANTS

The Starobinsky model [3] is described by the Lagrangian

$$S_{\rm S} = \frac{1}{2}\int {\rm d}^4x\sqrt{-g}\left(M_{\rm p}^2 R + \frac{1}{6M^2}R^2\right). \tag{2.1}$$

This theory propagates a spin-2 state (graviton) and a scalar degree of freedom. The latter is manifest in the so-called linear representation where one can rewrite the Lagrangian (2.1) as [25]

$$S_{\rm S} = \int {\rm d}^4x\sqrt{-g}\left(\frac{M_{\rm p}^2}{2}R + \frac{1}{M}R\psi - 3\psi^2\right). \tag{2.2}$$

It is easy to see that upon integrating out ψ, one gets back the original theory (2.1). After writing the expression (2.2) in the Einstein frame by means of the conformal transformation

$$g_{\mu\nu} \to e^{-\sqrt{2/3}\phi/M_{\rm p}} g_{\mu\nu} = \left(1 + \frac{2\psi}{MM_{\rm p}^2}\right)^{-1} g_{\mu\nu}, \tag{2.3}$$

we get the equivalent scalar field version of the Starobinsky model

$$S_{\rm S} = \int {\rm d}^4x\sqrt{-g}\Bigg[\frac{M_{\rm p}^2}{2}R - \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{3}{4}M_{\rm p}^4 M^2\left(1 - e^{-\sqrt{\frac{2}{3}}\phi/M_{\rm p}}\right)^2\Bigg]. \tag{2.4}$$

We see that during inflation (large values of ϕ), the dynamics are dominated by the vacuum energy

$$V_{\rm S} = \frac{3}{4}M_{\rm p}^4 M^2. \tag{2.5}$$

Equation (2.4) is the linear representation of the Starobinsky model where the extra scalar degree of freedom is manifest. The theory described by the action given by Eq. (2.4) leads to inflation with scalar tilt and tensor-to-scalar ratio

$$n_s - 1 \approx -\frac{2}{N}, \qquad r \approx \frac{12}{N^2}. \tag{2.6}$$

Note that r has an additional $1/N$ suppression with respect to the scalar tilt and thus this theory predicts a tiny amount of gravitational waves. It is therefore consistent with the Planck constraints. The normalization of the CMB anisotropies fixes $M \approx 10^{-5}$.

A. Higgs inflation as a descendant of the Starobinsky model

Let us now consider the Higgs inflation model which is described by an action of the form [16]

$$S_{\rm HI} = \int {\rm d}^4x\sqrt{-g}\Bigg[\frac{M_{\rm p}^2}{2}R + \xi H^\dagger H R - \partial_\mu H^\dagger\partial^\mu H - \lambda(H^\dagger H - v^2)^2\Bigg], \tag{2.7}$$

where H is the standard model (SM) Higgs doublet and v its vacuum expectation value. In the unitary gauge $H = h/\sqrt{2}$ and for $h^2 \gg v^2$, the theory is described by

$$S_{\rm HI} = \int {\rm d}^4x\sqrt{-g}\left(\frac{M_{\rm p}^2}{2}R + \frac{1}{2}\xi h^2 R - \frac{1}{2}\partial_\mu h\partial^\mu h - \frac{\lambda}{4}h^4\right). \tag{2.8}$$

In this case, successful inflation exists for $\xi^2/\lambda \approx 10^{10}$. During inflation, the kinetic term is, by definition, smaller than any potential term, and thus (2.8) is effectively described by the action

$$S_{\rm HI} = \int d^4x\sqrt{-g}\left(\frac{M_{\rm p}^2}{2}R + \frac{1}{2}\xi h^2 R - \frac{\lambda}{4}h^4\right). \qquad (2.9)$$

The Higgs field during inflation has been turned into an auxiliary field which can be integrated out. We find that

$$\xi hR - \lambda h^3 = 0, \qquad (2.10)$$

which leads to

$$h^2 = \frac{\xi R}{\lambda}. \qquad (2.11)$$

Plugging this value back into the action, we find that the theory during inflation can be equally well described by

$$S_{\rm HI} = \int d^4x\sqrt{-g}\left(\frac{M_{\rm p}^2}{2}R + \frac{\xi^2}{4\lambda}R^2\right). \qquad (2.12)$$

Therefore, during inflation, Higgs inflation is equivalent to the Starobinsky model, once we identify

$$M^2 = \frac{\lambda}{3\xi^2}. \qquad (2.13)$$

Since we know that $M \approx 10^{-5}$, we get that $\xi^2 \approx 10^{10}\lambda$, which is, not surprisingly, the value needed in Higgs inflation. In addition, the vacuum energy which drives inflation is then

$$V_{\rm HI} = \frac{3}{4}M^2 M_{\rm p}^4 = \frac{\lambda}{4\xi^2}M_{\rm p}^4. \qquad (2.14)$$

B. Universal attractor models as a descendant of the Starobinsky model

The equivalence of the Starobinsky and Higgs inflation models is not merely an accident. In fact, the Starobinsky model is also equivalent during inflation to the general form of nonminimal coupling proposed in Ref. [17]

$$S_{\rm att} = \int d^4x\sqrt{-g}\left[\frac{1}{2}\Omega(\phi)R - \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - V_J(\phi)\right], \qquad (2.15)$$

with

$$\Omega(\phi) = M_{\rm p}^2 + \xi f(\phi), \qquad V_J = f^2(\phi). \qquad (2.16)$$

It should be noted that this class of models was discussed first in Ref. [22] where it was pointed that they are not technically "natural," as there is no obvious way for a symmetry, for example, to preserve the relation between the nonminimal coupling and the scalar potential.

