PHYSICAL REVIEW D 91, 063518 (2015)

Gravitational waves and the scale of inflation

Mehrdad Mirbabayi,[1] Leonardo Senatore,[2] Eva Silverstein,[2] and Matias Zaldarriaga[1]

[1]*Institute for Advanced Study, Princeton, New Jersey 08540, USA*

[2]*Department of Physics and SLAC, Stanford University, Stanford, California 94305, USA*

(Received 8 January 2015; published 12 March 2015)

We revisit alternative mechanisms of gravitational wave production during inflation and argue that they generically emit a non-negligible amount of scalar fluctuations. We find the scalar power is larger than the tensor power by a factor of order $1/\epsilon^2$. For an appreciable tensor contribution, the associated scalar emission completely dominates the zero-point fluctuations of the inflaton, resulting in a tensor-to-scalar ratio $r \sim \epsilon^2$. A more quantitative result can be obtained if one further assumes that gravitational waves are emitted by localized subhorizon processes, giving $r_{\max} \simeq 0.3\epsilon^2$. However, ϵ is generally time dependent, and this result for r depends on its instantaneous value during the production of the sources, rather than just its average value, somewhat relaxing constraints from the tilt n_s. We calculate the scalar 3-point correlation function in the same class of models and show that non-Gaussianity cannot be made arbitrarily small, i.e. $f_{NL} \gtrsim 1$, independently of the value of r. Possible exceptions in multifield scenarios are discussed.

DOI: 10.1103/PhysRevD.91.063518 PACS numbers: 98.80.Cq, 98.80.-k

I. TENSOR EMISSION DURING INFLATION

Inflation stretches the vacuum fluctuations of the graviton field to nearly scale-invariant superhorizon gravitational waves,

$$\langle \gamma^s_{\mathbf{k}} \gamma^{s'}_{\mathbf{k}'} \rangle_{\rm vac} = (2\pi)^3 \delta^{(3)}(\mathbf{k} + \mathbf{k}') \delta^{ss'} \frac{1}{2k^3} \mathcal{P}_{t,\rm vac},$$
$$\mathcal{P}_{t,\rm vac} = \frac{2H^2}{M_{\rm Pl}^2}. \tag{1}$$

These tensor modes lead to B-type polarization of the cosmic microwave background [1,2]. If experiments with sensitivity comparable to those currently taking data were to make a conclusive detection (see e.g. [3]), Eq. (1) would imply that the scale of inflation is $H \simeq 10^{14}$ GeV. However, it has been suggested in [4–6], that there are other (secondary) mechanisms of gravity-wave production during inflation which can dominate the primary effect (1). This invalidates the above inference of the scale of inflation. In these examples the gravitons are not zero-point quantum fluctuations, and therefore a measurement of primordial B modes would not be direct check of the quantization of gravity [3]. Given these prospects now seems to be the right time to reexamine those mechanisms in more detail.

The basic idea behind these secondary mechanisms is to assume there is a sector X that constantly absorbs energy via its coupling to the inflaton field ϕ and emits gravitational waves. This emission can result from localized and nearly incoherent processes that occur periodically. This possibility can be motivated in field theory by assuming an approximate discrete shift symmetry [7]. For instance, the X sector can be composed of particles (strings) whose mass (tension) is a function of ϕ, and naturally arises in string theory motivated models of inflation where the inflaton is a monodromy-extended direction in field space [8], but in a regime of parameters where a sector of the spectrum becomes light each time the field traverses an underlying circle of sub-Planckian period (as in e.g. [9,10]). As a concrete field-theory model consider

$$M_X^2 = M^2 \sin^2\left(\frac{\phi}{f}\right), \tag{2}$$

with $\dot{\phi}_{\rm sr}/f \gg H$. Each time the mass goes through zero, there is a burst of particle production. If these massive particles subsequently decay or scatter each other, they will emit soft gravitons via Bremsstrahlung [4].

Another possibility is to have a process that acts coherently over a Hubble distance. The primary example is a $U(1)$ gauge field with a coupling to the inflaton [5,11],

$$\frac{\alpha}{f} \phi F \tilde{F}. \tag{3}$$

This causes a tachyonic instability of one of the two helicities of the gauge field. If the instability rate is faster than expansion rate (but not too faster to destabilize the inflation), a large helical field is generated and sources a polarized tensor field. We refer to these mechanisms as "coherent emission by extended configurations."

The natural question to ask is whether these scenarios can compete with the zero-point fluctuations (1). Obviously the energy density ρ_X of the auxiliary sector must be a small fraction of total energy density of the Universe $3M_{\rm Pl}^2 H^2$. The Friedmann equations imply

$$\rho_X + p_X + \rho_\phi + p_\phi = -2M_{\rm Pl}^2 \dot{H}, \tag{4}$$

where $\rho_X + p_X \sim \rho_X$ is sourced by $\rho_\phi + p_\phi$ (the kinetic energy of the inflaton in the slow-roll models). We therefore expect

$$\rho_X \lesssim M_{\rm Pl}^2 H^2 \epsilon, \tag{5}$$

where $\epsilon \equiv -\dot{H}/H^2$.

The level of gravitational waves with frequency ω which can potentially be emitted by ρ_X is roughly

$$\gamma_\omega \sim \frac{\rho_X}{M_{\rm Pl}^2 \omega^2}. \tag{6}$$

Taking $\omega \sim H$ and using the upper bound (5) yields

$$\gamma \lesssim \epsilon, \tag{7}$$

which can still be much larger than $H/M_{\rm Pl}$ of vacuum fluctuations (1).

Because of the coupling to the inflaton field, which is the only source of energy during inflation, the X sector necessarily emits scalar waves as well. The second natural question, thus, regards the level of scalar emission as compared to the vacuum fluctuations:

$$\langle \zeta_{\mathbf{k}} \zeta_{\mathbf{k}'} \rangle_{\rm vac} = (2\pi)^3 \delta^{(3)}(\mathbf{k} + \mathbf{k}') \frac{1}{2k^3} \mathcal{P}_{s,\rm vac},$$
$$\mathcal{P}_{s,\rm vac} = \frac{H^2}{2M_{\rm Pl}^2 \epsilon}. \tag{8}$$

Can the scalar emission be kept subdominant to $\mathcal{P}_{s,\rm vac}$? Even if not, can an observably large tensor-to-scalar ratio be explained by these mechanisms?

For the purpose of order of magnitude estimates and heuristic arguments, it is useful to characterize the primordial scalar and tensor power in terms of the number of quanta (N_s and N_t) in each logarithmic interval of wavelength and in a volume of the same wavelength size. In the absence of emission (or absorption) these numbers remain conserved during the expansion of the Universe. Since the fluctuations of vacuum and first excited level in a box are of the same order, these excitation numbers can be approximated by comparing the ratio of the actual power to the zero-point power: $N_t \sim \mathcal{P}_t/\mathcal{P}_{t,\rm vac}$ and $N_s \sim \mathcal{P}_s/\mathcal{P}_{s,\rm vac}$. Hence, in the presence of a secondary mechanism the tensor-to-scalar ratio $r = 4\mathcal{P}_t/\mathcal{P}_s$ would be modified from its usual value, $r = 16\epsilon$, to[1]

$$r \sim 16\epsilon \frac{N_t}{N_s}. \tag{9}$$

What allows us to make general statements about N_t/N_s is the nearly exponential expansion of the Universe:

(i) For any emission process that operates at a physical frequency ω, there is only a short period of time of order H^{-1} during which a given k mode can be excited.

