PHYSICAL REVIEW D **93,** 043510 (2016)

Large field inflation and gravitational entropy

Nemanja Kaloper,[1,*] Matthew Kleban,[2,†] Albion Lawrence,[3,‡] and Martin S. Sloth[4,§]

[1]*Department of Physics, University of California, Davis, California 95616, USA*
[2]*CCPP, Department of Physics, New York University, New York 10003, USA*
[3]*Martin Fisher School of Physics, Brandeis University, Waltham, Massachusetts 02453, USA*
[4]*CP^3-Origins, Center for Cosmology and Particle Physics Phenomenology, University of Southern Denmark, Campusvej 55, 5230 Odense M, Denmark*

(Received 26 December 2015; published 8 February 2016)

Large field inflation can be sensitive to perturbative and nonperturbative quantum corrections that spoil slow roll. A large number N of light species in the theory, which occur in many string constructions, can amplify these problems. One might even worry that in a de Sitter background, light species will lead to a violation of the covariant entropy bound at large N. If so, requiring the validity of the covariant entropy bound could limit the number of light species and their couplings, which in turn could severely constrain axion-driven inflation. Here we show that there is no such problem when we correctly renormalize models with many light species, taking the *physical* Planck scale to be $M_{pl}^2 \gtrsim N\mathcal{M}_{\rm UV}^2$, where $\mathcal{M}_{\rm UV}$ is the cutoff for the quantum field theory coupled to semiclassical quantum gravity. The number of light species then cancels out of the gravitational entropy of de Sitter or near-de Sitter backgrounds at leading order. Working in detail with N scalar fields in de Sitter space, renormalized to one loop order, we show that the gravitational entropy automatically obeys the covariant entropy bound. Furthermore, while the axion decay constant is a strong coupling scale for the axion dynamics, we show that it is *not* in general the cutoff of 4d semiclassical gravity. After renormalizing the two point function of the inflaton, we note that it is also controlled by scales much below the cutoff. We revisit N-flation and Kachru-Kallosh-Linde-Trivedi-type compactifications in this light, and show that they are perfectly consistent with the covariant entropy bound. Thus, while quantum gravity might yet spoil large field inflation, holographic considerations in the semiclassical theory do not obstruct it.

DOI: 10.1103/PhysRevD.93.043510

I. INTRODUCTION

Inflation is the best theoretical explanation of the large, old and smooth universe with small nearly scale invariant perturbations. It fits experimental tests perfectly. If, in addition to the observed scalar density fluctuation, tensor mode fluctuations of the CMB are directly observed, we could probe inflation in great detail. If these modes are generated by quantum fluctuations of the gravitational field, a direct observation by itself would imply that the effective field theory (EFT) models of inflation required to generate large tensor modes would have to operate up to scales around the grand unified theory (GUT) scale, which is close to the string or 10d/11d Planck scale in many conservative string theory scenarios. Moreover, to yield sufficiently long inflation such models would have to have very flat potentials over super-Planckian field ranges [1,2].

At such scales and over such field ranges, quantum field theory and quantum gravity effects correcting the dynamics can be significant. In this light it is reasonable to ask whether quantum gravity might provide any general constraints on these models, independent of a specific realization in string theory. While we do not have a complete theory of inflation in quantum gravity yet, there are hints and clues about how quantum gravity might influence the low energy theory. The two lines of inquiry which have received particular attention recently are the "weak gravity conjecture" (WGC) of [3] and the covariant entropy bounds [4]. The WGC uses some features of black hole entropy to place an upper bound on the mass of charged particles and/or the action of instantons leading to corrections to the axion potential. This can constrain some axion inflation models such as [5–8] for which the axion potential is of the sinusoidal form expected from the dilute gas approximation for instantons. There are more tenuous arguments [9–11] that axion monodromy inflation [12–29] is constrained by such considerations as well, although at present they are not excluded.

The argument from the covariant entropy bound [30] is based on the following logic. If there are many light weakly interacting species and a sufficiently long inflationary epoch[1] (or, a long lived metastable de Sitter space), the large number of species will thermalize with gravity and

*kaloper@physics.ucdavis.edu
†kleban@nyu.edu
‡albion@brandeis.edu
§sloth@cp3.dias.sdu.dk

[1]For earlier considerations of bounds on duration of inflation see [31,32].

overwhelm the geometric entropy. The setup is especially interesting as string-motivated inflationary constructions often consist of low energy theories with N light species, $N \gg 1$, which are weakly coupled to each other in the IR. According to the argument above, such models violate the covariant entropy bound, which requires that the entropy in a given Hubble-sized patch does not exceed the Gibbons-Hawking entropy. Requiring the validity of the covariant entropy bound thus constrains N, the couplings of these species, and the duration of inflation (or the lifetime of the metastable de Sitter space).

In what follows we will focus on this latter issue.[2] We specifically show that even for field theories with many light species, a correctly renormalized low energy theory, including the gravitational couplings, obeys the covariant entropy bound since the number of light species cancels from the entropy formulas to the leading order. Essentially, this arises since even if one starts with many weakly coupled species of particles, when gravity is turned on the renormalized effective field theory becomes strongly coupled well below the Planck scale, $\mathcal{M}_{\rm UV} \lesssim M_{pl}/\sqrt{N}$, which sets a cutoff on calculations using weakly coupled, semiclassical gravity. Beyond this scale, one needs a full-blown UV completion to follow the details of the dynamics. Nevertheless, the low energy description remains self-consistent, and obeys the semiclassical limits, as one might have expected from decoupling.

In more detail, the argument that entropy bounds are violated begins with an estimate of the contribution to the total de Sitter entropy of a field ϕ. The first step is the identification of a cutofff $\mathcal{M}_{\rm UV}$ for the dynamics of ϕ. For axions, this is identified with the axion decay constant f, which sets the periodicity in field space $\phi \equiv \phi + 2\pi f$. The next step is to count the number of patches of size $\mathcal{M}_{\rm UV}$ on the de Sitter horizon, leading to a contribution $S_\phi \sim \mathcal{M}_{\rm UV}^2/H^2$, where H is the Hubble scale. One then demands that $\sum_i S_{\phi_i} \leq \frac{M_{pl,4}^2}{H^2}$. Then if $\sqrt{N}\mathcal{M}_{\rm UV} > M_{pl,4}$, there appears to be an apparent violation of the covariant entropy bound due to a species problem.

The hole in this argument is that the formula $S \sim A/(4G_N)$ for 4d de Sitter or inflating universes is valid only when the underlying (semiclassical) 4d gravity is valid. This means that one must take into account the loop corrections to the gravitational sector, and consistently analyze the *renormalized* 4d effective field theory of gravity, accounting for the contributions of all the many light species to the relevant physical quantities which control the dynamics, including the cutoff and the dimensional couplings. In particular the route to the proper renormalization procedure must incorporate the following:

(i) The correct cutoff to impose is the scale at which 4d semiclassical gravity breaks down. This is generically at a scale $\mathcal{M}_{\rm UV}^2 \leq M_{pl}^2/N$. This cutoff follows from the well-known behavior of perturbative renormalization of gravity which involves the inclusion of higher dimension of irrelevant operators in the gravitational sector, which introduce a perturbative spin-2 ghost, with a mass $\sim 1/\mathcal{M}_{\rm UV}$ [33].

(ii) The correct value of $G_N \sim 1/M_{pl}^2$ to use is the renormalized Newton's constant at this scale, $1/G_{N,\rm ren} \sim M_{pl}^2 \sim M_{pl,\rm bare}^2 + N\mathcal{M}_{\rm UV}^2$. The inclusion of the contribution of N species at the scale $\mathcal{M}_{\rm UV}$ evades the species problem.

(iii) The axion decay constant f and the scale $\mathcal{M}_{\rm UV}$ may be very different. Specifically, $\mathcal{M}_{\rm UV}$ can be much smaller than the period. This may arise very simply in setups with intermediate-mass particles, as we will show explicitly using a two-axion model [34,35].

Footnote 2 in [30] dismisses significant renormalization of Newton's constant in models with many species of light fields, claiming there can be cancellations between corrections. The calculations in [36,37] show that while minimally coupled scalars and Weyl fermions contribute with the same sign, Abelian gauge fields contribute with the opposite sign. However, this is not a way out. First, for any field with spin less than 2, the divergent contributions to the entanglement entropy defined via the replica trick [38] have been shown to be *precisely* taken into account by the same fields' contribution to the renormalization of the gravitational action—see [39,40], and the references therein.[3] Second, Newton's constant is not the only place that quantum corrections to the gravitational action will occur, and the relative contributions of different fields will differ in these other terms, so that the estimate $M_{pl}/\sqrt{N}$ of the strong coupling scale remains appropriate.

While the itemized points above are individually discussed in the literature, the recurring confusions suggest that a unified and coherent discussion in the context of large field inflation is warranted. Hence we will provide a thorough review of these arguments (Sec. II), some additional calculations supporting them in cosmological settings (Sec. III), and a re-examination (Sec. IV) of the claimed constraints on N-flation and Kachru-Kallosh-Linde-Trivedi (KKLT)-type compactifications that are explicitly discussed in [30].