Let us consider the case of large values of the coupling ξ As in the Higgs inflation case, during inflation, the dynamics are completely dominated by the potential so that we may ignore the scalar kinetic term. Therefore, the theory turns out to be written as

$$S_{\rm att} = \int d^4x\sqrt{-g}\left[\frac{M_{\rm p}^2}{2}R + \frac{1}{2}\xi f(\phi)R - f(\phi)^2\right]. \qquad (2.17)$$

We may integrate out the scalar through its equation of motion which is

$$\frac{1}{2}\xi R f' - 2f'f = 0, \qquad f' = \partial f/\partial\phi. \qquad (2.18)$$

The scalar field equation admits two solutions

$$f' = 0 \qquad (2.19)$$

and

$$f = \frac{1}{4}\xi R. \qquad (2.20)$$

Equation (2.19) is solved by a constant configuration $\phi = \phi_*$. Therefore, it corresponds to Einstein gravity with Planck mass $M_{\rm p}^2 + \xi f(\phi_*)$ and cosmological constant $\lambda^2 f(\phi_*)^2$. However, the second solution (2.20) gives

$$S_{\rm att} = \int d^4x\sqrt{-g}\left[\frac{M_{\rm p}^2}{2}R + \frac{\xi^2}{16}R^2\right] \qquad (2.21)$$

i.e., the Starobinsky model (2.1) again with the identification

$$M^2 = \frac{4}{3\xi^2}. \qquad (2.22)$$

The vacuum energy that drives inflation turns out to be for in this case

$$V_{\rm att} = \frac{3}{4}M^2 M_{\rm p}^4 = \frac{M_{\rm p}^4}{\xi^2}. \qquad (2.23)$$

C. Higher dimensional Starobinsky model descendants

Let us now discuss the higher dimensional generalization of the Starobinsky model with the action of the form

$$S = \int d^dx\sqrt{-g}\left(\frac{M_*^{d-2}}{2}\mathcal{R} + a\mathcal{R}^b\right), \qquad (2.24)$$

where $\mathcal{R}$ is the $(4+d)$-dimensional Ricci scalar, M_* is the corresponding Planck mass and a and b are dimensionless parameters. This higher dimensional theory can be linearized in the scalar curvature as usual by introducing an auxiliary field ϕ

$$S = \int \mathrm{d}^d x \sqrt{-g}\left(\frac{M_*^{d-2}}{2}\mathcal{R} + w\phi^2\mathcal{R} - \phi^{\frac{2b}{b-1}}\right), \tag{2.25}$$

where

$$w = \frac{b}{b-1}((b-1)a)^{\frac{1}{b}}. \tag{2.26}$$

By making the conformal transformation to the metric $g_{\mu\nu} \to \Omega^2 g_{\mu\nu}$, where

$$\Omega^{d-2} = \left(1 + \frac{2w\phi^2}{M_*^{d-2}}\right)^{-1}, \tag{2.27}$$

we may write the action (2.25) as

$$S = \int \mathrm{d}^d x \sqrt{-g}\left(\frac{M_*^{d-2}}{2}\mathcal{R} - \frac{1}{2}(d-1)(d-2)M_*^{d-2}(\partial_\mu \log\Omega)^2 - V_0\{(\Omega^{2-d}-1)\Omega^{\frac{(b-1)d}{b}}\}^{\frac{b}{b-1}}\right), \tag{2.28}$$

where

$$V_0 = \frac{M_*^{\frac{b(d-2)}{b-1}}}{(2w)^{\frac{b}{(b-1)}}}. \tag{2.29}$$

Clearly, in order to get a Starobinsky-like model, we need

$$d - 2 = \frac{b-1}{b}d \quad \text{or} \quad b = \frac{d}{2}. \tag{2.30}$$

Then the action (2.28) turns out to be

$$S = \int \mathrm{d}^d x \sqrt{-g}\left[\frac{M_*^{d-2}}{2}\mathcal{R} - \frac{1}{2}(d-1)(d-2)M_*^{d-2}(\partial_\mu \log\Omega)^2 - V_0(1-\Omega^{d-2})^{\frac{d}{d-2}}\right]. \tag{2.31}$$

After parametrizing Ω as

$$\log\Omega = -\frac{1}{\sqrt{(d-1)(d-2)}}\frac{\psi}{M_*{}^{(d-2)/2}}, \tag{2.32}$$

we get that

$$S = \int \mathrm{d}^d x \sqrt{-g}\left[\frac{M_*^{d-2}}{2}\mathcal{R} - \frac{1}{2}\partial_\mu\psi\partial^\mu\psi - V_0\left(1 - e^{-\sqrt{\frac{d-2}{d-1}}\frac{\psi}{M_*^{\frac{d-2}{2}}}}\right)^{\frac{d}{d-2}}\right]. \tag{2.33}$$