(ii) Suppose this process transfers a total energy of E per Hubble volume into gravity waves within a logarithmic interval of frequency around ω. By the time the waves exit the horizon and freeze, the power dilutes by a factor of ω^{-4}, a fact that follows from our definition,

$$N_t \sim \frac{E}{\omega} \frac{H^3}{\omega^3}. \tag{10}$$

This strongly suppresses waves produced at $\omega \gg H$. In Appendix A, we will illustrate this point in a more concrete example by considering emission of hard gravitons in strong gravity regime. But in the main text concentrate on soft emission at $\omega \sim H$.[2]

A. INCOHERENT EMISSION BY LOCALIZED SUBHORIZON EVENTS

The above observation leads to a dramatic simplification if the emission process happens deep inside the horizon: the details of the production mechanism does not matter anymore. The emission of Hubble wavelength tensor and scalar modes depends only on a few coarse-grained features of the process.

The X sector can be thought of as effectively being composed of particles of mass M, with N_X spontaneous tunneling events per Hubble time per Hubble volume. At each event a total energy M is transferred from the time-dependent background field to X sector. We assume different events are spatially out of phase with respect to one another, while temporally they can be correlated. In addition to its invariant mass, each X particle is characterized by quadrupole and higher moments. The gravitational emission at Hubble wavelength due to the time variation of these moments is suppressed by powers of lH, where l is the characteristic size of the

[1] The inclusion of scalar sound-speed would lead to $\epsilon \to \epsilon c_s$ in all of our constraints and makes them stronger. See complementary discussions in [12–15] and [12,14,16] on the relation between r and respectively the scalar and tensor sound-speed, in the absence of secondary emission.

[2] Both arguments also apply to the emission of scalar modes.

X particle.[3] Large tensor emission requires relativistic and asymmetric processes.

Consider, for instance, the decay of a particle of mass M into two relativistic jets. This process is accompanied by considerable emission of soft gravitons. In flat space, the energy radiated per unit solid angle per unit frequency can be calculated using the standard Bremsstrahlung formula (e.g. [18])

$$\left(\frac{dE_t}{d\Omega d\omega}\right) = \frac{\omega^2}{2\pi^2 M_P^2}\sum_{N,M}\frac{\eta_N\eta_M}{(P_N\cdot k)(P_M\cdot k)} \times \left[(P_N\cdot P_M)^2 - \frac{1}{2}m_N^2 m_M^2\right]. \tag{14}$$

Here the sum is over external particles, $\{P_N\}$ are their momenta, and $\eta = +1$ for in-going particles and -1 otherwise. The result for the decay process is

$$\frac{dE_t}{d\Omega d\omega} = \frac{1}{(2\pi)^3}\left(\frac{M}{2M_{\mathrm{Pl}}}\right)^2. \tag{15}$$

(This formula also serves as an order of magnitude estimate for soft graviton emission from other relativistic processes involving particles of mass M.) Comparing the flat spectral index $dE_t/d\omega \propto \omega^0$ to the dilution factor ω^{-3} of (10) implies that such a decay process during the inflation contributes mainly to the power of modes with $\omega \sim H$, as already anticipated.

Note also that the above formula is valid only asymptotically, that is, if the final states traverse distances much longer than the wavelength of gravitons. If they are caught by multiple subsequent scatterings, although there is emission from each scattering, the waves interfere coherently and destructively (Landau-Pomeranchuk effect). Since we are interested in emission at $\omega \sim H$ and since each mode spends roughly a Hubble time H^{-1} at Hubble frequency, the net effect of multiple scattering is to suppress the power. (Different regimes of Bremsstrahlung from multiple scattering are reviewed in Appendix C.)

Hence we expect, in order to have the largest gravitational emission, each X particle should participate in a single relativistic event. The exact amount of emition depends on details of the process. However among such events the decay (15) seems to be the most efficient one. As long as different tunneling events are independent the total number of gravitons produced is obtained by summing over individual decays:

$$N_t \sim \frac{n_X(t_f)}{H^3}\frac{dE_t}{d\omega} \sim N_X\left(\frac{M}{M_{\mathrm{Pl}}}\right)^2\left(\frac{a(t_i)}{a(t_f)}\right)^3. \tag{16}$$

This implies a suppression factor if the lifetime of X particles, $\Delta t = t_f - t_i$, is comparable or longer than the Hubble time. We henceforth assume the opposite regime.

Let us make two final remarks. First, if subgroups of X particles emit coherently, for instance, by n of them merging into a bound state which subsequently decays, (16) gets enhanced by a factor of n. We were unable to find a realistic model of this kind and leave this as an open possibility.

Second, one should ask if the tunneling event itself leads to considerable gravitational emission. As already mentioned, if each event results in a localized massive object of small size (and, hence, small quadrupole moment), the emission is negligible. In more realistic scenarios particles are produced in pairs [4]. If these are two nonrelativistic particles of mass $M/2$, that conclusion still holds. (In fact the most efficient emission process would then be for the two particles annihilating into two relativistic jets, so that the pair can effectively be treated as a single particle of mass M all along.) On the other hand, if the pair is relativistic, there is gravitational emission given exactly by (15).

We postpone the actual calculation of the de Sitter correlation function resulting from the decay process to Appendix B.

B. COHERENT EMISSION BY EXTENDED CONFIGURATIONS

As in the example of the $U(1)$ gauge field, the tensor emission can happen coherently over horizon-size patches. Obtaining precise universal results seems impossible in this case. Nevertheless, an order of magnitude estimation of the maximum tensor emission, given energy M per Hubble volume of an extended configuration, is easy. In the weak gravity regime, we expect

[3] More explicitly, the emission in this regime is dominated by the interaction [17]

$$\int d^4x \int dt_X \gamma^{-1}\delta^4(x^\mu - x_X^\mu(\tau))Q^{ij}(t_X)R_{0i0j}(x) = \int d\tau Q^{ij}(t_X)R_{0i0j}(x_X) \tag{11}$$

where τ is the proper time of the particle emitting gravity waves, $R_{\mu\nu\rho\sigma}$ is the Riemann tensor, and Q^{ij} is the intrinsic quadrupole of the particle. The standard formula of quadrupole emission gives (see e.g. [18])

$$\frac{dE_t}{d\omega} \sim \frac{\omega^6 Q(\omega)^2}{M_{\mathrm{Pl}}^2} \sim \frac{\omega^4 Q(t)^2}{M_{\mathrm{Pl}}^2} \tag{12}$$

where in the last step we have used that $Q(\omega) = \int dt e^{i\omega t}Q(t) \sim Q(t)/\omega$. To relate this to the mass of the object, We write $Q \sim Ml^2$, where l is the typical size of the object. Taking $\omega \sim v/l$, where $\nu(\ll c)$ is the characteristic velocity, we obtain

$$\frac{dE_t}{d\omega} \sim \frac{\frac{v^4}{l^4}M^2 l^4}{M_{\mathrm{Pl}}^2} \sim v^4\frac{M^2}{M_{\mathrm{Pl}}^2}. \tag{13}$$

$$\gamma(\omega \sim H) \sim \frac{MH}{M_{\rm Pl}^2}. \tag{17}$$

(As before, the main contribution comes from emission into nearly Hubble frequency modes.) We, therefore, get $\mathcal{P}_t \sim M^2H^2/M_{\rm Pl}^4$. If there are several extended configurations emitting independently (e.g. N species of $U(1)$ gauge fields in the example of [6]), their contribution adds up to give

$$N_t \sim N\frac{M^2}{M_{\rm Pl}^2}. \tag{18}$$

Note that symmetries can highly suppress tensor emission compared to this naive expectation.