II. THE STRONG COUPLING SCALE FOR GRAVITY AND THE "SPECIES PROBLEM"

The classic calculations of the Bekenstein-Hawking entropy for black holes, and the Gibbons-Hawking entropy

[2]We stress that while this issue and its analysis at present appear distinct from the arguments about WGC in [3], Ref. [9] has suggestion a strong form of the WGC may be related to the entropic arguments in [30]. We have nothing to say about this possibility, but it is true that both are ultimately related to entropic considerations.

[3]The interpretation of the gauge field contributions to the entanglement entropy is a subject of ongoing research [41–48].

for cosmological spacetimes, are based crucially on semiclassical gravity. One can only use these results in the regime of validity of the (renormalized) semiclassical theory, and one must use the physical, renormalized couplings and scales at these energies. In this section we will describe the renormalization procedure and extract the behavior of the renormalized quantities, particularly the renormalized Planck scale, as a function of the number of light species of particles which appear in the loops that contribute to the effective action of gravity. We will also note that the higher dimension irrelevant operators in the gravitational sector, which arise from the loop corrections, yield a clear cutoff that determines the validity of the semiclassical approximation. These points have been noted in the context of black holes, via a variety of arguments elucidated below. We will close the section by recalling the emergence of the strong coupling scale in compactifications from $d > 4$ theories, and in the Randall-Sundrum scenario.

A. Perturbation theory arguments

We will first revisit perturbative renormalization of gravity due to exchange of virtual field theory degrees of freedom, starting with Einstein-Hilbert theory perturbatively quantized around a vacuum with maximal symmetry. The case of the Minkowski vacuum has been studied extensively [33,49,50]. The generalization to de Sitter vacua is straightforward, and has been considered in the context of computing black hole entropy in curved backgrounds. Here we will follow [37,51,52]. In this subsection we will focus on the renormalization of the action. Later on we will see how those results affect the horizon entropy.

If we start with the bare gravitational Lagrangian

$$\mathcal{L}_g = \frac{1}{16\pi G_N}(R - 2\Lambda) - \mathcal{L}_m(\phi) + \frac{1}{4\pi}[aR^2 + bR_{\mu\nu}R^{\mu\nu} + cR_{\mu\nu\kappa\rho}R^{\mu\nu\kappa\rho}] \quad (1)$$

where $\mathcal{L}_m$ is the quantum field theory of some matter, the bare quantities are G_N, the bare Newton's constant, Λ, the bare cosmological constant, and a, b, c, the higher dimension irrelevant operator bare couplings. These terms are set to cancel the one loop divergences in the theory due to matter couplings [49]. For simplicity we will take the matter sector to consist of N minimally coupled scalar fields, with only quadratic Lagrangians,

$$\mathcal{L}_m = \sum_{j=1}^{N} \frac{1}{2}[\partial_\mu \phi_j \partial^\mu \phi_j + m_j^2 \phi_j^2]. \quad (2)$$

This is sufficient for our purposes. Generalizations to other matter sectors are straightforward, and as we will discuss below, do not change the essential conclusions.

The one-loop contributions to the effective action from the matter sector will generically exhibit quartic, quadratic and logarithmic UV divergences [49]. The quartic UV divergence is the usual divergent contribution to the cosmological constant. If we truncate the matter theory to the quadratic Lagrangian (2) it may or may not appear depending on the regulator. The quadratic divergences are the wave-function renormalizations of the kinetic terms in (1), and include renormalizations of the additional "R^2" terms in the action.

To compute the one loop integrals, one first needs to regulate the divergent terms. We do so by introducing a system of Pauli-Villars regulators for every matter field in (2). The scheme is conceptually the same as in flat space, where one introduces a regulator for every divergent loop. If the cutoff is above the inverse curvature scale we can start with a locally flat region of space, introduce the regulators with minimal coupling to gravity. Because there are five distinct types of required counterterms, reflecting five different divergences, one needs five regulators for each matter scalar [52], denoted by ϕ_i ($i = 1, \ldots, 5$) and coming with different statistics, Δ_i (where $\Delta_i = \pm 1$ for commuting and anticommuting fields respectively). The regulator masses m_i are much larger than the matter ones in order to formally cancel the UV divergences, and define the UV cutoff, μ. The choice of the regulators and their statistics is determined by the requirements

$$\sum_{i=0}^{5} \Delta_i = 0 \quad \text{and} \quad \sum_{i=0}^{5} \Delta_i m_i^2 = 0 \quad (3)$$

ensuring finiteness of the regulated theory. Here m_0 is the mass of the original scalar field $\phi \equiv \phi_0$. Using this regularization procedure, the one-loop effective action is given by [52]

$$\mathcal{L}_g = -\frac{1}{8\pi}\left(\frac{\Lambda}{G_N} + \frac{\gamma}{4\pi}\right) + \frac{R}{16\pi}\left(\frac{1}{G_B} + \frac{\delta}{12\pi}\right) + \frac{1}{4\pi}\left[\left(a + \frac{\alpha}{576}\right)R^2 + \left(b - \frac{\alpha}{1440\pi}\right)R_{\mu\nu}R^{\mu\nu} + \left(c + \frac{\alpha}{1440\pi}\right)R_{\mu\nu\kappa\rho}R^{\mu\nu\kappa\rho}\right] \quad (4)$$

where

$$\alpha = N\sum_{i=0}^{5} \Delta_i \log m_i^2, \qquad \delta = N\sum_{i=0}^{5} \Delta_i m_i^2 \log m_i^2, \qquad \gamma = \frac{N}{2}\sum_{i=0}^{5} \Delta_i m_i^4 \log m_i^2. \quad (5)$$

These expressions are in fact dimensionless once the logs are summed up, due to the fact that the Δ_i are alternating

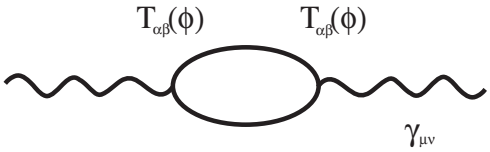

FIG. 1. One loop graviton vacuum polarization diagram.

numbers. One thus obtains the renormalized cosmological constant and Newton's constant

$$\frac{\Lambda^{\rm ren}}{G_N^{\rm ren}} = \frac{\Lambda}{G_N} + \frac{\gamma}{4\pi}, \qquad \frac{1}{G_N^{\rm ren}} = \frac{1}{G_N} + \frac{\delta}{12\pi}. \tag{6}$$

Furthermore, the second line of the expression (4) shows how, due to covariantization of the action, the wave function renormalization of the graviton depicted by the Feynman diagram of Fig. 1 forces the introduction of the counterterms involving the "R^2" terms.

The renormalized values of Newton's constant and the cosmological constant $G_N^{\rm ren}$ and $\Lambda^{\rm ren}$ are not calculable but are completely arbitrary. They are inputs to the theory which need to be *measured*.[4] One needs to put in two renormalization conditions, which specify the values of $G_N^{\rm ren}$ and $\Lambda^{\rm ren}$ at the subtraction point. This implies that the renormalized quantities depend on the subtraction point scale in the same way as the regulator masses. Taking the subtraction point to be the field theory UV cutoff implies that the renormalized Newton's constant depends on it in the same way as on the regulator masses, $1/G_N \sim \mathcal{O}(1)N\mathcal{M}_{\rm UV}^2$, since $m_i \sim \mathcal{M}_{\rm UV}$ for $i = 1, \ldots, 5$. Therefore, the renormalized Planck scale is

$$M_{pl}^2 = \mathcal{O}(1)N\mathcal{M}_{\rm UV}^2 + M_{pl,\rm bare}^2, \tag{7}$$

where the last term includes any additional contributions from the gravitational sector, additional UV degrees of freedom, and finite IR corrections.

When fermions are included, similar conclusions apply. Indeed, [37,56,57] point out that upon integrating out N_0 minimally coupled scalars and $N_{1/2}$ fermions above the cutoff Λ, the one-loop effective action for the gravitation field has the form

$$S_{1-\rm loop} \sim \int d^4x\sqrt{g}(M_{pl,\rm bare}^2 + c_1 N_0 \mathcal{M}_{\rm UV}^2 + c_2 N_{1/2}\mathcal{M}_{\rm UV}^2)R, \tag{8}$$

where c_1, c_2 have the same sign. The renormalized Planck mass is thus

$$M_{pl}^2 = M_{pl,\rm bare}^2 + (c_1 N_0 + c_2 N_{1/2})\mathcal{M}_{\rm UV}^2. \tag{9}$$

In [37,58,59], it was shown that gauge fields and nonminimally coupled scalars contribute negative shifts to the renormalization of Newton's constant. Thus, one may object that in some theories Newton's constant may receive a small renormalization due to cancellations between different fields running in the loops, and our resolution of the species problem does not apply. However, if we compute the entanglement entropy via the replica trick [38], the divergences in that calculation are nonetheless precisely taken care of by the renormalization of the gravitational action: see [39,40,45] and the references therein. The interpretation of the contribution of gauge fields is an active subject of research [41,42,44–47]. Nonetheless, it appears to be consistent to take the entanglement entropy as computed by the replica trick to account for the contribution of light fields to the gravitational entropy [39,40]. If we do so there is still no species problem.

Furthermore, Newton's constant is not the only term in the gravitational action to get renormalized: the loop contributions to the (curvature)^2 will involve different combinations of the effects of different species, without cancellations.