After a dimensional reduction in a $d-4$ torus T^{d-4}, we get the four-dimensional action

$$S = \int \mathrm{d}^4 x \sqrt{-g}\left[\frac{M_\mathrm{p}^2}{2}R - \frac{1}{2}\partial_\mu\chi\partial^\mu\chi - V_0\left(1 - e^{-\sqrt{\frac{d-2}{d-1}}\frac{\chi}{M_\mathrm{p}}}\right)^{\frac{d}{d-2}}\right] \tag{2.34}$$

after identifying

$$\chi = V_{d-4}^{1/2}\psi, \qquad V_{d-4}M_*^{d-2} = M_\mathrm{p}^2, \tag{2.35}$$

where V_{d-4} is the volume of T^{d-4}. We assume of course that the torus moduli or at least its volume modulus is stabilized. The potential of this generalized Starobinsky model is of the general form

$$V = V_0\left(1 - e^{\alpha\frac{\phi}{M_\mathrm{p}}}\right)^\beta, \tag{2.36}$$

which is a kind of T model [20]. For such a potential, it is straightforward to calculate the inflationary predictions. We find that

$$n_s \approx 1 - \frac{2}{N}, \qquad r \approx \frac{8}{\alpha^2 N^2}, \tag{2.37}$$

where $1/N_0 = \alpha\sqrt{2}$ and we have taken the limit $N \gg N_0$. In this limit, this is the same with the T-model predictions [20,26] as during inflation; β can be absorbed, to leading order, by an appropriate shift of ϕ.

We conclude this section with a comment on the conformally invariant SO(1,1) two-field model of Ref. [20] described by the Lagrangian

$$\mathcal{L} = \sqrt{-g}\left[\frac{1}{2}\partial_\mu\chi\partial^\mu\chi + \frac{\chi^2}{12}R - \frac{1}{2}\partial_\mu\phi\partial^\mu\phi + \frac{\phi^2}{12}R - \frac{\lambda}{4}(\phi^2 - \chi^2)^2\right]. \tag{2.38}$$

The field χ has a wrong kinetic term, and it was called conformon in Ref. [20]. Clearly the Lagrangian (2.38) is invariant under SO(1,1) rotations of (ϕ, χ). Therefore, one may fix this symmetry either by going to the Einstein frame $\chi^2 - \phi^2 = 6M_\mathrm{p}^2$ or to the Jordan frame $\chi = \sqrt{6}M_\mathrm{p}$. Both gauge fixings lead to

$$\mathcal{L} = \sqrt{-g}\left(\frac{M_\mathrm{p}^2}{2}R - \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - 9\lambda M_\mathrm{p}^4\right). \tag{2.39}$$

Here, we ignore as we did above the kinetic terms, assuming that they are small compared to the potential term. In this case, ϕ and χ are auxiliaries which can be integrated out to give

$$\mathcal{L} = \sqrt{-g}\frac{1}{144\lambda}R^2. \tag{2.40}$$

This is nothing else than the Starobinsky model in the $M_{\rm p} \to \infty$ limit. Therefore, again the conformally invariant SO(1,1) symmetric two-field model is a particular limit of the Starobinsky theory, at least in the region where scalar kinetic terms can be ignored. Note that (2.40) propagates a graviton and a scalar, as can be seen in the linear representation

$$\mathcal{L} = \sqrt{-g}(\varphi R - 36\lambda\varphi^2). \tag{2.41}$$

By integrating out φ we get the R^2 theory in (2.40). By going to the Einstein frame by means of the conformal transformation

$$g_{\mu\nu} \to \frac{M_{\rm p}^2}{2\varphi} g_{\mu\nu} \tag{2.42}$$

we get

$$\mathcal{L} = \sqrt{-g}\left(\frac{M_{\rm p}^2}{2} R - \frac{3}{2\varphi^2}\partial_\mu\varphi\partial^\mu\varphi - 9\lambda M_{\rm p}^4\right) \tag{2.43}$$

which is (2.39) after the transformation $\varphi = e^{\phi/\sqrt{3}}$.

III. DISTINGUISHING THE STAROBINSKY MODEL FROM ITS DESCENDANTS

From the discussion in the previous section, one can conclude that the Starobinsky model and its descendants differ only in their kinetic terms. Therefore, a reasonable question to ask is to which level this difference may be appreciated in the observables. Since the first slow-roll parameter $\epsilon = -\dot{H}/H^2$ (where H is the Hubble rate during inflation) parametrizes the kinetic energy [2], it is expected that differences between the Starobinsky model and its descendants appear at the level of differences in the slow-roll parameter ϵ. For the Starobinsky model, the slow-roll parameters are given by

$$\epsilon_{\rm S} \approx -\frac{3}{4N^2}, \tag{3.1}$$

$$\eta_{\rm S} \approx -\frac{1}{N}. \tag{3.2}$$

Now let us consider the Higgs inflation model and rewrite it in the Einstein frame. Redefining the metric as

$$g_{\mu\nu} \to \left(1 + \xi\frac{h^2}{M_{\rm p}^2}\right)^{-1} g_{\mu\nu}, \tag{3.3}$$

the action turns out to be

$$S_{\rm HI} = \int {\rm d}^4x\sqrt{-g}\left\{\frac{M_{\rm p}^2}{2}R - \frac{1}{2}\left(\frac{1}{1+\xi\frac{h^2}{M_{\rm p}^2}} + 6\xi^2\frac{h^2}{M_{\rm p}^2}\frac{1}{(1+\xi\frac{h^2}{M_{\rm p}^2})^2}\right)\partial_\mu h\partial^\mu h - \frac{\lambda}{4}\frac{h^4}{(1+\frac{\xi h^2}{M_{\rm p}^2})^2}\right\}. \tag{3.4}$$