II. SCALAR EMISSION AND ENERGY CONSERVATION

Since the energy M of each X event is provided by the coupling of X sector to the time-dependent inflaton background field, each tunneling event leads to scalar emission. By the same arguments as for gravitons [see (10)], only soft scalars of wavelength $\lambda \sim 1/H$ need to be considered. This emission can be calculated in a model-independent way using the effective field theory of inflation (EFTofI) [12,19]. However, to build intuition let us first consider a single-field slow-roll model.

For deeply subhorizon events curvature can be ignored and energy conservation implies that the energy M must come from the background inflaton. Therefore, the X particle production must be accompanied by a scalar wave $\delta\phi(t,x)$ which is responsible for that energy deficit. The stress-energy tensor of ϕ is

$$T_{\mu\nu} = \partial_\mu\phi\partial_\nu\phi - g_{\mu\nu}\left[\frac{1}{2}(\partial\phi)^2 - V(\phi)\right]. \tag{19}$$

The perturbed energy density is

$$\delta T_{00} = \dot\phi_{\rm sr}\delta\dot\phi + V'\delta\phi + O(\delta\phi^2). \tag{20}$$

Terms quadratic in $\delta\phi$ correspond to the energy carried by the wave of $\delta\phi$, and are negligible compared to the linear terms. Moreover, the scalar profile would be a sharp pulse with characteristic width determined by the size and duration of the tunneling event. As such most of the energy is carried away by frequencies higher than the Hubble rate. Therefore, the second term of (20) which can be written as $-3H\dot\phi_{\rm sr}\delta\phi$, is suppressed compared to the first one. By energy conservation δT_{00} must integrate to $-M$, implying

$$\int d^3x\delta\dot\phi = -\frac{M}{\dot\phi_{\rm sr}}. \tag{21}$$

(Note that the integral on the left is the Noether charge associated to the approximate shift symmetry of $\delta\phi$, whose mass is much less than H; it remains conserved after the tunneling event.)

Now consider the flat-space expression for the energy emitted in the wave of $\delta\phi$:

$$E_s = \int d^3x\frac{1}{2}[\delta\dot\phi^2 + (\nabla\delta\phi)^2]. \tag{22}$$

Using (21) and dimensional analysis, we get at small frequencies

$$\frac{dE_s}{d\omega} \sim \frac{M^2}{\dot\phi_{\rm sr}^2}\omega^2. \tag{23}$$

Because of the dilution effects associated with the expansion the scalar power would be most affected by the lower end of the spectrum, frequencies of order H. For these frequencies our flat space analysis is only an order of magnitude estimate.

A. Generalization

Consider an independent localized tunneling event in which energy M is transferred to X sector in a period much shorter than a Hubble time. By translational invariance of the background the total momentum transfer must be zero. As argued before for emission of long-wavelength tensor and scalars the detailed structure of the event is unimportant. Therefore we can describe it by the production of a particle at rest whose mass grows from 0 to M. Choosing a frame where the constant-time hypersurfaces coincide with constant-inflaton hypersurfaces, the X particle would be described by a Dirac-Born-Infeld action with a time-dependent mass $M(t)$

$$S_X = -\int d\tau M(t)\sqrt{g_{\mu\nu}\dot x_X^\mu \dot x_X^\nu}. \tag{24}$$

Following the EFTofI [12], we can restore time-diffeomorphism invariance by shifting $t \to t + \pi$ and prescribing the right symmetry transformation to π. Hence any explicit time-dependence [such as $M(t)$] leads to a linear coupling $\pi\partial_t\mathcal{L}$. Since explicit time-dependence in the action results in energy nonconservation

$$\partial_\mu T_0^\mu = -\partial_t\mathcal{L}, \tag{25}$$

on subhorizon scales π couples at leading order to $-\partial_\mu T_0^\mu$. This coupling results in a universal emission of π whenever energy is transferred from background to X sector, and is a consequence of the conservation of total energy. Due to the mixing of π and metric fluctuations $h_{\mu\nu}$ the gravitational stress-energy tensor of π starts linear and is sign indefinite. Therefore, in any particle production event the total stress-energy tensor (X plus π) is conserved well inside the horizon.

Up to slow-roll corrections the dynamics of scalar modes during inflation is adequately described by π alone. Restricting to the leading derivative operators in the EFTofI, the relevant action for π is therefore

$$S = -\int d^4x\sqrt{-g}M_{\rm Pl}^2\dot{H}[\dot{\pi}^2 - a^{-2}(\partial\pi)^2] + \int d^4x M(t+\pi)\int dt_X\gamma^{-1}\delta^4(x^\mu - x_X^\mu(\tau)), \quad (26)$$

where $\gamma = p_X^0/M_X$. For each independent event momentum conservation forces $\gamma = 1$. The above action results in a cubic coupling between the canonically normalized field $\pi_c = \sqrt{2\epsilon}M_{\rm Pl}H\pi$ and X particles with strength[4]

$$g_{\rm eff} \equiv \frac{\dot{M}}{\sqrt{2\epsilon}M_{\rm Pl}H}. \quad (27)$$

The relative factor of $\sqrt{-g}$ between the first and second line of (26) is responsible for a dilution effect similar to (16). In the following we assume $H(t_f - t_i) \ll 1$.

The flat space solution for π in the presence of a single source X at $\mathbf{x} = 0$ is

$$\pi_c(t,\mathbf{x}) = \frac{1}{\sqrt{2\epsilon}M_{\rm Pl}H}\int\frac{d^3\mathbf{k}}{(2\pi)^3}e^{i\mathbf{k}\cdot\mathbf{x}} \times\left[\frac{ie^{-ikt}}{2k}\int_{t_i}^{t_f}\dot{M}e^{ik\tau}dt' + \text{c.c.}\right], \quad (28)$$

which after integration by parts gives

$$\pi_c(t,\mathbf{x}) = \frac{M(t_f)}{\sqrt{2\epsilon}M_{\rm Pl}H}\int\frac{d^3\mathbf{k}}{(2\pi)^3k}e^{i\mathbf{k}\cdot\mathbf{x}}\sin k(t-t_f), \quad (29)$$

plus terms suppressed by $O(k(t_f - t_i))$ which is negligible because we are interested in $k \sim H$. Using the expression for the energy density of the canonically normalized field in flat space: $(\dot{\pi}_c^2 + (\nabla\pi_c)^2)/2$ we obtain the following result for the energy emitted per unit frequency per unit solid angle

$$\frac{dE_s}{d\omega d\Omega} = \frac{1}{(2\pi)^3}\frac{M^2\omega^2}{4\epsilon M_{\rm Pl}^2H^2}. \quad (30)$$