The upshot of this is that any truncated effective theory of quantum field theory (QFT) coupled to gravity, with higher dimension irrelevant operators constructed from geometric invariants, is strongly coupled beyond $\mathcal{M}_{\rm UV}$. The action (4) already shows this, since it contains at least a spin-2 massive ghost, with a mass $m_{\rm ghost} \simeq M_{pl}/\sqrt{c_{\rm ren}} \simeq M_{pl}/\sqrt{N}$, where[5] $c_{\rm ren} = c + \frac{\alpha}{1440\pi}$ as in (4) [33]. The ghost will generically remain present at any finite loop order of the expansion, and without a full UV completion it is impossible to determine if the theory can be extended above this scale. This has been noted previously in the cosmological context in [60–62].

In summary, to consistently do semiclassical 4d gravity calculations, one must restrict the theory to scales below the physical cutoff

$$\mathcal{M}_{\rm UV} \lesssim \frac{M_{pl}}{\sqrt{N}}. \tag{10}$$

If M_{pl} is fixed by classical gravitational measurements and N is increased, the cutoff of the QFT *must* in general be correspondingly lowered. The inequality $M_{pl}^2 \geq (c_1 N_0 + c_2 N_{1/2})\mathcal{M}_{\rm UV}^2$, implied by (9), will be saturated when $N \gg 1$, yielding the scaling $M_{pl}^2 = N\mathcal{M}_{\rm UV}^2$. This occurs, for example, in RS2 braneworlds [63,64] and in induced gravity [50,65].

[4] See [53–55] for a discussion of the measurement subtleties.

[5] Here we ignore the numerical factors in the renormalized value of c because they can be compensated by the logs in realistic models with very light particles.

B. Black hole entropy

The species problem appears in considerations of black hole entropy when one tries to include the effects of large numbers of matter fields. There are two apparently different approaches which however yield the same answer (see [39,40] for up-to-date versions of the relevant arguments, and for surveys of prior work).

One approach is to compute the free energy for a black hole as a function of the temperature, via computing fluctuations about the Euclidean saddle point, take the appropriate derivatives with respect to temperature. This gives a thermal entropy whose classical contribution is the Hawking-Bekenstein entropy.

The other approach is to interpret the black hole entropy as an entanglement entropy. Then the one-loop contributions from the matter fields can be computed following the prescription of [38,66,67]. Technically, they involve changing the temperature without changing the horizon radius, introducing a conical deficit angle into the spacetime. However, the result for the entropy is the same as the saddle point approach given above. In four dimensions, the calculation involves divergences which are quadratic in the cutoff and which scale with the number of species. Such divergences are absorbed precisely by the renormalization of the gravitational effective action. The resulting entropy will be the Hawking-Bekenstein or Wald entropy for the black hole, with the renormalized Newton's constant that is the correct physical gravitational coupling at low energies. The species problem never appears so long as one writes the Bekenstein-Hawking entropy in terms of physical couplings. The calculation also confirms that the appropriate cutoff is precisely the formula (10) [68].

The arguments reviewed above agree with the following qualitative picture of the black hole from the point of view of a static Schwarzschild observer. Static observers a proper distance ϵ from the black hole see a local Unruh temperature of order $T_u = 1/\epsilon$. One can consider the region within a distance ϵ from the horizon to be a thermal membrane or "stretched horizon" [69] with temperature T_u. For N species lighter than T_u, the thermal entropy of this membrane is $S = NT_u^3 V$ where $V = \epsilon A$ is the volume of the stretched horizon, A is the area of the black hole. Setting $\epsilon = 1/\mathcal{M}_{\rm UV}$, we find $S = N\mathcal{M}_{\rm UV}^2 A$ which parametrically matches the Hawking-Bekenstein entropy when $\mathcal{M}_{\rm UV} = M_4/\sqrt{N}$ [70–73]. Increasing the cutoff further takes us out of the range of 4d semiclassical gravity, and requires knowledge of the UV completion.

The precise details of the analogous calculations for the gravitational entropy of cosmological horizons are still not available, although there are arguments that the considerations above should extend to such cases [71,74]. Furthermore, the qualitative discussion of the paragraph above applies directly to the de Sitter horizon as seen by a static observer.

Since the work of [75], it has been clear that the Bekenstein-Hawking entropy for black holes in string theory should be considered as a statistical entropy, counting the density of states as a function of the energy. Recent advances in our understanding of the emergence of thermodynamics in closed quantum systems [76–79] have shown that the entropy of a thermal system can be understood precisely as an entanglement entropy between observed and unobserved factors of the Hilbert space. It would be interesting to explore the equivalence between the semiclassical gravity calculations of each, in this light. It may also help shed light on the nature of de Sitter entropy.

C. Some examples of the strong coupling scale

We close this section by noting previous concrete examples of a strong coupling scale for gravity below the Planck scale. A very simple example [57,80] is a D-dimensional theory (say, string theory) with a fundamental Planck scale $M_D < M_4$, compactified on a $d = D - 4$-dimensional manifold with volume $V = L_{KK}^d$. The number of Kaluza-Klein (KK) modes between L_{KK} and M_D is, roughly, $(L_{KK}M_D)^d$; the strong coupling scale is then

$$\mathcal{M}_{\rm UV}^2 = \frac{M_4^2}{(L_{KK}M_D)^d} = \frac{L_{KK}^d M_{KK}^{d-2}}{(L_{KK}M_D)^d} = M_D^2 \qquad (11)$$

so 4d semiclassical gravity breaks down at the fundamental D-dimensional Planck scale M_D.

We can also consider a general case in which N_3 D3-branes are close together within some 6d compactification with volume $V \sim R_{KK}^6$; this corresponds to the UV completion of the previously mentioned Randall-Sundrum (RS) braneworld scenario [63,81]. In this case, $N_3^2 = (R_{\rm ads}/\ell_{p,10})^8$ is the number of light species, where $R_{\rm ads}$ is the scale of warping near the D3-branes; deep in the core of the D3-brane geometry, it is the curvature radius of the associated anti-de Sitter background. The strong coupling scale for 4d gravity is

$$\mathcal{M}_{\rm UV}^2 = \frac{m_{pl}^2}{N_3^2} \sim \frac{R_{KK}^6}{R_{\rm ads}^8}. \qquad (12)$$

Now, if $R_{\rm ads} < R_{KK}$, so that the D3-brane throat is smaller than the KK scale, $\mathcal{M}_{\rm UV}^2 \gg 1/R_{\rm ads}^2$, and the theory becomes effectively five-dimensional at scales below $\mathcal{M}_{\rm UV}^2$. As we increase the number of D3-branes, we have a strongly warped compactification that is well described by a RS braneworld scenario [64,82]. The 4d Planck scale is $\delta m_{pl}^2 = m_{pl,5}^3 R_{\rm ads}$, where $m_{pl,5}$ is the 5d Planck scale; the central charge of the dual CFT is $c \propto (R m_{pl,5})^3$. This implies a UV scale $\mathcal{M}_{\rm UV} \sim \frac{m_{pl}}{\sqrt{c}} \sim 1/R_{\rm ads}$. Thus the theory becomes effectively five-dimensional at $\mathcal{M}_{\rm UV}$.

III. RENORMALIZATION OF DE SITTER ENTROPY

To determine the gravitational entropy of N massive (light) scalars in de Sitter space we work in the "static patch" of $3+1$-dimensional de Sitter space, with the metric

$$ds^2 = -g(r)d\tau^2 + \frac{1}{g(r)}dr^2 + r^2 d\Omega_3^2,$$
$$g(r) = \left(1 - \frac{r^2}{r_H^2}\right), \tag{13}$$

where $H^2 = 1/r_H^2$ is the de Sitter Hubble constant. The section of the geometry $r \leq r_H$ is the causal patch of a single observer at $r = 0$, and r_H is the location of her event horizon. One can define a Hamiltonian as the infinitesimal generator of translations in static time τ. Let $S(\beta)$ be the thermal entropy at the Gibbons-Hawking temperature $T = 1/\beta = H/2\pi$ computed in the canonical ensemble defined with respect to this static patch Hamiltonian. The covariant entropy bound states that $S(r_H) = A/4G_N$.

We want to determine the contribution of N scalars to this entropy.[6] The blueshift near the horizon implies that a large number of modes are concentrated there, and leads to a divergence. To deal with these and relate them to the renormalization of the gravitational effective action, we follow the strategy of [52] which ensures that regularization of the entropy and the gravitational effective action are done in the same scheme [58]. First, we impose "brick wall" boundary conditions (i.e., Dirichlet boundary conditions) on all scalar fields at a small but finite distance from the horizon. This surface acts not only as a position-space regulator, but also a momentum space regulator, by isolating the leading quadratic divergence coded by the blueshift in the near horizon limit. We use the renormalization prescription for the one-loop gravitational effective action discussed above to renormalize the entropy. The resulting contribution to the gravitational entropy precisely matches the renormalization of Newton's constant, extending the Bekenstein-Gibbons-Hawking formula to one loop in de Sitter space. We close with some comments regarding the relationship to entanglement entropy.