Let us now compare this theory with the Starobinsky theory in the representation (2.9) which in the Einstein frame is written similarly as

$$S_{\rm S} = \int {\rm d}^4x\sqrt{-g}\left(\frac{M_{\rm p}^2}{2}R - \frac{6}{2}\xi^2\frac{h^2}{M_{\rm p}^2}\frac{1}{(1+\xi\frac{h^2}{M_{\rm p}^2})^2}\partial_\mu h\partial^\mu h - \frac{\lambda}{4}\frac{h^4}{(1+\frac{\xi h^2}{M_{\rm p}^2})^2}\right). \tag{3.5}$$

The difference between the two theories is evident. They differ by a factor of

$$\Delta\mathcal{L} = -\frac{1}{2}\frac{1}{1+\xi\frac{h^2}{M_{\rm p}^2}}\partial_\mu h\partial^\mu h, \tag{3.6}$$

which is precisely the Higgs kinetic term we neglected to arrive at with the Starobinsky theory in the Einstein frame. Here we should stress that the fundamental difference between the Higgs inflation and the Starobinsky model resides in the scalar kinetic term in the Jordan frame. For the Starobinsky model, there is no kinetic term for the auxiliary field ϕ in the linear representation of the model. This has the effect of making the parameter ξ irrelevant, as it can be completely absorbed in the scalar field and it is redundant. In the case of Higgs inflation, there is a kinetic term for the Higgs field to start with, as it is a real field in the Jordan frame and not an auxiliary. In this case, therefore, ξ cannot anymore be absorbed; it is not redundant, and as we see, it lowers the cutoff by a factor ξ^{-1} as compared to the Starobinsky model.

The slow-roll parameters for Higgs inflation and the Starobinsky theory are given by

$$\epsilon_{\rm HI,S} = \frac{M_{\rm p}^2}{2}\left(\frac{1}{V}\frac{\partial V}{\partial\chi}\right)^2 = \frac{M_{\rm p}^2}{2}\left(\frac{1}{V}\frac{\partial V}{\partial h}\right)^2\left(\frac{\partial h}{\partial\chi}\right)^2, \tag{3.7}$$

where χ is the canonically normalized scalar, different for Higgs and Starobinsky models, and V is the common potential

$$V = \frac{\lambda}{4}\frac{h^4}{(1+\frac{\xi h^2}{M_\mathrm{p}^2})^2}. \tag{3.8}$$

Then, since

$$\frac{\partial h}{\partial \chi} = \left(\frac{1}{1+\xi\frac{h^2}{M_\mathrm{p}^2}} + 6\xi^2\frac{h^2}{M_\mathrm{p}^2}\frac{1}{(1+\xi\frac{h^2}{M_\mathrm{p}^2})^2}\right)^{-1/2} \tag{3.9}$$

for Higgs inflation and

$$\frac{\partial h}{\partial \chi} = \left(6\xi^2\frac{h^2}{M_\mathrm{p}^2}\frac{1}{(1+\xi\frac{h^2}{M_\mathrm{p}^2})^2}\right)^{-1/2} \tag{3.10}$$

for the Starobinsky model, we find that

$$\epsilon_\mathrm{HI} = \frac{M_\mathrm{p}^2}{2}\left(\frac{1}{V}\frac{\partial V}{\partial h}\right)^2\left(\frac{1}{1+\xi\frac{h^2}{M_\mathrm{p}^2}} + 6\xi^2\frac{h^2}{M_\mathrm{p}^2}\frac{1}{(1+\xi\frac{h^2}{M_\mathrm{p}^2})^2}\right)^{-1}, \tag{3.11}$$

$$\epsilon_\mathrm{S} = \frac{M_\mathrm{p}^2}{2}\left(\frac{1}{V}\frac{\partial V}{\partial h}\right)^2\left(6\xi^2\frac{h^2}{M_\mathrm{p}^2}\frac{1}{(1+\xi\frac{h^2}{M_\mathrm{p}^2})^2}\right)^{-1}. \tag{3.12}$$

Since the number of e-folds till the end of inflation is related to h as $N \approx (6\xi h^2/8M_\mathrm{p}^2)$, we get that

$$\frac{\epsilon_\mathrm{HI}}{\epsilon_\mathrm{S}} = \frac{8N\xi}{1+\frac{4}{3}N+8N\xi} = 1-\frac{1}{6\xi} \simeq 1-\frac{10^{-5}}{6\lambda}. \tag{3.13}$$

Even though the slow-roll parameter enters with a factor of 6ϵ in the spectral index n_s, the difference is too small to be detectable. Another difference between the Starobinsky model and the Higgs inflation model is their corresponding reheating temperatures [19]: $T_\mathrm{RH} \simeq 3\times 10^9$ GeV and $T_\mathrm{RH} \simeq 6\times 10^{13}$ GeV, respectively. This leads to a difference in the predicted value of spectral index at the level of 10^{-3} [19]. As we mentioned in the Introduction, this difference is larger than the typical Planck error only if strong assumptions are made about the reionization history, the primordial Helium abundance and the effective number of neutrino.