As before, the total scalar emission is obtained by summing over independent X events. The maximum ratio N_t/N_s can therefore be calculated by comparing (15), which appears to be to the most efficient gravity wave production scenario, with the above formula at $\omega \sim H$. This yields[5]

$$\frac{N_{t,\max}}{N_s} \sim \epsilon. \quad (31)$$

Evidently, in order for the production mechanism to have any significance, that is $N_t \sim 1$ or larger, the scalar byproducts would completely dominate the vacuum fluctuations (8). Therefore according to (9), the largest possible tensor-to-scalar ratio which can be obtained in this scenario is of order ϵ^2. The more careful calculation of scalar and tensor correlation functions in de Sitter space (Appendix B) yields

$$r_{\max} \simeq 0.3\epsilon^2. \quad (32)$$

An observable level of B modes, $r \gtrsim 10^{-3}$, requires $\epsilon > 0.05$ in these scenarios. The Hubble parameter would therefore drop by about one order of magnitude during the 60 e-folds of inflation. Nevertheless, the scalar and tensor tilt can remain sufficiently small as we will estimate in the next section. From this estimate a value of $r \gtrsim 0.1$ seems hard to be explained with this class of models, although that may depend on the extent to which ϵ varies.[6]

The scalar emission by extended objects can be estimated as follows. The $\pi\partial_\mu T_0^\mu$ coupling and the overall normalization of π kinetic term in (26) imply

$$\pi(\omega \sim H) \sim \frac{MH^2}{M_{\rm Pl}^2\dot{H}}, \quad (33)$$

for energy M of the configuration in a Hubble volume. Using $\zeta \simeq -H\pi$ gives $P_s \sim H^2M^2/M_{\rm Pl}^4\epsilon^2$. Comparison with $\mathcal{P}_{s,\rm vac}$ gives

$$N_s \sim N\frac{M^2}{\epsilon M_{\rm Pl}^2} \quad (34)$$

[4]The minimal scenario considered above (and in [4]) corresponds to $M = g\phi$ and $\delta\phi = \pi_c$.

[5]We note that the ϵ appearing here is the instantaneous ϵ at the time of X production. It is natural to expect ϵ to have periodic variations in the inflationary models with particle production; however, the relative amplitude of oscillations is small in the conventional models [7]. Models with significant variations in ϵ can perhaps be constructed by considering nonmonotonic inflationary potentials. In this case, one of the ϵ factors in the bound (32) has to be replaced by the instantaneous value of ϵ at scalar emission.

[6]In the example of tensor emission by long string pairs considered in [4], the scalar emission during the process of energy transfer from background to the X sector was missed. This led to a much larger estimates for N_t/N_s. In fact, the tensor-to-scalar ratio is much smaller in this case ($r \sim \epsilon^2/N_{\rm loop}$), because the initial scalar emission by the long string pair is coherent, while the subsequent gravitational emission by the decay of the pair into $N_{\rm loop}$ pieces is incoherent.

where we also inserted the number of species N. Comparing with (18) gives $r_{\max} \sim \epsilon^2$.

B. Multifield inflation

In the context of tensor emission by $U(1)$ gauge field production, it has been suggested [6] to decouple scalar emission by introducing another scalar field ψ which is slow-rolling (say with $\dot{\psi}_{\rm sr} \ll \dot{\phi}_{\rm sr}$):

$$S = \int \sqrt{-g}\left[\frac{M_{\rm Pl}^2}{2}R + \frac{1}{2}(\partial\phi)^2 + \frac{1}{2}(\partial\psi)^2 - V(\phi,\psi) - \frac{1}{4}F^2 - \frac{\psi}{4f}F\tilde{F}\right]. \tag{35}$$

By replacing (3) with $\frac{1}{f}\psi F\tilde{F}$ the energy source of the auxiliary sector becomes $\dot{\psi}_{\rm sr}^2/2$. If in addition the scalar spectrum is exclusively determined by ϕ fluctuations, then energy conservation doesn't seem to enforce any correlation between tensor and scalar emission. Note that this idea, if viable, can be used in other production mechanisms as well. But is it viable?

Let us first understand this proposal in view of the above argument for the universal π emission as a consequence of energy conservation. In the EFTofI the fluctuations of fields are decomposed into parallel (adiabatic) and perpendicular (isocurvature) to the background trajectory in field space [19]. The π field, which we refer to as the inflaton, is the fluctuations along the background trajectory. As such, it is a linear combination of the fluctuations of ϕ and ψ, and after canonical normalization reads

$$\pi_c = \frac{1}{\sqrt{\dot{\phi}_{\rm sr}^2 + \dot{\psi}_{\rm sr}^2}}(\dot{\phi}_{\rm sr}\delta\phi + \dot{\psi}_{\rm sr}\delta\psi). \tag{36}$$

In particular it couples to the gauge fields:

$$L_{\pi A} = \frac{\alpha}{4f}\pi_c F\tilde{F}, \qquad \alpha \simeq \frac{\dot{\psi}_{\rm sr}}{\dot{\phi}_{\rm sr}}. \tag{37}$$

There is also a light field σ which characterizes perpendicular fluctuation in field space, and has a similar coupling to $F\tilde{F}$ but with $\alpha \sim 1$.

The copious production of gauge fields excites π and results in a contribution to the scalar power that is the same as the single-field version of the model [11]

$$\mathcal{P}_s = \mathcal{P}_{s,\rm vac}\left[1 + 7.5\times 10^{-5}\mathcal{P}_{s,\rm vac}\frac{e^{4\pi\xi}}{\xi^6}\right], \tag{38}$$

with $\xi = \alpha\dot{\phi}_{\rm sr}/2fH$. Comparing to the tensor production [6]

$$\mathcal{P}_t = 16\epsilon\mathcal{P}_{s,\rm vac}\left[1 + 3.4\times 10^{-5}\epsilon\mathcal{P}_{s,\rm vac}\frac{e^{4\pi\xi}}{\xi^6}\right], \tag{39}$$

we see the same ϵ^2 suppression. As expected π couples to $\partial_\mu T_0^\mu$ and whenever there is particle production (gauge fields in this case) there is an associated emission of π.[7]

However, the isocurvature fluctuations (σ) in this model are also sourced by the gauge fields. The resulting isocurvature modes can later convert into adiabatic ones [19]:

$$\zeta \simeq -H\pi + \zeta_{,\sigma}\sigma. \tag{40}$$

Therefore, if $\zeta_{,\sigma} \simeq H\dot{\psi}_{\rm sr}/\dot{\phi}_{\rm sr}^2$ the net contribution of the gauge fields to observed scalar spectrum can be made negligible.

Since this conversion happens at superhorizon scales, it is indeed easier to work in terms of the background model formulated in terms of ϕ and ψ, and ask if it is possible to decouple ζ from $\delta\psi$, up to possible slow-roll suppressed corrections. This problem is studied in more detail in Appendix D. It is argued that choosing the reheating surface to be determined by ϕ and $\epsilon \ll 1/N_e$ (where $N_e \sim 60$ is the number of e-folds of inflation), decouples $\delta\psi$ and corresponds to the above value for $\zeta_{,\sigma}$.

Another multifield inflationary model that evades our conclusion because of a fundamentally different reason is chromonatural inflation [23]. Here there is a non-Abelian gauge field background which causes the perturbations of the gauge field mix with the tensor modes. The assumption that tensor modes couple universally via a cubic coupling of strength $1/M_{\rm Pl}$ does not hold in this example since the gauge field fluctuations can directly oscillate into gravitons.