A. Renormalizing the entropy of N scalar fields

Consider a free, massive scalar field in de Sitter space. To regulate the theory in this background, we impose the "brick-wall" boundary condition

$$\Phi = 0 \quad \text{at } r = r_H - \epsilon. \tag{14}$$

Here ϵ is the coordinate distance of the brick wall regulator from the horizon $r_H = 1/H$. An an infrared cutoff is not necessary since the static patch of de Sitter is finite. We now compute the free energy of this scalar at a temperature T, which in the end we will set to be the Gibbons-Hawking temperature $T_{\rm GH} = H/2\pi$.

We next determine the mode expansion for the energy levels $E(n, l, l_3)$ of the field Φ. Here l is the total angular momentum and l_3 the angular momentum along some fixed axis. The field equation for the modes with energy E and angular momentum quantum numbers l, l_3 are $\Phi = e^{-iE\tau}Y_l^{l_3}(\theta,\varphi)\phi(r)$, where the radial modes obey

$$\frac{1}{r^2}\partial_r(r^2 g(r)\partial_r\phi) + \left(\frac{1}{g(r)}E^2 - m^2 - \frac{l(l+1)}{r^2}\right)\phi = 0. \tag{15}$$

Close to the horizon, the energy blueshift as $g(r) \to 0$ guarantees that the WKB approximation $\exp(\pm i\int k(r)dr)$ where

$$k^2(r, l, E) = \frac{1}{g^2(r)}E^2 - \frac{1}{g(r)}\left(m^2 + \frac{l(l+1)}{r^2}\right), \tag{16}$$

will give a good accounting of the behavior and multiplicity of modes with a given energy, $\pi n = \int_L^{r_H-\epsilon} dr k(r, l, E)$. Moreover, the blueshift also guarantees that this region gives the dominant contribution to the entropy which diverges in the near horizon limit. Hence, the leading order contributions to the entropy will come from precisely the modes in this regime. Within this approximation, the number of states up to energy E is

$$\rho(E) = \frac{1}{\pi}\int_0^{r_H-\epsilon} dr \int_0^{l_{\max}(E)} dl(2l+1) \times \frac{\sqrt{E^2 - g(r)\left(m^2 + \frac{l(l+1)}{r^2}\right)}}{g(r)}, \tag{17}$$

where $l_{\max}$ is the value at which the argument of the square root vanishes, and the summation of angular momenta l is approximated by an integral, which is valid for $l \gg 1$ near the horizon.

Now, for N identical scalars, the free energy at inverse temperature β is given by

$$e^{-\beta F} = \prod_{n,l,l_3}\frac{1}{(1 - e^{-\beta E(n,l,l_3)})^N}. \tag{18}$$

Hence, $\beta F = N\int dE\frac{d\rho(E)}{dE}\ln(1-\exp(-\beta E)) = -\beta N\int dE\rho(E)/(e^{\beta E}-1)$ after integration by parts. Further, following [83], since the dominant contributions will come from the highest energy modes, which have large l, their density of states and total number of modes with a given energy behave as $l(l+1) \sim l^2$, $(2l+1) \sim 2l$. Integrating over l in (17) then gives

$$F = -\frac{2N}{3\pi}\int_0^\infty dE\frac{1}{e^{\beta E} - 1} \times \int_0^{r_H-\epsilon} dr\frac{r^2}{g^2(r)}(E^2 - g(r)m^2)^{3/2}. \tag{19}$$

[6] A similar calculation was done in [74] for $2+1$-dimensional de Sitter space.

Rewriting $g(r) = -\frac{2}{r_H}(r-r_H) - \frac{1}{r_H^2}(r-r_H)^2$ to extract the divergences in the limit $\epsilon \to 0$ we find

$$\begin{aligned} F &= -\frac{2N}{3\pi}\int_0^\infty dE \frac{E^3}{e^{\beta E}-1}\frac{r_H^4}{4} \\ &\quad \times \left[\frac{1}{\epsilon} + \left(\frac{1}{r_H} + \frac{3m^2}{r_H E^2}\right)\log(\epsilon/r_H) + \mathcal{O}(1)\right] \\ &= -\frac{N\pi^3}{90}\frac{r_H^4}{\beta^4}\frac{1}{\epsilon} - \frac{N\pi^3}{90}\frac{r_H^4}{\beta^4}\left(\frac{1}{r_H} + \frac{15}{2}\frac{\beta^2 m^2}{\pi^2 r_H}\right) \\ &\quad \times \log(\epsilon/r_H) + \text{finite terms}. \end{aligned} \tag{20}$$

The free energy has two divergences, an inverse power and a logarithmic one. These are the cutoff-dependent contributions which are subtracted off by the counterterms of the theory, defined by the regulators. In order to subtract these divergences using the prescription for the renormalization of the effective action (20), we must compute these quantities in the same scheme [58]. Since—as [52]—we are using the Pauli-Villars regulators, the total free energy for N scalar fields and the system of Pauli-Villars regulators for each of them is

$$\beta F = \beta \sum_{i=0}^{5} \Delta_i F^i \tag{21}$$

where F^i is (20) computed using the mass, m_i, of the ith species. Because the individual free energies are replicas of each other, their divergences will be the same as in (20). So when we extract them from the total free energy (21), the sum rules (3) imply that these terms vanish. Of course, this means that in the regulated theory the divergences reappear as the $m_i \to \infty$ divergences. These terms are renormalized by the counterterms in the effective action. In principle, we would have to compute the finite terms in (20) to identify them. However, here we can use a shortcut, noting that the counterterms are defined by taking the limits $m_i \to \infty$ *at the same rate*. Thus these divergences will behave in exactly the same way as the blueshift divergences which occur when we move the brick wall to the horizon. Since the blueshift formula yieds $E_{\text{blue}} = E/g^{1/2} = E\sqrt{r_H/(2\epsilon)}$, we can simply trade $m_i^2 \ln(\epsilon/r_H)$ for $m_i^2 \ln(m_i^2)$ in each contribution $\propto m_i^2 \ln(\epsilon/r_H)$ in the sum of F^i's. This yields the dominant contribution to the regulated free energy in the limit $m_i \to \infty$[7]

$$F = -\frac{N\pi}{12}\frac{r_H^3}{\beta^2}\sum_{i=0}^{5}\Delta_i m_i^2 \log m_i^2 + \ldots. \tag{22}$$

Given the free energy $F(\beta)$, $S = \beta^2 \partial F/\partial\beta$. So the leading divergent contribution to the entropy as the cutoff is taken to infinity is

$$S = \frac{N\pi}{6}\frac{r_H^3}{\beta}\sum_{i=0}^{5}\Delta_i m_i^2 \log m_i^2`. \tag{23}$$

This is the leading contribution of N species of particles and their Pauli-Villars regulators to the total entropy in de Sitter causal patch, coming predominantly from the modes which are accumulated near the horizon. Now, since this system is in equilibrium with the background, we set the temperature β^{-1} to the Bekenstein-Gibbons-Hawking temperature of de Sitter space $T_{GH} = H/2\pi$, or alternatively we use $r_H = 1/H = \beta/(2\pi)$. Recalling that the horizon area is $A = 4\pi r_H^2$,

$$S = \frac{N}{48\pi}A\sum_{i=0}^{5}\Delta_i m_i^2 \log m_i^2 = \frac{A}{48\pi}\delta, \tag{24}$$

where we have employed the definition of the counterterm δ from (5). When we add this UV contribution to the bare Bekenstein-Gibbons-Hawking entropy of de Sitter, we obtain simply the finite renormalized entropy

$$S^{\text{ren}} = S_{dS} + S = \frac{A}{4G_N} + \frac{A}{48\pi}\delta = \frac{A}{4G_N^{\text{ren}}}. \tag{25}$$

The divergences match: the leading order N dependence precisely cancels. So the species problem never appears when the de Sitter entropy is correctly calculated using the physical renormalized Newton's constant.

We note that this conclusion is expected to remain correct even if the entropy is calculated as the entanglement entropy (from fields with spins < 2). In the case of black holes, the contribution of background fields to the entanglement entropy and Gibbons-Hawking free energy match precisely. A similar argument holds for quantum fields on de Sitter backgrounds [40,84].[8]

IV. EFFECTIVE FIELD THEORY AND INFLATION

The discussion in the previous sections sets the stage for the analysis of inflation in theories with many light matter species. Only after we have renormalized the gravitational sector of the theory, as well as the standard local QFT matter sector, can we consider the question of corrections to

[7]There is also a purely logarithmic divergence coming from the first logarithmic term in (20). We expect that this should match the renormalization of the (curvature)^2 couplings as in [52], if we extend the Gibbons-Hawking entropy to the Wald entropy, but we will not pursue that here.

[8]Calculations of the de Sitter entropy in flat slicing can be found in [85–87].

the scalar and tensor power spectrum of CMB fluctuations, the inflaton sector dynamics and the viability of long inflation against both perturbative and nonperturbative corrections from both field theory and quantum gravity.

With these points in mind, we will review the general aspects of large field inflation driven by axions, and the conditions and reasons for its viability as a QFT. We will then revisit some of the explicit arguments in [30] regarding nonperturbative quantum gravity effects, and reconsider their implications for the correctly renormalized low energy theory.