Let us now turn to the universal attractor models. The general class of models (2.15) can be written in the Einstein frame by the conformal transformation

$$g_{\mu\nu} \to \left(1+\frac{\xi f(\phi)}{M_\mathrm{p}^2}\right)^{-2} g_{\mu\nu}, \tag{3.14}$$

and it is explicitly written as

$$S_\mathrm{att} = \int \mathrm{d}^4x\sqrt{-g}\left[\frac{M_\mathrm{p}^2}{2}R - \frac{3}{4}\frac{\xi^2 f'^2}{M_\mathrm{p}^2}\frac{\partial_\mu\phi\partial^\mu\phi}{\left(1+\frac{\xi f}{M_\mathrm{p}^2}\right)^2} - \frac{1}{2}\frac{\partial_\mu\phi\partial^\mu\phi}{1+\frac{\xi f}{M_\mathrm{p}^2}} - \frac{f^2}{\left(1+\frac{\xi f}{M_\mathrm{p}^2}\right)^2}\right]. \tag{3.15}$$

Similarly, the Starobinksy model in the representation (2.17) can be written as

$$S_\mathrm{S} = \int \mathrm{d}^4x\sqrt{-g}\left[\frac{M_\mathrm{p}^2}{2}R - \frac{3}{4}\frac{\xi^2}{M_\mathrm{p}^2}f'^2\frac{\partial_\mu\phi\partial^\mu\phi}{\left(1+\frac{\xi f}{M_\mathrm{p}^2}\right)^2} - \frac{f^2}{\left(1+\frac{\xi f}{M_\mathrm{p}^2}\right)^2}\right]. \tag{3.16}$$

Clearly, the two models differ in their kinetic terms

$$\Delta\mathcal{L} = -\frac{1}{2}\sqrt{-g}\frac{\partial_\mu\phi\partial^\mu\phi}{1+\frac{\xi f}{M_\mathrm{p}^2}}, \tag{3.17}$$

and the difference is tiny for large values of ξ. The slow-roll parameters for the above general classes of inflation models and the Starobinsky theory are given by

$$\epsilon_\mathrm{att,\,S} = \frac{M_\mathrm{p}^2}{2}\left(\frac{1}{V}\frac{\partial V}{\partial \chi}\right)^2 = \frac{M_\mathrm{p}^2}{2}\left(\frac{1}{V}\frac{\partial V}{\partial \phi}\right)^2\left(\frac{\partial \phi}{\partial \chi}\right)^2, \tag{3.18}$$

where χ is the canonically normalized scalar, different for the two models, and V is the common potential

$$V = \frac{f^2}{(1+\frac{\xi f}{M_\mathrm{p}^2})^2}. \tag{3.19}$$

Let us discuss the particular, but sufficiently generic case of $f = \phi^n/M_\mathrm{p}^{n-2}$, for which

$$V = \frac{\phi^{2n}}{M_\mathrm{p}^{2n-4}(1+\xi\frac{\phi^n}{M_\mathrm{p}^n})}. \tag{3.20}$$

Then, since

$$\frac{\partial \phi}{\partial \chi} = \left(\frac{1}{1+\xi\frac{\phi^n}{M_\mathrm{p}^n}} + \frac{3\xi^2 n^2}{2}\frac{\phi^{2n-2}}{M_\mathrm{p}^{2n-2}}\frac{1}{\left(1+\xi\frac{\phi^n}{M_\mathrm{p}^n}\right)^2}\right)^{-1/2} \tag{3.21}$$

for general models of nonminimally coupled inflation and

$$\frac{\partial \phi}{\partial \chi} = \left(\frac{3\xi^2 n^2}{2}\frac{\phi^{2n-2}}{M_\mathrm{p}^{2n-4}}\frac{1}{(1+\xi\frac{\phi^n}{M_\mathrm{p}^n})^2}\right)^{-1/2} \tag{3.22}$$

for the Starobinsky model, we find that

$$\epsilon_{\rm att} = \frac{M_{\rm p}^2}{2}\left(\frac{1}{V}\frac{\partial V}{\partial\phi}\right)^2\left(\frac{1}{1+\xi\frac{\phi^n}{M_{\rm p}^n}} + \frac{3\xi^2 n^2}{2}\frac{\phi^{2n-2}}{M_{\rm p}^{2n-2}}\frac{1}{(1+\xi\frac{\phi^n}{M_{\rm p}^n})^2}\right)^{-1} \tag{3.23}$$

$$\epsilon_{\rm S} = \frac{M_{\rm p}^2}{2}\left(\frac{1}{V}\frac{\partial V}{\partial\phi}\right)^2\left(\frac{3\xi^2 n^2}{2}\frac{\phi^{2n-2}}{M_{\rm p}^{2n-2}}\frac{1}{(1+\xi\frac{\phi^n}{M_{\rm p}^n})^2}\right)^{-1}. \tag{3.24}$$

Since the number of e-folding is related to ϕ as

$$N \approx \frac{3\xi\phi^n}{4M_{\rm p}^n}, \tag{3.25}$$

we infer that

$$\frac{\epsilon_{\rm att}}{\epsilon_{\rm S}} \approx 1 - \frac{N^{\frac{2}{n}-1}}{2n^2\xi^{\frac{2}{n}}}\left(\frac{4}{3}\right)^{2/n}. \tag{3.26}$$

This always deviates from unity by a quantity smaller that 10^{-3}, and therefore the difference is not observable.

IV. EFFECTIVE CUTOFF SCALES

One (somewhat controversial) issue is the natural cutoff of the theories we have discussed so far. As there exists another mass M (or $1/\xi^{1/2}$), which enters besides the dimensionful Planck mass $M_{\rm p}$, it is natural to expect that the cutoff of the theory may not be $M_{\rm p}$, but a ratio of it by appropriate power of M (or ξ). If this power is high enough, it may happen that the cutoff is quit low, lower than the inflationary scale. In such a case, the discussion of inflation cannot be trusted, or it is questionable, to say the least. Below we find the cutoffs of the models discussed so far by considering the scalar field in the Einstein frame as a one-dimensional σ model. Then, as mentioned, the expansion of its kinetic term for small values of the field reveals the cutoff of the theory and, above all, the differences among the models.

