Isocurvature perturbations and non-Gaussianity of gravitationally produced nonthermal dark matter
Daniel J. H. Chung and Hojin Yoo
Phys. Rev. D 87, 023516 — Published 22 January 2013
DOI: 10.1103/PhysRevD.87.023516
Isocurvature Perturbations and Non-Gaussianity of Gravitationally Produced Nonthermal Dark Matter

Daniel J. H. Chung∗ and Hojin Yoo†
Department of Physics, University of Wisconsin-Madison
1150 University Avenue, Madison, WI 53706, USA

Gravitational particle production naturally occurs during the transition from the inflationary phase to the non-inflationary phase. If the particles are stable and very weakly interacting, they are natural nonthermal dark matter candidates. We show that such nonthermal dark matter particles can produce local non-Gaussianities large enough to be observed by ongoing and near future experiments without being in conflict with the existing isocurvature bounds. Of particular interest is the fact that these particles can be observable through local non-Gaussianities even when they form a very small fraction of the total dark matter content.

I. INTRODUCTION

Standard slow-roll inflationary models with a single dynamical field degree of freedom (e.g. see the review article [1]) cannot generate large local non-Gaussianities (NGs) [2–5], which have been widely discussed and speculated upon in the context of the cosmic microwave background (CMB) data (e.g. [6–12]) and large scale structure data (e.g. [13–22]). Many multifield mechanisms have been proposed to generate observably large local NGs (e.g. [23–47]). Most of these models utilize coherent condensate field degrees of freedom instead of incoherent many-particle states.

In this paper, we explore the possibility that nonthermal dark matter (DM) particles gravitationally produced during the phase transition out of the quasi-de-Sitter phase of inflation [48, 49] generate observably large NGs. These dark matter particles can be viewed as the remnants of de Sitter (dS) temperature driven radiation during inflation, and no non-standard ingredients are needed for the inflationary scenario for the purposes of this work. The only nontrivial model requirement is that the dark matter either be very heavy and/or superweakly interacting.

This class of scenarios effectively possesses only three important independent dimensionful parameters: the Hubble expansion rate during inflation, the reheating temperature, and the dark matter mass. Hence, the physics is dominantly controlled by only two of these parameters since the third converts the other two into dimensionless numbers. We choose these to be the dark matter mass \( m_X \) and the reheating temperature \( T_{RH} \). The existing cosmological data constraining the isocurvature perturbation amplitude and the dark matter abundance place bounds on the allowed parametric range for these parameters. We find that to generate large observable local non-Gaussianities characterized by an effective \( f_{NL} \) parameter of around 30, there is an upper bound of \( m_X \lesssim 4H_e \), where \( H_e \) is the expansion rate at the end of inflation. We also find that \( f_{NL} \) will be suppressed if \( T_{RH} \gtrsim 10^6 \) GeV if the dark matter is absolutely stable. Somewhat surprisingly, even when the \( X \) particles make up a small fraction of the total dark matter content while thermal relics make up the rest of the dark matter, observably large non-Gaussianities may be imprinted.

The isocurvature perturbations in this class of scenarios have been studied previously [50]. We note that this was also briefly considered in [23], which arrived at a pessimistic conclusion. However, that paper did not consider the model as carefully as [50], which reached a more realistic conclusion regarding the viability of such scenarios. The purpose of this paper is to point out that within this framework, large local non-Gaussianities can be generated with a single \( O(10^{-1}) \) tuning of the dark matter mass.

The order of presentation is as follows. In Sec. II, we discuss the class of dark matter and inflation models for which the current non-Gaussianity computation is relevant. In Sec. III, we present a computation of the two-point function, including the cross-correlation function between the isocurvature and the curvature components. The computation of the bispectrum and a presentation of detailed arguments as to how the local non-Gaussianity can be large is given in Sec. IV. In Sec. V, we check our general analytic arguments by computing in detail numerically the observables in the context of \( m^2 \phi^2 / 2 \) inflationary model. We then close the main body of the paper with a summary and conclusions in Sec. VI. In Appendix A, we present an analytic approximation to the mode function during inflation that accounts for the small deviation from the pure dS phase. Finally, in Appendix B, we justify how the effective classical background variable about which the classical perturbations are defined is

∗Electronic address: danielchung@wisc.edu
†Electronic address: hyoo6@wisc.edu
1 These particles are sometimes called gravitationally produced WIMPZILLAs.
2 This mass scale has part of its origins from the maximum dark matter abundance today.
given by the expectation value of the quantum operator. Throughout this work, our metric convention is $\left(+,-,-,-\right)$.

II. CLASS OF DARK MATTER AND INFLATIONARY MODELS

We begin by defining the class of dark matter and inflationary models considered in this work. One requirement is that the dark matter field $X$ be sufficiently long lived to be a viable dark matter candidate. Since we are considering an isocurvature dark matter scenario, another requirement is that $X$ is very weakly interacting with thermalized Standard Model (SM) particles that are assumed to arise from the inflaton decay chain. Although it is straightforward to include dark matter interactions that allow transformations to different particles, we will assume that the self-annihilation (or coannihilation) interactions of the dark matter are too weak to change the dark matter number density appreciably after it is produced at the end of inflation. Next, we note that even if $X$ is sufficiently weakly interacting as not to thermalize, it may be a byproduct of a slow-roll inflaton decay with a strength larger than gravitational strength. In such situations, the isocurvature nature of the $X$ particles produced gravitationally will be made impure by the inflaton decay contribution. To keep the analysis simple for the purposes of this paper, we will assume that the decay contribution is negligible. Finally, we will assume that $X$ is a boson minimally coupled to gravity. We will explore the complexities that arise when some of these requirements are relaxed in future work.