A. Inflation with many species

The standard picture of inflation and its main observable prediction, the CMB fluctuations, rely on the validity of semiclassical 4d gravity at the Hubble scale H, which is the curvature scale of the background during the inflationary epoch. In light of the discussions above, it is clear that $M_{pl}^{\rm ren}/\sqrt{N} \geq \mathcal{M}_{\rm UV} \gg H$ is required for this picture to be valid. This is manifest in calculations of loop corrections to the scalar and tensor power spectrum. The metric couples to matter fields with Planck-suppressed couplings, so a 4-d calculation of the density fluctuations (assuming the locally Lorentzian vacuum, aka the Bunch-Davies vacuum, for the inflaton and graviton fluctuations) will produce a spectrum of perturbations [88–91][9]

$$P = P_{\rm tree}\left(1 + cN\frac{H^2}{M_{pl}^2} + \ldots\right) = P_{\rm tree}\left(1 + c'\frac{H^2}{\mathcal{M}_{\rm UV}^2} + \ldots\right) \tag{26}$$

for both scalar and tensor modes, where $c' \ll 1$ if $\mathcal{M}_{\rm UV} \ll M_4/\sqrt{N}$. For example, the N species could be Kaluza-Klein modes; the resulting UV scale is the 10d Planck scale [88]. This picture arises from a general effective field theory analysis of the inflaton-graviton sector [88]: higher powers of H come from terms in the effective action that are of higher power in the curvature, are dictated by the graviton wave function renormalization, and will be suppressed some UV scale $\mathcal{M}_{\rm UV}$, which plays the role of a cutoff of the low energy theory. It is clear that the corrections are only small if $\mathcal{M}_{\rm UV} > H$, which is at any rate required for the validity of semiclassical gravity at the scale H. For inflation with the minimal required number of efoldings, there is also the question of initial states which deviate from the Bunch-Davies vacuum, which we will ignore in what follows. As long as the inflationary dynamics obeys the standard rules of EFT, these deviations are limited [92]. The main physical observables do not depend significantly on the number of light species to the leading order, because they are automatically expressed in terms of the renormalized 4d physical quantities.

[9]Up to logarithmic corrections [89–91].

B. Axions: inflation and N-flation

1. Axions as inflatons

Axions are perfect candidates for inflatons: the periodicity of the axion $\phi \equiv \phi + 2\pi f_a$ protects the potential from perturbative corrections, allowing for a relatively shallow potential. In the cases that the potential is generated by a dilute gas of instantons and takes the form $V \sim \Lambda^4 \cos(\phi/f_a)$, inflation requires a fairly large value of $f_a \sim M_{pl}$, to support a long and uninterrupted slow roll regime that can sustain at least ~60 efolds of inflation [5]. There has been much work on constructing axion inflation models, and we will not review that work here. Instead, we will focus on the aspects of axion-driven inflation, with a quasi-de Sitter geometry, relevant for understanding possible entropy bounds.

As noted above, to ask questions about the validity of entropy bounds in a quasi-de Sitter space, we must first determine the regime of validity of the semiclassical theory. Reference [30] argues that the proper cutoff at which to evaluate the de Sitter entropy in a theory with an axion is the axion decay constant f_a. The argument is that the composite operator $\phi^2(0)$, evaluated with a momentum cutoff $\mathcal{M}_{\rm UV}$, scales as $\mathcal{M}_{\rm UV}^2$. So, the argument goes, when $\mathcal{M}_{\rm UV} \sim f_a$, the fluctuations in the scalar field completely delocalize it on the circle $\phi \equiv \phi + 2\pi f_a$, preventing the semiclassical description of the scalar as a rolling in a single perturbative sector, and smearing it over the full covering space. An alternative reading is that the two-point function $\langle\phi(x)\phi(y)\rangle \sim 1/|x-y|^2$ at short distances, and for separations $|x-y| \sim f_a^{-1}$ the fluctuations between points cover the entire target space circle.

As we will now explain, this argument is not correct. In fact, the cutoff $\mathcal{M}_{\rm UV}$ can be either larger or small than f_a without leading to any inconsistency. First of all, if the cutoff $\mathcal{M}_{\rm UV}$ is *smaller* than the period f_a, fluctuations at the cutoff would obviously do little in the way of smearing the expectation value of the axion over the scale f_a. This might be countered by claiming that the cutoff should be close to M_{pl}. Yet, as we have seen previously, this is not the case in many models of interest. Second, the estimate of the scale of $\phi^2(0)$ in [30] ignores the renormalization of this composite operator. In fact the estimate $\langle\phi^2(0)\rangle \sim \mathcal{M}_{\rm UV}^2$ is really the regularized quadratic divergence of the cosmological constant term in de Sitter space with a massive scalar field, and it will be subtracted off in the correct renormalization procedure. Indeed, in the flat space vacuum, $\langle\phi^2(0)\rangle = 0$ after properly renormalizing the operator. In de Sitter space, there is an IR contribution only to the renormalized operator, and $\langle\phi^2(0)\rangle \sim H^2$.

The alternate point that $\langle\phi(x)\phi(y)\rangle \sim 1/|x-y|^2$ leads to the axion being delocalized at the scale $1/f_a$ is true. However, the correct interpretation of this phenomenon is that f_a is the strong coupling scale for the *axion* dynamics.

The axion potential typically arises from instantons which couple to the axion via an irrelevant operator $\sim\frac{\phi}{f_a}F\wedge F$. Periodicity of the axion guarantees that any direct dependence on ϕ (as opposed to its derivatives) must be a periodic function of ϕ/f, ie a harmonic series in $\cos\phi/f_a$. If we write a low-energy effective field theory by expanding this about a minimum, the expansion will be in powers of ϕ/f_a, thus indicating that f_a is a natural scale for strong coupling, beyond which the potential cannot be approximated by the first few terms in the expansion.

There is no reason for the axion sector strong coupling scale to be the one at which 4d semiclassical gravity breaks down. For example, for the cases where $f_a<\mathcal{M}_{\rm UV}$, the UV completion of the axion sector can be described fully in the four-dimensional EFT below $\mathcal{M}_{\rm UV}$. The axion can be UV-completed as a phase of a Peccei-Quinn complex doublet $\Phi=\varphi e^{i\theta}$, with potential $V(\Phi)=\lambda(|\Phi|^2-f_a^2)^2$. After symmetry breaking and integrating out the heavy radial mode, the axion decay constant is the Peccei-Quinn *vev* f_a. The mass of the radial mode is $m_\varphi\sim\sqrt{\lambda}f_a$, and as long as the theory is weakly coupled, $\lambda\ll 1$, we have $m_\varphi\ll f_a$. The cutoff of the low energy theory with only the phase retained is $\sim m_\varphi$, where the low energy theory of the axion with interactions governed by higher dimension operators, generated after integrating the radial mode, becomes strongly coupled and violates unitarity. To resolve this, all one needs is to integrate the radial mode back in at scales above m_φ. This happens entirely within the realm of the EFT with gravity below $\mathcal{M}_{\rm UV}$. In more complicated cases, as in string theory compactifications, the UV completion will depend on the details of moduli stabilization. Of course one should understand this UV completion properly when accounting for the axion sector's contribution to the de Sitter entropy below the scale $\mathcal{M}_{\rm UV}$.

For axion decay constants near the Planck scale, four-dimensional gravity typically breaks down at scales well *below* f_a, as also noted in [10]. In particular, if the Kaluza-Klein scale is below f_a, most string theory axions lift in 10 or 11 dimensions to a higher-form gauge field at this scale. The worry arises if one requires $f_a>M_{pl}$ as in the early models of axion inflation [5]. The QFT sector of such models appears to behave without a problem. However nonperturbative gravity effects—as exemplified by wormhole calculations of the corrections to the low energy actions—may be very dangerous for such models [3,93–96]. Moreover, the WGC is in tension with elementary axion theories with such large f_a [3,9–11,97,98].[10]

However it is possible to realize low energy axion models with a very large effective f_a, above the actual strong coupling scale of the theory (whether the strong coupling dynamics is from local QFT degrees of freedom, or from quantum gravity). Examples are provided by various realizations of axion monodromy. We provide a specific example here for illustrative purposes, inspired by [34,35]. The purpose is not to build a complete model of inflation, but to illustrate how to generate a hierarchy between an effective f_a and the actual strong coupling scale $\mathcal{M}_{\rm UV}\ll f_a$, within field theory.

Consider a simple case involving two axions, coupling via topological terms to three different gauge groups, in nonorthogonal linear combinations:

$$\mathcal{L}_{\rm int}=\frac{\phi_1}{f_1}{\rm tr}F_1\wedge F_1+\frac{\phi_2}{f_2}{\rm tr}F_2\wedge F_2+\left(\frac{\phi_1}{f_1}-\frac{n\phi_2}{f_2}\right){\rm tr}F_3\wedge F_3, \tag{27}$$

where n is an integer. Provided that all the gauge sectors are weakly coupled just below the cutoff, we can calculate the instanton potential generated by the gauge theories in the dilute gas approximation (when it applies), and find to leading order

$$V_{\rm eff}=\mu_1^4\cos\left(\frac{\phi_1}{f_1}\right)+\mu_2^4\cos\left(\frac{\phi_2}{f_2}\right)+\mu_3^4\cos\left(\frac{\phi_1}{f_1}-\frac{n\phi_2}{f_2}\right). \tag{28}$$

Let the axion decay constants be comparable, $f_1\sim f_2<\mathcal{M}_{\rm UV}$, but let there be a hierarchy between the gauge sector strong coupling scales $\mu_1\ll\mu_2\ll\mu_3$. This can be arranged by a choice of the fermionic charges in the theory, gauge groups, and their coupling constants.