III. NON-GAUSSIANITY AND TILT

In the last section we argued that for a large class of models, large gravitational emission implies dominant scalar emission. The scalar spectrum is naturally expected to be non-Gaussian. In the case of localized emission, the non-Gaussianity is calculated in Appendix B, and in terms of the conventional f_{NL} parameter,

$$f_{NL}\zeta \sim \frac{\langle\zeta^3\rangle}{\langle\zeta^2\rangle^{3/2}} \sim N_X^{-1/2}, \tag{41}$$

where $N_X \sim n_X H^{-3}$, the number of X particles per Hubble volume, is related to N_s through (30) and the assumption of incoherent emission, namely,

[7]We notice in passing that the signal in this model is exponentially sensitive to the value of ξ, so that it is observationally relevant only for a very small range of values of ξ. See [20,21] for a more detailed study of this model, and [22] for a related work on multi-field scenarios.

$$N_s \sim N_X \frac{M^2}{\epsilon M_{\rm Pl}^2}. \tag{42}$$

In (41), ζ is a shorthand:

$$\zeta \equiv \mathcal{P}_s^{1/2} = N_s^{1/2}\mathcal{P}_{s,\rm vac}^{1/2} \sim 10^{-5}. \tag{43}$$

By energy conservation N_X cannot be made arbitrarily large, therefore there is a lower bound on f_{NL}.[8] Let us rewrite the constraint (5) using the average number density n_X of tunneling events and the final mass M:

$$\rho_X = n_X M \lesssim M_{\rm Pl}^2 H^2 \epsilon. \tag{44}$$

Multiplying both sides by a Hubble volume $1/H^3$ gives

$$N_X \frac{MH}{\epsilon M_{\rm Pl}^2} \lesssim 1. \tag{45}$$

Using (42) the lhs can be written as $N_X^{1/2} N_s^{1/2} H/M_{\rm Pl}\sqrt{\epsilon}$, which together with (41) and (43) results in

$$f_{NL} \gtrsim 1. \tag{46}$$

The large non-Gaussianity can be used to break the degeneracy between the production scenarios and the conventional one. In Appendix B we provide the shape of the induced non-Gaussianity. We find that the cosine with the standard equilateral [24] and orthogonal [25] templates is quite large, 0.97, making it rather challenging to actually distinguish this shape from the standard ones.

We should emphasize that the above degree of non-Gaussianity (41), which arises from stochasticity of the emission process, is the minimum level based on very general assumptions. In any concrete model there may be other sources of non-Gaussianity, which we do not expect to cancel the stochastic piece. For instance, if after tunneling the X particles remain coupled to the inflaton for a longer period of time, their scalar emission would be influenced and, hence, correlated with the previously emitted waves. This is the case in the ‘t‘rapped inflation” model [9]. The already existing non-Gaussianity bounds may then lead to more stringent constraints on the maximum level of gravitational wave emission as argued in [20]. Note, however, that various sources of non-Gaussianity are expected to be independent.

Tilt: We finally calculate the tilt of scalar and tensor spectra assuming that the coupling strength g and mass M are dictated by UV physics and remain approximately constant during inflation. Since $\dot{M}$ is the only relevant dimensionful parameter at the time of tunneling, the number density of X events is expected to be $n_X \sim \dot{M}^{3/2}$ or $N_X \sim (\dot{M}H^2)^{3/2}$. Using the formulas (27), (43), and (42), we get

$$\zeta^2 \sim g^{3/2}\frac{M^2}{M_{\rm Pl}^2}\left(\frac{H}{M_{\rm Pl}}\right)^{1/2}\epsilon^{-5/4}. \tag{47}$$

The scalar tilt is, therefore,

$$n_s - 1 = -\frac{1}{2}\epsilon - \frac{5}{4}\epsilon_2, \tag{48}$$

where $\epsilon_2 \equiv \dot{\epsilon}/\epsilon H$. A similar calculation gives for the tensor tilt

$$n_t = -\frac{1}{2}\epsilon + \frac{3}{4}\epsilon_2. \tag{49}$$

Thus, even $\epsilon \simeq 0.1$ does not require significant fine-tuning. Here we did not make a distinction between the instantaneous and average ϵ, but that only relaxes the constraints from n_s.

IV. CONCLUSIONS

We have analyzed a large class of inflationary models and found that there is a generic upper bound on the tensor-to-scalar ratio in secondary mechanisms of tensor production. This is because emitting gravitons requires energy, and transferring energy from time-dependent background to another sector leads to coupling to the inflaton π and scalar emission. This emission is expected to be larger than the tensor emission by a factor of $1/\epsilon^2$. Therefore, unless there are isocurvature modes which are produced by the same process and cancel the former contribution to ζ, tensor-to-scalar ratio will be $O(\epsilon^2)$. These models are also associated with large non-Gaussianity, i.e. $f_{NL} \gtrsim 1$, independently of the value of r. In the event of a detection of r, non-Gaussianities would provide a way to potentially, though challengingly, distinguish this scenarios from the standard one.

ACKNOWLEDGMENTS

We thank S. Adler, G. D’Amico, R. Flauger, A. Gruzinov, M. Kleban, S. Mukohyama, R. Namba, M. Peloso, and G. Shiu for stimulating discussions. M. M. is supported by National Science Foundation Grants No. PHY-1314311 and No. PHY-0855425. L. S. is supported by Department of Energy Early Career Award No. DE-FG02-12ER41854 and the National Science Foundation Grant No. PHY-1068380. E. S. is supported in part by the National Science Foundation under Grants No. PHY-0756174 and No. PHY11-25915 and by the Department of Energy under Contract No. DE-AC03-76SF00515. M. Z. is supported in part by the National

[8]Another possibility is to have yet another contribution to the scalar spectrum ($N_s{}'$) with a Gaussian distribution and $N_s{}' \gg N_s$. This is unlikely to allow any observable gravity wave signal as it suppresses the ratio (32) by another factor of $N_s/N_s{}'$.

Science Foundation Grants No. AST- 0907969, No. PHY-1213563 and No. AST-1409709.

APPENDIX A: HARD GRAVITONS

Only well-localized objects can emit hard gravitons. Clearly the effect is larger in the strong gravity regime. So consider the production of n_X mass-M black hole pairs per unit volume which subsequently merge and emit an order-one fraction of their mass into gravitons of frequency

$$\omega \sim \frac{1}{r_g} \sim \frac{M_{\rm Pl}^2}{M}. \tag{A1}$$

The number of tensor modes is of order

$$N_t \sim \frac{n_X M}{\omega^4} \sim \frac{n_X}{\omega^3} \frac{M^2}{M_{\rm Pl}^2}, \tag{A2}$$

where $n_X \ll \omega^3$ but $M \gg M_{\rm Pl}$, so there is a chance of having $N_t > 1$. However, when compared to the associated soft scalar emission,

$$N_s \sim \frac{n_X}{H^3} \frac{M^2}{\epsilon M_{\rm Pl}^2}, \tag{A3}$$

one obtains the much smaller tensor-to-scalar ratio,

$$r \sim \epsilon^2 \frac{H^3}{\omega^3}. \tag{A4}$$

APPENDIX B: DE SITTER CORRELATORS

To calculate de Sitter correlation functions of π, one has to take the expansion of the Universe and periodicity of the production mechanism into account. Suppose there is a sequence of random production of X particles with average proper density $\bar{n}_X$ and at moments η_n. Let us at each η_n divide the space into small cells i, each of comoving volume δv_i, so small that $p_i \equiv \bar{n}_X a^3 \delta v_i \ll 1$. To each cell assign a random variable $X_{i,n} = 1$ with probability p, and 0 otherwise. As we saw in (29) each event is practically a delta function source for π. Hence, the field $\pi_{\mathbf{k}}(\eta)$ resulting from creation events can be written

$$\pi_{\mathbf{k}}(\eta) = \frac{M}{2\epsilon M_{\rm Pl}^2 H^2} \sum_n G_{\mathbf{k}}(\eta, \eta_n) \sum_i X_{i,n} e^{i\mathbf{k}\cdot\mathbf{x}_i} \tag{B1}$$

where the first sum is over the production times and the second on the cells. The de Sitter retarded Green's function in the limit $k\eta \to 0$ simplifies to

$$G_{\mathbf{k}}(0, \eta_n) = \frac{H^2}{k^3}(\sin k\eta_n - k\eta_n \cos k\eta_n) \equiv \frac{H^2}{k^3} g(k\eta_n). \tag{B2}$$