The Starobinsky model (3.5) can be expanded as

$$S = \int d^4x\sqrt{-g}\left[\frac{M_{\rm p}^2}{2}R - \frac{1}{2}\left(\frac{\xi h^2}{M_{\rm p}^2} + 6\frac{\xi^2 h^2}{M_{\rm p}^2} + \cdots\right)\partial_\mu h\partial^\mu h - \frac{\lambda}{4}h^4\left(1 - 2\frac{\xi h^2}{M_{\rm p}^2} + \cdots\right)\right]. \tag{4.1}$$

We should canonically normalize the leading kinetic term. Thus, after defining $h^2 = \frac{M_{\rm p}\psi}{\sqrt{3}\xi}$, we get that the action turns out to be

$$S = \int d^4x\sqrt{-g}\left[\frac{M_{\rm p}^2}{2}R - \frac{1}{2}\left(1 - \frac{\psi}{\sqrt{3}M_{\rm p}} + \cdots\right)\partial_\mu\psi\partial^\mu\psi - \frac{\lambda}{12}\frac{M_{\rm p}^2}{\xi^2}\psi^2\left(1 - 2\frac{\psi}{\sqrt{3}M_{\rm p}} + \cdots\right)\right]. \tag{4.2}$$

From the above form of the action we see that the cutoff $\Lambda_{\rm S}$ of the Starobinsky theory is, as already found in [23],

$$\Lambda_{\rm S} = M_{\rm p}. \tag{4.3}$$

A simple inspection of Eq. (2.5) shows that

$$V_{\rm S} \ll \Lambda_{\rm S}^4, \tag{4.4}$$

indicating the internal consistency of the model [23]. The Higgs inflation action (3.4) on the other hand can be expanded as

$$S_{\rm HI} = \int d^4x\sqrt{-g}\left[\frac{M_{\rm p}^2}{2}R - \frac{1}{2}\left(1 + \frac{\xi h^2}{M_{\rm p}^2} + 6\frac{\xi^2 h^2}{M_{\rm p}^2} + \cdots\right)\partial_\mu h\partial^\mu h - \frac{\lambda}{4}h^4\left(1 - 2\frac{\xi h^2}{M_{\rm p}^2} + \cdots\right)\right]. \tag{4.5}$$

Here the leading kinetic term is canonically normalized, and therefore, since $\xi \gg 1$, we find that the cutoff is [21–23]

$$\Lambda_{\rm HI} = \frac{M_{\rm p}}{\xi}. \tag{4.6}$$

This should be compared with the vacuum energy that drives inflation in Eq. (2.14), from where we get that

$$V_{\rm HI} \gg \Lambda_{\rm HI}^4, \tag{4.7}$$

making the consistency of the model questionable. This simple argument has been criticized in Ref. [24] where it was observed that the cutoff should be field dependent as the kinetic term is noncanonical. This argument would give a cutoff that during inflation, when $h \gg M_{\rm p}/\xi^{1/2}$, is even larger then the Planckian scale. However, we disagree with this approach. The presence of a cutoff $\Lambda_{\rm HI} \sim M_{\rm p}/\xi$ at lower values of the field cannot be avoided, and it signals the breakdown of the model in that field range. The small field region is "tested" by the dynamics during the reheating stage, and one may not simply disregard this point by invoking that the inflationary field range is the one of interest. It should also be mentioned here a related problem, the naturalness of the model. The only way to solve this inconsistency is to add new degrees of freedom at energies $\sim M_{\rm p}/\xi^{1/2}$ in a way that does not spoil the flatness of the inflation potential, as for example in the model discussed in Ref. [27]. In such a case, however, the predictability of Higgs inflation is lost, as there is now a strong dependence on the new physics assumed to appear at $\sim M_{\rm p}/\xi^{1/2}$.

Similar considerations can be made for the attractor models. To find the cutoff $\Lambda_{\rm att}$, we expand the action (3.15) as

$$S_{\rm att} \approx \int d^4x\sqrt{-g}\left\{\frac{M_{\rm p}^2}{2}R - \frac{1}{2}\left[1 - \frac{\xi f}{M_{\rm p}} + \frac{\xi^2 f^2}{M_{\rm p}^2} + \cdots 3\frac{\xi^2 f'^2}{M_{\rm p}^2}\left(1 - 2\frac{\xi f}{M_{\rm p}^2} + \frac{\xi^2 f'^2}{M_{\rm p}^4} + \cdots\right)\right]\partial_\mu\phi\partial^\mu\phi - f^2(1 - 2\xi f + \cdots)\right\}. \tag{4.8}$$

For a polynomial form $f(\phi) = \phi^n/M_{\rm p}^{n-2}$, with $n \neq 1$, the cut-off is determined by $\xi^2 f'^2$ term in Eq. (4.8) and reads

$$\Lambda_{\rm att} = \frac{M_{\rm p}}{\xi^{\frac{1}{n-1}}}, \tag{4.9}$$

which is below $M_{\rm p}$ as ξ is large (and n is different from unity). Moreover, the vacuum energy during inflation is given in Eq. (2.23), which in terms of the cutoff (4.9) is written as

$$V_{\rm att} = \xi^{\frac{6-2n}{n-1}}\Lambda_{\rm att}^4. \tag{4.10}$$

Clearly, only for $n > 7/3$ the vacuum energy satisfies $V_{\rm att} \ll \Lambda_{\rm att}^4$ and the model makes sense.