To understand the perturbative behavior of the theory, pick a particular vacuum of the theory, say $\phi_1=\phi_2=0$, and consider small fluctuations. The potential (28) yields the mass matrix of the small fluctuations,

$$V_{\rm masses}=\frac{\mu_1^4}{2f_1^2}\phi_1^2+\frac{\mu_2^4}{2f_2^2}\phi_2^2+\frac{n^2\mu_3^4}{2f_2^2}\left(\phi_2-\frac{f_2}{nf_1}\phi_1\right)^2. \tag{29}$$

Given the scale hierarchy, (29) shows that the heaviest field in the theory is really the linear combination $\chi_{\rm heavy}\propto\phi_2-\frac{f_2}{nf_1}\phi_1$, which is mostly ϕ_2, with a small admixture of ϕ_1. So to understand the low energy dynamics, we can pick it as one of the two normal modes of the system, and choose the direction orthogonal to it as the other. Picking canonical normalizations for these fields yields

$$\chi_{\rm heavy}=\frac{f_1f_2}{f_{\rm eff}}\left(\frac{n\phi_2}{f_2}-\frac{\phi_1}{f_1}\right),\qquad \chi_{\rm light}=\frac{f_1f_2}{f_{\rm eff}}\left(\frac{\phi_2}{f_1}+\frac{n\phi_1}{f_2}\right), \tag{30}$$

[10]Reference [99], on the other hand, claims these are not problematic in principle, but that there are problems with using them for inflation in specific string models.

where $f_{\rm eff} = \sqrt{n^2 f_1^2 + f_2^2}$. Substituting these field redefinitions into the potential (28) yields

$$V_{\rm eff} = \mu_1^4 \cos\left(\frac{n\chi_{\rm light}}{f_{\rm eff}} - \frac{f_2\chi_{\rm heavy}}{f_1 f_{\rm eff}}\right) + \mu_2^4 \cos\left(\frac{nf_1\chi_{\rm heavy}}{f_2 f_{\rm eff}} + \frac{\chi_{\rm light}}{f_{\rm eff}}\right) + \mu_3^4 \cos\left(\frac{f_{\rm eff}\chi_{\rm heavy}}{f_1 f_2}\right) \tag{31}$$

with canonically normalized kinetic terms.

The hierarchy of strong coupling scales means that the last term strongly localizes $\chi_{\rm heavy}$ in a minimum of the final term in (31). The first term gives a small sinusoidal modulation of the second term, which is a cosine potential for $\chi_{\rm light}$ with periodicity $f_{\rm eff} \sim nf_1$. The trajectory of this field in the coordinates ϕ_1, ϕ_2 can be seen in Fig. 2; we have produced a form of axion monodromy.

Let us study the dynamics of $\chi_{\rm light}$ in more detail, and understand when the low-energy effective action becomes strongly coupled. Vacua of the theory correspond to nonzero $\chi_{\rm light}, \chi_{\rm heavy}$: the arguments of the cosines must be odd integer multiples of π. Specifically, pick a vacuum $\chi_{\rm heavy} = (2l+1)\frac{f_1 f_2}{f_{\rm eff}}\pi$ for the heavy degree of freedom. Expanding the potential (31) about it, and taking $n \gg \mu_2^2/\mu_1^2$ the effective potential for $\chi_{\rm light}$ becomes

$$V_{\rm eff} \simeq \frac{\mu_2^4}{2f_{\rm eff}^2}\left(\chi_{\rm light} + (2l+1)\frac{nf_1^2}{f_{\rm eff}}\pi\right)^2 + \mu_1^4 \cos\left(\frac{n\chi_{\rm light}}{f_{\rm eff}} - (2l+1)\frac{f_2}{f_{\rm eff}}\pi\right). \tag{32}$$

We do not expand the second term: the frequency of this harmonic is much larger than that in the first term, by $n \gg \mu_2^2/\mu_1^2$; and the phase shift is small, $\sim f_2/f_{\rm eff} \sim 1/n$, for many vacua in the theory. The cosine is merely a harmonic modulation on top of the first term as long as the field $\chi_{\rm light}$ is far away from the vacuum, which is approximately at $\simeq (2l+1)\frac{nf_1^2}{f_{\rm eff}}\pi \simeq (2l+1)f_1$. The potential

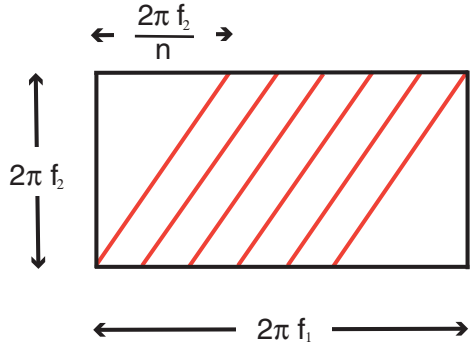

FIG. 2. A simple axion monodromy from two axions.

energy stored in this vacuum displacement is quite small, on the order of $V_{\rm eff} \lesssim (2l+1)^2\mu_2^4$, safely in the regime of semiclassical gravity as long as $\sqrt{2l+1}\mu_2 \lesssim M_{pl}$.

The field displacement of $\chi_{\rm light}$ from its vacuum can be larger than M_{pl} even when f_1 is safely below the Planck scale, if $l \gtrsim M_{pl}/f_1$, and the low energy action (32) can still remain. This situation is depicted in Fig. 2, where the slanted lines denote the trajectory of the light field $\chi_{\rm light}$, and the field region between different segments belongs to field configurations with energies much larger than those along the slanted lines.

The cutoff limiting the regime of validity of the single light axion $\chi_{\rm light}$ low energy theory is given by the mass of the heavy field which has been integrated out, $\mathcal{M}^{\rm eff} \simeq \frac{n\mu_3^2}{f_2} \sim \frac{n^2\mu_3^2}{f_{\rm eff}}$. This can be easily arranged to be smaller than the apparent strong coupling scale $f_{\rm eff}$ of the dynamics for $\chi_{\rm light}$, by choosing $\mu_3 \lesssim f_{\rm eff}/n \sim f_1$. Thus the theory can be fully constructed in the regime where the standard perturbation theory operates, with low energy shift symmetry protecting the flatness of the effective potential, and the phase misalignment between the two axions simulating the trans-Planckian field displacements required for supporting long inflation. The nonperturbative corrections from gauge theory, and even wormhole induced terms from quantum gravity remain small. Finally, in light of the discussion beforehand, when the theory is correctly renormalized it also automatically obeys the covariant entropy bounds. More general monodromy models work in a similar way.

2. N-flation

N-flation [8] proposes to achieve effectively super-Planckian inflaton range from a large number of axions with sub-Planckian decay constants. The underlying assumption is that there are many axions which are displaced from their potential minima and are light. The total energy that drives inflation comes from the sum of the individual energies for each field, and this is what supports the slow roll of each individual field. So for N axions rolling in unison, the effective inflaton range can scale as $\sqrt{N}f_a$, where f_a is a characteristic fundamental axion decay constant and N is the number of axions.

There is an active, ongoing discussion in the literature as to whether such a theory can be embedded in a good string theory model, and whether it is consistent with a properly interpreted version of the weak gravity conjecture: see for example [9,10,100–105]. This is an interesting question, and we will not address it here. Our point is merely that there is no obvious violation of the covariant entropy bounds in this example. Reference [8] already noted that a large number of species can run in loops and correct the bare value of Newton's constant. As argued above, 4d semiclassical gravity will break down at some scale $\mathcal{M}_{\rm UV} \leq M_4/\sqrt{N}$, and there is no intrinsic problem with the covariant entropy bound.

The real issue is the value of $\mathcal{M}_{\rm UV}$. In addition to requiring $\mathcal{M}_{\rm UV} > H$, there is also an open question for the calculability of the underlying string compactification if $\mathcal{M}_{\rm UV}$ is lower than all of the compactification scale, the string scale, and the 10d Planck scale. In this case, the 4d dynamics relevant for the compactification intrinsically requires understanding strongly coupled quantum gravity; this could, for example, complicate considerations of moduli stabilization.

C. Flux compactifications with large gauge groups

Reference [30] examines the one-Kähler-modulus KKLT model [106] and its generalization to racetrack models [107], when the nonperturbative potential for the Kahler modulus is generated by gaugino condensation on wrapped D7-branes, and claims that the species problem renders them inconsistent. As we have argued, it does not. However it is interesting to ask at what scale 4d semiclassical gravity is expected to break down, and what dynamics becomes relevant.