The late-time 2-point function of π can be calculated by noting that

$$\langle X_{i,n} X_{j,m} \rangle = \bar{n}_X^2 a_n^2 a_m^2 \delta v_i \delta v_j + \bar{n}_X a_n^3 \delta v_i \delta_{nm} \delta_{ij}. \tag{B3}$$

The first term gives a disconnected contribution proportional to $\delta^3(\mathbf{k}_1)\delta^3(\mathbf{k}_2)$, but the second term gives

$$\langle \pi_{\mathbf{k}_1} \pi_{\mathbf{k}_2} \rangle' = \frac{M^2}{(2\epsilon M_{\rm Pl}^2)^2 k^3} \left[\frac{\bar{n}_X}{H^3} \sum_n \frac{g^2(k\eta_n)}{-k^3 \eta_n^3} \right]. \tag{B4}$$

Where prime means that $(2\pi)^3 \delta^3(\mathbf{k}_1 + \mathbf{k}_2)$ has been omitted. The sum gets contribution only from those η_n for which $k\eta_n = O(1)$. Thus the expression in brackets is $O(N_X)$. The calculation of 3-point function of π is very similar and gives

$$\langle \pi_{\mathbf{k}_1} \pi_{\mathbf{k}_2} \pi_{\mathbf{k}_3} \rangle' = \frac{M^3}{(2\epsilon M_{\rm Pl}^2)^3 k_1^2 k_2^2 k_3^2} \left[\frac{\bar{n}_X}{H^3} \sum_n \prod_i \frac{g(k_i \eta_n)}{-k_i \eta_n} \right]. \tag{B5}$$

The expression in the bracket is $O(N_X)$. From (B2) it follows that in the squeezed limit $k_1 \to 0$ the above expression scales as $O(k_1^0)$. The self-interactions of π and slow-roll suppressed higher-order terms in the conversion between π and ζ will change the squeezed limit behavior of the 3-point function.

Tensors: The calculation of tensor power is similar. One first defines the helicity components of the transverse-traceless part of the spatial metric

$$\gamma_{\mathbf{k},ij} = \gamma_{\mathbf{k}}^+ \varepsilon_{ij}^+ + \gamma_{\mathbf{k}}^- k \varepsilon_{ij}^-, \tag{B6}$$

where $\varepsilon_{ij}^\pm$ are $\mathbf{k}$-dependent polarization tensors: $\varepsilon_{ij}^\pm = \varepsilon_i^\pm \varepsilon_j^\pm / \sqrt{2}$, with $\varepsilon^\pm = \hat{x} \pm i\hat{y}$ for $\mathbf{k} \propto \hat{z}$. The linear equation of motion for the helicity modes is

$$\gamma_{\mathbf{k}}^{r\prime\prime} - \frac{2}{\eta} \gamma_{\mathbf{k}}^{r\prime} + k^2 \gamma_{\mathbf{k}}^r = 16\pi G a^2 \frac{1}{2\sqrt{2}} \varepsilon_i^r \varepsilon_j^r T_{\mathbf{k}}^{ij}. \tag{B7}$$

In the decay of a particle of mass M into two relativistic jets the rhs gets contribution only from the stress tensor of the jets:

$$T^{ij} = \frac{\eta}{\eta_n} E \hat{p}^i \hat{p}^j a^{-3} \delta^{(3)}(\mathbf{x} - \hat{p}(\eta - \eta_n)) \theta(\eta - \eta_n) \tag{B8}$$

where η_n is the decay time and $p = E\hat{p}$ the momentum of the jet. For an event at $\mathbf{x} = 0$ one obtains (after summing the contribution of the two jets)

$$\gamma_{\mathbf{k}}^r(\eta) = \frac{M}{\sqrt{2} M_{\rm Pl}^2} (\hat{\varepsilon}^r \cdot \hat{p})^2 \frac{1}{H\eta_n} \int_{\eta_n}^0 d\eta' G_{\mathbf{k}}(\eta, \eta') \cos(\hat{p} \cdot k\eta'). \tag{B9}$$

Hence, to calculate the emission from multiple events one needs to keep track of $\hat{p}$ of each event in addition to its position $\mathbf{x}_i$ and time η_n. We therefore define the random variable $X_{i,\hat{p},n} = 1$ with probability $p_i = n_X a_n^3 \delta v_i d^2\hat{p}/4\pi$, and 0 otherwise. A similar calculation as in the scalar case leads to

$$\langle \gamma^r_{\mathbf{k}} \gamma^s_{\mathbf{k}'} \rangle' = \frac{\delta^{rs}}{2k^3} p_{t,\mathrm{vac}} \frac{M^2}{M_{\mathrm{Pl}}^2} \left[\frac{\bar{n}_X}{H^3} \sum_n \frac{1}{-2k^3\eta_n^3} \times \int_0^1 d\mu \left(\frac{1-\mu^2}{\eta_n} \int_{\eta_n}^0 d\eta g(k\eta) \cos\mu k\eta \right)^2 \right]. \tag{B10}$$

Comparing to (B4) yields

$$\frac{\mathcal{P}_t/\mathcal{P}_{t,\mathrm{vac}}}{\mathcal{P}_s/\mathcal{P}_{s,\mathrm{vac}}} = \epsilon \frac{\sum_n H(k\eta_n)}{\sum_n G(k\eta_n)} \tag{B11}$$

where G and H represent the summands in (B4) and (B10).[9]

To get a more accurate result, we assume uniform distribution of creation events in physical time, and approximate the sums in (B11) by the integral $\int \frac{dt}{\Delta t} = \int_{-\infty}^0 d\eta_n/(\eta_n \Delta t H)$, with Δt being the time-spacing of the events, giving $\frac{\mathcal{P}_t/\mathcal{P}_{t,\mathrm{vac}}}{\mathcal{P}_s/\mathcal{P}_{s,\mathrm{vac}}} \simeq 0.018\epsilon$ or

$$r_{\max} \simeq 0.3\epsilon^2. \tag{B13}$$

APPENDIX C: BREMSSTRAHLUNG EMISSION FROM MULTIPLE SCATTERINGS

In this appendix we briefly review different regimes of soft emission in multiple scattering processes following [26,27]. The original derivation is for radiation by a relativistic electron moving inside matter. The Bremsrahlung emission of soft photons now depends on the rate of scattering Γ and the average scattering angle per unit length $q = \langle \theta_{\mathrm{sc}}^2 \rangle / l$. Replacing $\alpha \to M^2/M_{\mathrm{Pl}}^2$, where M is the center of mass energy of the scattering process, gives a rough estimate of gravitational Bremsstrahlung. There are three regimes:

(I) Small angle: Consider a single scattering event with a very small angle $\theta_{\mathrm{sc}} \ll \gamma^{-1}$. The coherence time τ_1 over which the electron can influence the emitted photon is

$$\tau_1(1-v) \sim \lambda. \tag{C1}$$

Using $1 - v \sim \gamma^{-2}$ we get

$$\tau_1 \sim \frac{\gamma^2}{\omega} \gg \frac{1}{\omega}. \tag{C2}$$

Since we are interested in gravitational waves of Hubble wavelength emitted in a Hubble time, we should consider emission during

$$T \sim \frac{1}{\omega} \sim \frac{1}{H}. \tag{C3}$$

Since $\tau_1 \gg T$, the whole process can be approximated by a single event in this regime. Moreover, the energy emitted per unit frequency receives a small angle suppression

$$\frac{dE}{d\omega} \sim \alpha \frac{\Delta p^2}{m^2} \sim \alpha\gamma^2\theta_{\mathrm{sc}}^2 \ll \alpha. \tag{C4}$$

The above treatment is valid even if there are multiple scatterings, as long as $q\tau_1 \ll \gamma^{-2}$.

(II) Landau-Pomeranchuk: Suppose q is increased beyond that. The new coherence length τ_2 becomes shorter and will be determined in terms of $\theta_{\mathrm{sc}} = \sqrt{q\tau_2} > \gamma^{-1}$ according to

$$\tau_2 c(1 - \cos\theta_{\mathrm{sc}}) \sim \lambda, \tag{C5}$$

or

$$\tau_2 \sim \frac{1}{\omega\theta_{\mathrm{sc}}^2} \gg \frac{1}{\omega}, \tag{C6}$$

where we assumed θ_{sc} is still much less than unity. This implies

$$\tau_2 = \frac{1}{\sqrt{\omega q}} \quad \text{and} \quad \theta_{\mathrm{sc}}^2 = \sqrt{\frac{q}{\omega}}. \tag{C7}$$

The requirement $\gamma^{-1} \ll \theta_{\mathrm{sc}} \ll 1$ in one coherence length τ_2 gives

$$1 \ll \frac{\omega}{q} \ll \gamma^4. \tag{C8}$$

The emission rate can be computed as follows. One first cuts the particle trajectory into coherent pieces of length τ_2. Each segment can be thought of as a particle moving in a straight line whose charge is turned on at

[9] For $-k\eta_n \gg 1$, $G \simeq -\cos^2 k\eta_n/k\eta_n$ while $H \simeq (\cos 2k\eta_n - 2)/12k^3\eta_n^3$. Averaging over several periods $1/k \ll \Delta\eta \ll -\eta_n$ we get for the ratio

$$\frac{\langle G(k\eta_n)\rangle_{\Delta\eta}}{k^2\eta_n^2 \langle H(k\eta_n)\rangle_{\Delta\eta}} \simeq 3, \tag{B12}$$

which is in rough agreement with the ratio 2 inferred from the flat space results (15) and (30) (recall that (15) is the energy emitted in both tensor polarizations).

some moment and off after τ_2. The emission from different segments add up incoherently. Unlike (C4), there is no small angle suppression since $\theta_{sc}\gamma \gg 1$; therefore,

$$\frac{dE}{d\omega} \sim \frac{T}{\tau_2}\alpha = T\alpha\sqrt{q\omega}. \tag{C9}$$

This is the standard Landau-Pomeranchuk formula. Note that for $T = 1/\omega$ this is a suppression compared to a single large angle scattering, giving

$$\frac{dE}{d\omega} \sim \alpha\sqrt{\frac{q}{\omega}} \ll \alpha. \tag{C10}$$

(III) Large angle: Finally, when $\omega \ll q$ there will be a lot of scatterings of order-one angle in a wavelength. Now the emission can be obtained by dividing the electron trajectory into segments of length $\tau_3 = 1/q$. Each segment makes an order-one angle with the next one and, hence, emits incoherently. However, now the segment is shorter than the wavelength which results in a suppression of $\omega\tau_3$ in the amplitude. The emission formula becomes

$$\frac{dE}{d\omega} \sim Tq\alpha\frac{\omega^2}{q^2}. \tag{C11}$$

We conclude that the maximum amount of Bremsstrahlung emission during $T \sim 1/\omega$ is obtained by a single large angle scattering which results in $dE/d\omega \sim \alpha$.

APPENDIX D: SECONDARY GRAVITATIONAL EMISSION IN MULTIFIELD INFLATION

Suppose there are two scalar fields ϕ and ψ rolling during inflation, and the energy for the auxiliary sector responsible for the gravitational emission is provided by coupling to ψ. The energy transfer from the background $\dot\psi$ to this sector leads to emission of $\delta\psi$ quanta. The contribution of these fluctuations to adiabatic modes can be obtained using δN formalism [28], where $\zeta = \delta N$ is calculated by determining how much the expansion,

$$N = \int dt H(t), \tag{D1}$$

differs in different patches of the Universe with different values of ψ (and other fields). If eventually the ψ fluctuations source adiabatic modes with proportionality coefficient

$$N_{,\psi} \sim \frac{H\dot\psi_{\rm sr}}{\dot\phi_{\rm sr}^2}, \tag{D2}$$

then the $r < \epsilon^2$ bound still remains in place. (Comma denotes partial derivative.) This is because for $\dot\psi_{\rm sr} \ll \dot\phi_{\rm sr}$,

$$\pi \simeq \frac{\pi_c}{\dot\phi_{\rm sr}} \simeq \frac{\delta\phi}{\dot\phi_{\rm sr}} + \frac{\dot\psi_{\rm sr}}{\dot\phi_{\rm sr}^2}\delta\psi; \tag{D3}$$

hence, the contribution (D2) of $\delta\psi$ to ζ would be of the same order of magnitude as if $\zeta_{,\sigma}$ were absent from (40).

Let us see when a contribution of order (D2) should be expected. The slow-roll condition $3H\dot\psi_{\rm sr} \simeq -V_{,\psi}$ implies that fluctuations in ψ contain energy $\delta\rho_\psi \sim H\dot\psi_{\rm sr}\delta\psi$. Therefore, if these fluctuations perturb the kinetic energy of the inflaton $\dot\phi_{\rm sr}^2/2$, they result in (D2). This will be the case, for instance, if $\delta\psi$ fluctuations become massive as function of ϕ and before the end of inflation since as in the examples studied in the text the available source of energy during inflation is the kinetic energy of the inflaton. Immediately after the transition we expect $\delta\dot\phi^2 \sim \delta\rho_\psi$. On the other hand, if the ϕ and ψ sectors are two completely decoupled slow-rolling sectors, we expect

$$N_{,\psi} \sim \frac{U(\psi)}{V(\phi)}\frac{H}{\dot\psi}. \tag{D4}$$

This is $\epsilon_\phi/\epsilon_\psi$ times (D2), where $\epsilon_\phi = \dot\phi_{\rm sr}^2/2V(\phi)$ and similarly $\epsilon_\psi = \dot\psi_{\rm sr}^2/2U(\psi)$. This would lead to $r_{\rm max} \sim \epsilon_\psi^2$. For ψ to slowly roll ϵ_ψ must be small though perhaps it can be larger than ϵ.[10]

Therefore, to avoid (D2) one way is to couple the two sectors in such a way that the fractional contribution of $\delta\psi$ to expansion history be given by $\delta\rho_\psi/\rho_{\rm tot}$. This would be the case if the reheating surface is completely determined by ϕ independently of the value of ψ. As a toy model consider a potential

$$U(\phi,\psi) = \theta(\phi_0 - \phi)V(\phi,\psi), \tag{D5}$$

where $V(\phi,\psi)$ satisfies slow-roll condition for both fields. At $\phi = \phi_0$ all of the potential energy abruptly converts into the kinetic energy of ϕ at ϕ_0. The fluctuations of ψ induce fluctuations in $\dot\phi$ but since now $\dot\phi^2/2$ contains most of the energy density of the Universe the fractional variations do not have the previous $1/\epsilon$ enhancement.