The case $n = 1$ is special, and we consider it separately. The reason is that $\xi^2 f'^2$ dominates, and a constant rescaling of the scalar, similar to the one in the Starobinsky model, is needed to canonically normalize the leading kinetic term. It is known that the simplest chaotic inflation has severe problems with the recent Planck data. Its inflationary dynamics is described by the action

$$S_m = \int d^4x\sqrt{-g}\left(\frac{M_{\rm p}^2}{2}R - \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}m^2\phi^2\right), \tag{4.11}$$

which predicts [2]

$$n_s - 1 = -\frac{2}{N}, \qquad r = \frac{8}{N}, \tag{4.12}$$

for the primordial tilt n_s and the tensor-to-scalar ratio r, values which lie outside the joint 95% C.L. for the Planck data. Let us now consider instead of the action (4.11), a nonminimally coupled chaotic model

$$S = \int d^4x\sqrt{-g}\left(\frac{M_{\rm p}^2}{2}R + \xi M_{\rm p}\phi R - \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}m^2\phi^2\right), \tag{4.13}$$

where ξ is a dimensionless parameter. Clearly, as discussed above, during inflation the inflation kinetic term is small compared to the potential, and thus the model is described effectively by

$$S = \int d^4x\sqrt{-g}\left(\frac{M_{\rm p}^2}{2}R + \xi M_{\rm p}\phi R - \frac{1}{2}m^2\phi^2\right). \tag{4.14}$$

The field ϕ can be integrated out leading again to the Starobinsky model

$$S = \int d^4x\sqrt{-g}\left(\frac{M_{\rm p}^2}{2}R + \xi^2\frac{M_{\rm p}^2}{2m^2}R^2\right). \tag{4.15}$$

Therefore, the nonminimally coupled chaotic inflationary model (4.14) is equivalent during inflation to the Starobinsky gravity (2.1) with $M_{\rm p}^2/12M^2 = M_{\rm p}^2\xi^2/2m^2$. As a result, since $M \approx 10^{-5}$, we get that

$$\xi \approx 10^5 m, \tag{4.16}$$

whereas the primordial tilt and the tensor-to-scalar ratio are now $(n_s - 1) \simeq -2/N$ and $r = 12/N^2$. Let us now write the action (4.14) in the Einstein frame. For this, we need to make the following conformal transformation:

$$g_{\mu\nu} \to \left(1 + \frac{2\xi\phi}{M_{\rm p}}\right)^{-1} g_{\mu\nu}, \tag{4.17}$$

and the action becomes

$$S = \int d^4x\sqrt{-g}\left[\frac{M_{\rm p}^2}{2}R - 3\xi^2\frac{\partial_\mu\phi\partial^\mu\phi}{(1 + \frac{2\xi\phi}{M_{\rm p}})^2} - \frac{1}{2}\frac{\partial_\mu\phi\partial^\mu\phi}{1 + \frac{2\xi\phi}{M_{\rm p}}} - \frac{1}{2}m^2\phi^2\left(1 + \frac{2\xi\phi}{M_{\rm p}}\right)^{-2}\right]. \tag{4.18}$$

For large values of the scalar field ϕ ($\phi \gg M_{\rm p}/2\xi$), we have

$$S_{nm} \approx \int d^4x\sqrt{-g}\left\{\frac{M_{\rm p}^2}{2}R - \frac{1}{2}\left(\frac{3M_{\rm p}^2}{2\phi^2} + \frac{M_{\rm p}}{2\xi\phi}\right)\partial_\mu\phi\partial^\mu\phi - V_0\left(1 - \frac{M_{\rm p}}{\xi\phi} + \cdots\right)\right\} \tag{4.19}$$

where

$$V_0 = \frac{m^2 M_{\rm p}^2}{8\xi^2} \tag{4.20}$$

is the vacuum energy-driving inflation. Then one may easily verify that this theory is the Starobinksy theory for

$$\frac{3M_{\rm p}^2}{2\phi^2} \gg \frac{M_{\rm p}}{2\xi\phi}. \tag{4.21}$$

In other words, for ϕ in the range

$$\frac{M_{\rm p}}{2\xi} \ll \phi \ll \frac{3}{2}\xi M_{\rm p}, \tag{4.22}$$

the nonminimal chaotic inflation effectively coincides with the Starobinsky model. Note that (4.22) implies that $\xi \gg 1$.

The action (4.18) can be expanded also for small values of ϕ. However, in this case, there is no canonically normalized leading-order kinetic term for the scalar. Thus, after defining $\chi = \sqrt{6}\xi\phi$, we have

$$S \approx \int {\rm d}^4x\sqrt{-g}\left\{\frac{M_{\rm p}^2}{2}R - \frac{1}{2}\left(1-\frac{4\chi}{\sqrt{6}M_{\rm p}} - \frac{\chi}{3\xi M_{\rm p}}\right)\partial_\mu\chi\partial^\mu\chi - \frac{1}{12}\frac{m^2M_{\rm p}^2}{\xi^2}\chi^2\left(1-\frac{4\chi}{\sqrt{6}M_{\rm p}}\right)+\cdots\right\}. \tag{4.23}$$

From the form of the action, it follows that the cutoff Λ of the nonminimal chaotic inflation is indeed the Planckian mass, $\Lambda = M_{\rm p}$, with $V_0 \ll \Lambda^4$. This is exactly what happens in the Starobinsky theory, where the absence of the canonically normalized leading kinetic term pushes the cutoff to the Planck scale.