Let us consider the scenario described in [106]. The modulus $\sigma = (R_{KK}/\ell_{p,10})^4$ is the volume modulus; the gauge coupling is then $g^2_{\rm YM} = 4\pi/\sigma$. The superpotential generated by gaugino condensation is taken to be

$$W = W_0 + Ae^{-2\pi\sigma/N_7} \tag{33}$$

where N_7 is the rank of the D7-brane gauge group, and W_0 is the tree-level flux-induced term in the superpotential. If, following [30,106], we take $W_0 \ll 1$, in order to get a solution well described by classical 10d geometry. In this case, the Kahler modulus is parametrically

$$\sigma = \left(\frac{R_{KK}}{\ell_{p,10}}\right)^4 \sim \frac{N_7}{2\pi}|\ln W_0|. \tag{34}$$

Now using the fact that $m^2_{pl,4} \sim R^6_{KK}/\ell^8_{p,10}$, and taking the number of light species to be $\sim N_7^2$ (more precisely, we are assuming here that N_7 is the dominant contribution to this number), we find

$$\mathcal{M}^2_{\rm UV} = \frac{m^2_{p,4}}{N_7^2} \sim \frac{|\ln W_0|^2}{(2\pi R_{KK})^2} \tag{35}$$

in other words, for small W_0, the strong coupling scale is at or larger than the Kaluza-Klein scale of the string theory compactification. Thus, before (or when) this scale is reached, 4d semiclassical gravity has already broken down in a completely standard fashion.

While the potentials in [107] are more complicated, we may take the specific numbers used in Eqs. (3.11-3.13) of that paper and find once again that the strong coupling scale Λ is of the same order as the Kaluza-Klein scale. This is preserved under the rescalings of parameters in Sec. 3.3 of that paper.

ACKNOWLEDGMENTS

Part of this work was carried out while A. L. was at the Aspen Center for Physics, which is supported by National Science Foundation Grant No. PHY-1066293. A. L. would like to thank the organizers of and participants in the "Primordial Physics" workshop for stimulating conversations on subjects related to this work. N. K. would like to thank CP^3-Origins, University of Southern Denmark, and CCPP, NYU, for kind hospitality in the course of this work. We would like to thank A. Arvanitaki, A. Brown, G. D'Amico, M. Porrati, and G. Zahariade for useful discussions. N. K. is supported in part by the DOE Grant No. DE-SC0009999. A. L. is supported by the DOE Grant No. DE-SC0009987. M. S. S. is supported by the Lundbeck Foundation. M. K. is supported in part by the NSF through Grant No. PHY-1214302, and he acknowledges membership at the NYU-ECNU Joint Physics Research Institute in Shanghai.

[1] D. H. Lyth, What Would We Learn by Detecting a Gravitational Wave Signal in the Cosmic Microwave Background Anisotropy?, Phys. Rev. Lett. **78**, 1861 (1997).
[2] G. Efstathiou and K. J. Mack, The Lyth bound revisited, J. Cosmol. Astropart. Phys. 05 (2005) 008.
[3] N. Arkani-Hamed, L. Motl, A. Nicolis, and C. Vafa, The string landscape, black holes and gravity as the weakest force, J. High Energy Phys. 06 (2007) 060.
[4] R. Bousso, A covariant entropy conjecture, J. High Energy Phys. 07 (1999) 004.
[5] K. Freese, J. A. Frieman, and A. V. Olinto, Natural Inflation with Pseudo-Nambu-Goldstone Bosons, Phys. Rev. Lett. **65**, 3233 (1990).
[6] N. Arkani-Hamed, H.-C. Cheng, P. Creminelli, and L. Randall, Pseudonatural inflation, J. Cosmol. Astropart. Phys. 07 (2003) 003.
[7] N. Arkani-Hamed, H.-C. Cheng, P. Creminelli, and L. Randall, Extranatural Inflation, Phys. Rev. Lett. **90**, 221302 (2003).
[8] S. Dimopoulos, S. Kachru, J. McGreevy, and J. G. Wacker, *N*-flation, J. Cosmol. Astropart. Phys. 08 (2008) 003.

[9] J. Brown, W. Cottrell, G. Shiu, and P. Soler, Fencing in the swampland: Quantum gravity constraints on large field inflation, J. High Energy Phys. 10 (2015) 023.
[10] B. Heidenreich, M. Reece, and T. Rudelius, Weak gravity strongly constrains large-field axion inflation, J. High Energy Phys. 12 (2015) 108.
[11] E. Palti, On natural inflation and moduli stabilisation in string theory, J. High Energy Phys. 10 (2015) 188.
[12] E. Silverstein and A. Westphal, Monodromy in the CMB: Gravity waves and string inflation, Phys. Rev. D **78**, 106003 (2008).
[13] L. McAllister, E. Silverstein, and A. Westphal, Gravity waves and linear inflation from axion monodromy, Phys. Rev. D **82**, 046003 (2010).
[14] N. Kaloper and L. Sorbo, A Natural Framework for Chaotic Inflation, Phys. Rev. Lett. **102**, 121301 (2009).
[15] N. Kaloper, A. Lawrence, and L. Sorbo, An ignoble approach to large field inflation, J. Cosmol. Astropart. Phys. 03 (2011) 023.
[16] G. D'Amico, R. Gobbetti, M. Kleban, and M. Schillo, Unwinding inflation, J. Cosmol. Astropart. Phys. 03 (2013) 004.
[17] G. D'Amico, R. Gobbetti, M. Schillo, and M. Kleban, Inflation from flux cascades, Phys. Lett. B **725**, 218 (2013).
[18] G. D'Amico, R. Gobbetti, M. Kleban, and M. Schillo, Large-scale anomalies from primordial dissipation, J. Cosmol. Astropart. Phys. 11 (2013) 013.
[19] E. Palti and T. Weigand, Towards large r from $[p, q]$-inflation, J. High Energy Phys. 04 (2014) 155.
[20] A. Hebecker, S. C. Kraus, and L. T. Witkowski, D7-brane chaotic inflation, Phys. Lett. B **737**, 16 (2014).
[21] L. E. Ibanez and I. Valenzuela, The inflaton as a MSSM Higgs and open string modulus monodromy inflation, Phys. Lett. B **736**, 226 (2014).
[22] L. E. Ibanez, F. Marchesano, and I. Valenzuela, Higgs-otic inflation and string theory, J. High Energy Phys. 01 (2015) 128.
[23] S. Bielleman, L. E. Ibanez, F. G. Pedro, and I. Valenzuela, Multifield dynamics in Higgs-otic inflation, arXiv:1505.00221.
[24] S. Bielleman, L. E. Ibanez, and I. Valenzuela, Minkowski 3-forms, flux string vacua, axion stability and naturalness, J. High Energy Phys. 12 (2015) 119.
[25] T. W. Grimm, Axion inflation in F-theory, Phys. Lett. B **739**, 201 (2014).
[26] R. Blumenhagen and E. Plauschinn, Towards universal axion inflation and reheating in string theory, Phys. Lett. B **736**, 482 (2014).
[27] R. Blumenhagen, D. Herschmann, and E. Plauschinn, The challenge of realizing F-term axion monodromy inflation in string theory, J. High Energy Phys. 01 (2015) 007.
[28] F. Marchesano, G. Shiu, and A. M. Uranga, F-term axion monodromy in inflation, J. High Energy Phys. 09 (2014) 184.
[29] X. Gao, T. Li, and P. Shukla, Combining universal and odd RR axions for aligned natural inflation, J. Cosmol. Astropart. Phys. 10 (2014) 048.
[30] J. P. Conlon, Quantum gravity constraints on inflation, J. Cosmol. Astropart. Phys. 09 (2012) 019.
[31] T. Banks and W. Fischler, An upper bound on the number of e-foldings, arXiv:astro-ph/0307459.
[32] N. Kaloper, M. Kleban, and L. Sorbo, Observational implications of cosmological event horizons, Phys. Lett. B **600**, 7 (2004).
[33] K. S. Stelle, Renormalization of higher derivative quantum gravity, Phys. Rev. D **16**, 953 (1977).
[34] J. E. Kim, H. P. Nilles, and M. Peloso, Completing natural inflation, J. Cosmol. Astropart. Phys. 01 (2005) 005.
[35] M. Berg, E. Pajer, and S. Sjors, Dante's inferno, Phys. Rev. D **81**, 103535 (2010).
[36] D. N. Kabat, Black hole entropy and entropy of entanglement, Nucl. Phys. **B453**, 281 (1995).
[37] F. Larsen and F. Wilczek, Renormalization of black hole entropy and of the gravitational coupling constant, Nucl. Phys. **B458**, 249 (1996).
[38] C. G. Callan, Jr. and F. Wilczek, On geometric entropy, Phys. Lett. B **333**, 55 (1994).
[39] T. Jacobson and A. Satz, Black hole entanglement entropy and the renormalization group, Phys. Rev. D **87**, 084047 (2013).
[40] J. H. Cooperman and M. A. Luty, Renormalization of entanglement entropy and the gravitational effective action, J. High Energy Phys. 12 (2014) 045.
[41] W. Donnelly, Decomposition of entanglement entropy in lattice gauge theory, Phys. Rev. D **85**, 085004 (2012).
[42] H. Casini, M. Huerta, and J. A. Rosabal, Remarks on entanglement entropy for gauge fields, Phys. Rev. D **89**, 085012 (2014).
[43] K.-W. Huang, Central charge and entangled gauge fields, Phys. Rev. D **92**, 025010 (2015).
[44] S. Ghosh, R. M. Soni, and S. P. Trivedi, On the entanglement entropy for gauge theories, J. High Energy Phys. 09 (2015) 069.
[45] S. N. Solodukhin, Newton constant, contact terms and entropy, Phys. Rev. D **91**, 084028 (2015).
[46] W. Donnelly and A. C. Wall, Geometric entropy and edge modes of the electromagnetic field, arXiv:1506.05792.
[47] R. M. Soni and S. P. Trivedi, Aspects of entanglement entropy for gauge theoriesarXiv:1510.07455.
[48] C.-T. Ma, Entanglement with centers, arXiv:1511.02671.
[49] G. 't Hooft and M. J. G. Veltman, One loop divergencies in the theory of gravitation, Annales Poincare Phys. Theor. A **20**, 69 (1974).
[50] S. L. Adler, Einstein gravity as a symmetry breaking effect in quantum field theory, Rev. Mod. Phys. **54**, 729 (1982); **55**, 837(E) (1983).
[51] D. V. Fursaev and S. N. Solodukhin, On one loop renormalization of black hole entropy, Phys. Lett. B **365**, 51 (1996).
[52] J.-G. Demers, R. Lafrance, and R. C. Myers, Black hole entropy without brick walls, Phys. Rev. D **52**, 2245 (1995).
[53] N. Kaloper and A. Padilla, Sequestering the Standard Model Vacuum Energy, Phys. Rev. Lett. **112**, 091304 (2014).
[54] N. Kaloper and A. Padilla, Vacuum energy sequestering: The framework and its cosmological consequences, Phys. Rev. D **90**, 084023 (2014); **90**, 109901 (2014).
[55] N. Kaloper, A. Padilla, D. Stefanyszyn, and G. Zahariade, A manifestly local theory of vacuum energy sequestering, arXiv:1505.01492.