The fluctuations $\delta\rho_\psi/\rho_{\rm tot}$ also lead to variations in the number of e-folds (D1) which can be approximated as

$$N \simeq \frac{1}{M_{\rm Pl}^2}\int^{\phi_0}\frac{V}{V_{,\phi}}d\phi. \tag{D6}$$

[10]This scenario has been studied in more detail in [21]. To compare note that their ΔN would be of the order of our $1/\epsilon_\psi$. In particular, ΔN should be long enough for the cosmologically relevant range of modes to cross the horizon.

For concreteness suppose $V_{,\phi,\psi} = 0$. Then,

$$N_{,\psi} \sim N_e \frac{V_{,\psi}}{V} \sim N_e \epsilon \frac{H\dot{\psi}_{\rm sr}}{\dot{\phi}^2_{\rm sr}}, \qquad \text{(D7)}$$

where $N_e \sim 60$ is the total number of e-folds from the horizon crossing of $\delta\psi$ fluctuations until the end of inflation. Hence, if $N_e \epsilon \ll 1$ the contribution of $\delta\psi$ fluctuations to scalar power can be suppressed.

[1] M. Zaldarriaga and U. Seljak, An all sky analysis of polarization in the microwave background, Phys. Rev. D **55**, 1830 (1997).

[2] M. Kamionkowski, A. Kosowsky, and A. Stebbins, Statistics of cosmic microwave background polarization, Phys. Rev. D **55**, 7368 (1997).

[3] K. N. Abazajian, K. Arnold, J. Austermann, B. A. Benson, C. Bischoff, J. Bock, J. R. Bond, and J. Borrill, Inflation physics from the cosmic microwave background and large scale structure, Astropart. Phys. **63**, 55 (2015).

[4] L. Senatore, E. Silverstein, and M. Zaldarriaga, New sources of gravitational waves during inflation, J. Cosmol. Astropart. Phys. 08 (2014) 016.

[5] L. Sorbo, Parity violation in the cosmic microwave background from a pseudoscalar inflaton, J. Cosmol. Astropart. Phys. 06 (2011) 003; L. Cook and J. L. Sorbo, Particle production during inflation and gravitational waves detectable by ground-based interferometers, Phys. Rev. D **85**, 023534 (2012); **86**, 069901(E) (2012).

[6] N. Barnaby, J. Moxon, R. Namba, M. Peloso, G. Shiu, and P. Zhou, Gravity waves and non-Gaussian features from particle production in a sector gravitationally coupled to the inflaton, Phys. Rev. D **86**, 103508 (2012); S. Mukohyama, R. Namba, M. Peloso, and G. Shiu, arXiv:1405.0346.

[7] S. R. Behbahani, A. Dymarsky, M. Mirbabayi, and L. Senatore, (Small) resonant non-Gaussianities: Signatures of a discrete shift symmetry in the effective field theory of inflation, J. Cosmol. Astropart. Phys. 12 (2012) 036.

[8] E. Silverstein and A. Westphal, Monodromy in the CMB: Gravity waves and string inflation, Phys. Rev. D **78**, 106003 (2008); L. McAllister, E. Silverstein, and A. Westphal, Gravity waves and linear inflation from axion monodromy, Phys. Rev. D **82**, 046003 (2010); R. Flauger, L. McAllister, E. Pajer, A. Westphal, and G. Xu, Oscillations in the CMB from axion monodromy inflation, J. Cosmol. Astropart. Phys. 06 (2010) 009.

[9] D. Green, B. Horn, L. Senatore, and E. Silverstein, Trapped inflation, Phys. Rev. D **80**, 063533 (2009); J. Cosmol. Astropart. Phys. 06 (2010) 009.

[10] G. D'Amico, R. Gobbetti, M. Kleban, and M. Schillo, Unwinding inflation, J. Cosmol. Astropart. Phys. 03 (2013) 004.

[11] N. Barnaby and M. Peloso, Large Non-Gaussianity in Axion Inflation, Phys. Rev. Lett. **106**, 181301 (2011).

[12] C. Cheung, P. Creminelli, A. L. Fitzpatrick, J. Kaplan, and L. Senatore, The effective field theory of inflation, J. High Energy Phys. 03 (2008) 014.

[13] G. D'Amico and M. Kleban, Phys. Rev. Lett. **113**, 081301 (2014).

[14] L. Senatore, in *Proc. of Burke Institute Workshop on Primordial Gravitational Waves and Cosmology*, Caltech, May 2014, https://burkeinstitute.caltech.edu/workshops/BICEP2; also in *Proc. of ICTP Workshop on Frontiers of New Physics Colliders and Beyond*, June 2014, http://indico.ictp.it/event/a13203/session/3/contribution/8/material/2/.

[15] D. Baumann, D. Green, and R. A. Porto, arXiv:1407.2621.

[16] P. Creminelli, J. Gleyzes, J. Norea, and F. Vernizzi, J. Cosmol. Astropart. Phys. 01 (2015) 016.

[17] W. D. Goldberger and I. Z. Rothstein, An Effective field theory of gravity for extended objects, Phys. Rev. D **73**, 104029 (2006); Phys. Rev. Lett. **113**, 231301 (2014).

[18] S. Weinberg, *Gravitation and Cosmology* (Wiley, New York, 1972).

[19] L. Senatore and M. Zaldarriaga, The effective field theory of multifield inflation, J. High Energy Phys. 04 (2012) 024.

[20] O. zsoy, K. Sinha, and S. Watson, arXiv:1410.0016.

[21] R. Z. Ferreira and M. S. Sloth, arXiv:1409.5799.

[22] M. Biagetti, E. Dimastrogiovanni, M. Fasiello, and M. Peloso, J. High Energy Phys. 12 (2014) 139.

[23] P. Adshead, E. Martinec, and M. Wyman, Phys. Rev. D **88**, 021302 (2013).

[24] P. Creminelli, A. Nicolis, L. Senatore, M. Tegmark, and M. Zaldarriaga, Limits on non-Gaussianities from WMAP data, J. Cosmol. Astropart. Phys. 05 (2006) 004.

[25] L. Senatore, K. M. Smith, and M. Zaldarriaga, Non-Gaussianities in single field inflation and their optimal limits from the WMAP 5-year data, J. Cosmol. Astropart. Phys. 01 (2010) 028.

[26] S. Klein, Suppression of Bremsstrahlung and pair production due to environmental factors, Rev. Mod. Phys. **71**, 1501 (1999).

[27] A. V. Koshelkin, Long wave asymptote for the Landau-Pomeranchuk-Migdal effect, J. Phys. A **37**, 1051 (2004).

[28] M. Sasaki and E. D. Stewart, A General analytic formula for the spectral index of the density perturbations produced during inflation, Prog. Theor. Phys. **95**, 71 (1996).