V. CONCLUSIONS

In this paper we have discussed the relation of certain inflationary models to the Starobinsky theory. In particular, we have pointed out that the agreement of these models with the recent Planck measurements is due to the fact that during inflation they are effectively described by the Starobinsky theory. In this respect, the Starobinsky theory is a prototype of theories where the scalar potential has a plateau for large values of the scalar field. The examples we discussed here in detail are the Higgs inflation model and the universal attractor models at strong coupling, the dynamics of which coincides to leading order in the slow-roll parameter with that of the Starobinsky theory. However, they differ from the latter since the scalar in the Starobinsky theory is auxiliary in the Jordan frame and turns out to be propagating only in the Einstein frame.

Although these models are effectively equivalent to the Starobinsky theory for large values of the fields, they are not equivalent for small values. In particular, one expects large differences in the small-field regime. Therefore, one may correctly identify the range of the validity of the theory by determining its cutoff scale, if it is considered as an effective field theory. We have discussed the cutoff by looking in the scalar kinetic term, which is similar to kinetic term of a one-dimensional σ model. We have found that, although the cutoff of the Starobinsky theory is the Planck scale, for a polynomial function $f(\phi) = \phi^n/M_{\rm p}^{n-2}$ in the general universal attractor model, the cutoff is lower than the inflationary scale for $n < 7/3$ (this case includes also Higgs inflation for $n = 2$). However, the case $n = 1$ is particular and we have discussed it in more detail. In particular, beyond being in agreement with the data, it is valid up to Planckian scales.

ACKNOWLEDGMENTS

We thank J. R. Espinosa for useful discussions. A. R. is supported by the Swiss National Science Foundation (SNSF) project The non-Gaussian universe (Project No. 200021140236). The research of A. K. was implemented under the “Aristeia” Action of the Operational Programme Education and Lifelong Learning and is cofunded by the European Social Fund (ESF) and National Resources. It is partially supported by the European Union’s Seventh Framework Programme (FP7/2007-2013) under REA Grant Agreement No. 329083.

[1] P. A. R. Ade *et al.* (Planck Collaboration), arXiv:1303.5082.
[2] D. H. Lyth and A. Riotto, Phys. Rep. **314**, 1 (1999).
[3] A. A. Starobinsky, Phys. Lett. **91B**, 99 (1980).
[4] V. F. Mukhanov and G. V. Chibisov, Pisma Zh. Eksp. Teor. Fiz. **33**, 549 (1981) [JETP Lett. **33**, 532 (1981)]; A. A. Starobinsky, Sov. Astron. Lett. **9**, 302 (1983).
[5] J. Ellis, D. V. Nanopoulos, and K. A. Olive, Phys. Rev. Lett. **111**, 111301 (2013); J. Cosmol. Astropart. Phys. 10 (2013) 009.
[6] R. Kallosh and A. Linde, J. Cosmol. Astropart. Phys. 06 (2013) 028.
[7] F. Farakos, A. Kehagias, and A. Riotto, Nucl. Phys. **B876**, 187 (2013).
[8] S. Ferrara, R. Kallosh, A. Linde, and M. Porrati, Phys. Rev. D **88**, 085038 (2013); S. Ferrara, R. Kallosh, A. Linde, and M. Porrati, J. Cosmol. Astropart. Phys. 11 (2013), 046.
[9] S. Ferrara, R. Kallosh, and A. Proeyen, J. High Energy Phys. 11 (2013) 134.
[10] S. Ferrara, P. Fre, and A. S. Sorin, arXiv:1311.5059.
[11] S. V. Ketov and A. A. Starobinsky, J. Cosmol. Astropart. Phys. 08 (2012), 022; S. V. Ketov and S. Tsujikawa, Phys. Rev. D **86**, 023529 (2012); S. V. Ketov, arXiv:1309.0293.
[12] S. Cecotti, Phys. Lett. B **190**, 86 (1987).
[13] S. Cecotti, S. Ferrara, M. Porrati, and S. Sabharwal, Nucl. Phys. **B306**, 160 (1988).
[14] B. L. Spokoiny, Phys. Lett. **147B**, 39 (1984).
[15] R. Fakir and W. G. Unruh, Phys. Rev. D **41**, 1792 (1990).
[16] F. L. Bezrukov and M. Shaposhnikov, Phys. Lett. B **659**, 703 (2008).

[17] R. Kallosh, A. Linde, and D. Roest, Phys. Rev. Lett. **112**, 011303 (2014).
[18] R. Kallosh, A. Linde, and D. Roest, J. High Energy Phys. 11 (2013) 198.
[19] F. L. Bezrukov and D. S. Gorbunov, Phys. Lett. B **713**, 365 (2012).
[20] R. Kallosh and A. Linde, J. Cosmol. Astropart. Phys. 07 (2013) 002.
[21] C. P. Burgess, H. M. Lee, and M. Trott, J. High Energy Phys. 09 (2009) 103; 07 (2010) 007.
[22] J. L. F. Barbon and J. R. Espinosa, Phys. Rev. D **79**, 081302 (2009).
[23] M. P. Hertzberg, J. High Energy Phys. 11 (2010) 023.
[24] F. Bezrukov, A. Magnin, M. Shaposhnikov, and S. Sibiryakov, J. High Energy Phys. 01 (2011) 016.
[25] B. Whitt, Phys. Lett. **145B**, 176 (1984).
[26] R. Kallosh and A. Linde, J. Cosmol. Astropart. Phys. 10 (2013) 033; 12 (2013) 006.
[27] G. F. Giudice and H. M. Lee, Phys. Lett. B **694**, 294 (2011).