[56] G. Veneziano, Large N bounds on, and compositeness limit of, gauge and gravitational interactions, J. High Energy Phys. 06 (2002) 051.
[57] G. Dvali and M. Redi, Black hole bound on the number of species and quantum gravity at LHC, Phys. Rev. D **77**, 045027 (2008).
[58] D. N. Kabat and M. J. Strassler, A comment on entropy and area, Phys. Lett. B **329**, 46 (1994).
[59] S. N. Solodukhin, One loop renormalization of black hole entropy due to nonminimally coupled matter, Phys. Rev. D **52**, 7046 (1995).
[60] J. Z. Simon, Higher derivative Lagrangians, nonlocality, problems and solutions, Phys. Rev. D **41**, 3720 (1990).
[61] J. Z. Simon, The stability of flat space, semiclassical gravity, and higher derivatives, Phys. Rev. D **43**, 3308 (1991).
[62] J. Z. Simon, No Starobinsky inflation from self-consistent semiclassical gravity, Phys. Rev. D **45**, 1953 (1992).
[63] L. Randall and R. Sundrum, An Alternative to Compactification, Phys. Rev. Lett. **83**, 4690 (1999).
[64] S. S. Gubser, AdS/CFT and gravity, Phys. Rev. D **63**, 084017 (2001).
[65] A. D. Sakharov, Spectral density of eigenvalues of the wave equation and the vacuum polarization, Theor. Math. Phys. **23**, 435 (1975); Sov. Phys. Usp. **34**, 395 (1991).
[66] L. Susskind, Some speculations about black hole entropy in string theory, arXiv:hep-th/9309145.
[67] S. N. Solodukhin, The conical singularity and quantum corrections to entropy of black hole, Phys. Rev. D **51**, 609 (1995).
[68] G. Dvali and S. N. Solodukhin, Black hole entropy and gravity cutoff, arXiv:0806.3976.
[69] L. Susskind, L. Thorlacius, and J. Uglum, The stretched horizon and black hole complementarity, Phys. Rev. D **48**, 3743 (1993).
[70] T. Jacobson, Black hole entropy and induced gravity, arXiv:gr-qc/9404039.
[71] S. Hawking, J. M. Maldacena, and A. Strominger, de Sitter entropy, quantum entanglement and AdS/CFT, J. High Energy Phys. 05 (2001) 001.
[72] R. Emparan, Black hole entropy as entanglement entropy: A holographic derivation, J. High Energy Phys. 06 (2006) 012.
[73] N. Kaloper, Cutoffs, stretched horizons and black hole radiators, Phys. Rev. D **86**, 104052 (2012).
[74] W. T. Kim, Entropy of $(2+1)$-dimensional de Sitter space in terms of brick wall method, Phys. Rev. D **59**, 047503 (1999).
[75] A. Strominger and C. Vafa, Microscopic origin of the Bekenstein-Hawking entropy, Phys. Lett. B **379**, 99 (1996).
[76] J. M. Deutsch, Quantum statistical mechanics in a closed system, Phys. Rev. A **43**, 2046 (1991).
[77] M. Srednicki, Chaos and quantum thermalization, Phys. Rev. E **50**, 888 (1994).
[78] H. Tasaki, From Quantum Dynamics to the Canonical Distribution: General Picture and a Rigorous Example, Phys. Rev. Lett. **80**, 1373 (1998).
[79] M. Rigol, V. Dunjko, and M. Olshanii, Thermalization and its mechanism for generic isolated quantum systems, Nature (London) **452**, 854 (2008).
[80] G. Dvali, Black holes and large N species solution to the hierarchy problem, Fortsch. Phys. **58**, 528 (2010).
[81] L. Randall and R. Sundrum, A Large Mass Hierarchy from a Small Extra Dimension, Phys. Rev. Lett. **83**, 3370 (1999).
[82] H. L. Verlinde, Holography and compactification, Nucl. Phys. **B580**, 264 (2000).
[83] G. 't Hooft, On the quantum structure of a black hole, Nucl. Phys. **B256**, 727 (1985).
[84] O. Ben-Ami, D. Carmi, and M. Smolkin, Renormalization group flow of entanglement entropy on spheres, J. High Energy Phys. 08 (2015) 048.
[85] R. Muller and C. O. Lousto, Entanglement entropy in curved space-times with event horizons, Phys. Rev. D **52**, 4512 (1995).
[86] E. Keski-Vakkuri and M. S. Sloth, Holographic bounds on the UV cutoff scale in inflationary cosmology, J. Cosmol. Astropart. Phys. 08 (2003) 001.
[87] J. Maldacena and G. L. Pimentel, Entanglement entropy in de Sitter space, J. High Energy Phys. 02 (2013) 038.
[88] N. Kaloper, M. Kleban, A. E. Lawrence, and S. Shenker, Signatures of short distance physics in the cosmic microwave background, Phys. Rev. D **66**, 123510 (2002).
[89] S. Weinberg, Quantum contributions to cosmological correlations, Phys. Rev. D **72**, 043514 (2005).
[90] L. Senatore and M. Zaldarriaga, On loops in inflation, J. High Energy Phys. 12 (2010) 008.
[91] S. B. Giddings and M. S. Sloth, Semiclassical relations and IR effects in de Sitter and slow-roll space-times, J. Cosmol. Astropart. Phys. 01 (2011) 023.
[92] M. Porrati, Bounds on generic high-energy physics modifications to the primordial power spectrum from back reaction on the metric, Phys. Lett. B **596**, 306 (2004).
[93] S. B. Giddings and A. Strominger, Axion induced topology change in quantum gravity and string theory, Nucl. Phys. **B306**, 890 (1988).
[94] L. F. Abbott and M. B. Wise, Wormholes and global symmetries, Nucl. Phys. **B325**, 687 (1989).
[95] S. R. Coleman and K.-M. Lee, Wormholes made without massless matter fields, Nucl. Phys. **B329**, 387 (1990).
[96] R. Kallosh, A. D. Linde, D. A. Linde, and L. Susskind, Gravity and global symmetries, Phys. Rev. D **52**, 912 (1995).
[97] A. de la Fuente, P. Saraswat, and R. Sundrum, Natural Inflation and Quantum Gravity, Phys. Rev. Lett. **114**, 151303 (2015).
[98] M. Montero, A. M. Uranga, and I. Valenzuela, Transplanckian axions!?, J. High Energy Phys. 08 (2015) 032.
[99] K. Kooner, S. Parameswaran, and I. Zavala, Warping the weak gravity conjecture, arXiv:1509.07049.
[100] J. Brown, W. Cottrell, G. Shiu, and P. Soler, On axionic field ranges, loopholes and the weak gravity conjecture, arXiv:1504.00659.
[101] T. C. Bachlechner, C. Long, and L. McAllister, Planckian axions in string theory, J. High Energy Phys. 12 (2015) 042.
[102] T. C. Bachlechner, C. Long, and L. McAllister, Planckian axions and the weak gravity conjecture, arXiv:1503.07853.
[103] T. Rudelius, On the possibility of large axion moduli spaces, J. Cosmol. Astropart. Phys. 04 (2015) 049.

[104] T. Rudelius, Constraints on axion inflation from the weak gravity conjecture, J. Cosmol. Astropart. Phys. 09 (2015) 020.

[105] B. Heidenreich, M. Reece, and T. Rudelius, Sharpening the weak gravity conjecture with dimensional reduction, arXiv:1509.06374.

[106] S. Kachru, R. Kallosh, A. D. Linde, and S. P. Trivedi, De Sitter vacua in string theory, Phys. Rev. D **68**, 046005 (2003).

[107] J. J. Blanco-Pillado, C. P. Burgess, J. M. Cline, C. Escoda, M. Gomez-Reino, R. Kallosh, A. D. Linde, and F. Quevedo, Racetrack inflation, J. High Energy Phys. 11 (2004) 063.