The simplest model that satisfies the above criteria has two scalar fields minimally coupled to gravity as follows:

$$S = \int d^4 x \sqrt{g} \frac{1}{2} \left[ (\partial \phi)^2 - 2U(\phi) + (\partial X)^2 - m_X^2 X^2 \right] + S_{\text{SM}}[g_{\mu\nu}, \{ \psi_{\text{SM}} \}] + S_{\text{RH}}[g_{\mu\nu}, \phi, \{ \psi_{\text{SM}} \}],$$

(1)

where $S_{\text{SM}}$ is the SM sector, $S_{\text{RH}}$ is responsible for reheating, and $\phi$ is the inflaton realizing a slow-roll inflationary scenario with $U(\phi) = m^2 \phi^2/2$. We note that since particle production during the dS to non-dS phase leads to a dS horizon temperature population of SM one-particle states, one might naively expect a minimum reheating temperature of $T_{\text{RH}} \gtrsim H_e/\left(2\pi \right)$, where $H_e$ is the Hubble expansion rate the end of inflation. However, this is incorrect since during the coherent oscillations period between the end of the inflation and the radiation dominated epoch, the SM radiation dilutes with respect to the inflaton energy density.

Since we will carry out numerical analysis of the mode equations, this simple model is useful. Furthermore, it is clear that the results generalize to a large class of models where the interactions are very small. We note also that the requirement of the interactions being weak enough to avoid thermalization does not require a particularly stringent limit on the interaction couplings. For example, for typical $O(1)$ coupling strengths for self-annihilations, it is well known that large values of $m_X$ will naturally lead to the nonthermal DM behavior desired in this paper if

$$\left(\frac{200\text{TeV}}{m_X}\right)^2 \left(\frac{T_{\text{max}}}{m_X}\right) \lesssim 1,$$

(2)

where $T_{\text{max}} \gtrsim T_{\text{RH}}$ is the maximum effective temperature reached during reheating [51]. Of course, one possible model-building obstacle with large masses is that in situations with accidental global symmetries protecting the stability of the particles, higher-dimension operators must be suppressed to avoid early decay. Nonetheless, many viable beyond SM (BSM) models that contain superheavy DM candidates have been proposed [52–64].

Given that the $X$ particles have negligible non-gravitational interactions and minimal couplings to gravity, they can only be produced gravitationally or through initial conditions. We consider the dynamics of an inflationary patch whose Bunch-Davies vacuum [65] satisfies

$$\lim_{k \to \infty} \hat{a}_k \langle BD, 0 \rangle = 0,$$

(3)

where $\hat{a}_k$ is the annihilation operator associated with the curvature perturbations which is approximately dominated by the inflaton. This vacuum is also assumed to satisfy the “no-particle” condition of the adiabatic vacuum during inflation [48, 66–68]:

$$\hat{a}_k |BD,0\rangle = 0,$$

(4)

where $\hat{a}_k$ is the annihilation operator associated with the $X$ field. The stress-energy tensor is renormalized such that

$$\langle BD, 0 | T_{\mu \nu}^{(X, \text{ren})} | BD, 0 \rangle = 0,$$

(5)

which in practice is accomplished by the normal ordering of the creation-annihilation operators in the adiabatic vacuum basis. This means that the classical initial condition dependent DM density vanishes.

Nonetheless, because the transition from quasi-dS phase of inflation to the non-dS phase after inflation represents a non-adiabatic transition, it is well known that non-negligible particle production occurs through Bogoliubov mixing of the creation-annihilation operators, giving rise to a significant DM abundance today [48, 69]. The physics of this particle production mechanism is similar to that of Hawking radiation. In the intermediate mass case where $m_X \sim H_e$ with minimal gravitational couplings,
we find numerically that the $X$ energy density at the end of inflation can be approximated as $\rho_X(\tau_e) \approx 10^{-2}H_e^4$, which leads to the relic abundance of $X$ particles today to be

$$\Omega_X h^2 \approx 10^{-1} \left( \frac{H_e}{10^{12} \text{ GeV}} \right)^2 \left( \frac{T_{RH}}{10^9 \text{ GeV}} \right).$$

What is interesting about this scenario is that although the classical picture of particle production occurs at the end of inflation, the correlations that are relevant at the CMB scale are set long before the bulk of the particle production occurs. This is intuitively self-consistent from the Heisenberg time-energy uncertainty considerations. Although Eq. (6) yields the simplest possible scenario, we later generalize the situation to the case of mixed dark matter contributions in which the total cold dark matter (CDM) abundance is given by

$$\Omega_{CDM} = \Omega_{therm} + \Omega_X,$$

where $\Omega_{therm}$ are thermal relics that have only adiabatic perturbations and $\Omega_X$ are relics that have dominantly isocurvature perturbations. In this case, we define

$$\omega_X = \frac{\Omega_X}{\Omega_{CDM}},$$

and will scale some of our computations to generalize our results to a wider class of scenarios.

We now comment further on the inflationary model relevant for the above DM scenario. The main features of the inflationary model that are numerically important for the isocurvature and non-Gaussianity analyses are $H_e$ (the Hubble expansion rate when the modes of interest leave the horizon), $H_c$ and $T_{RH}$ (the reheating temperature). As we will see, the primary role of $H_e$ is to determine the spectral index of the isocurvature spectrum, $H_c < H_e$ controls the particle production, and $T_{RH}$ partially controls the map between the comoving wave vector and the physical momentum. We assume that there are curvature perturbations from the inflaton sector with the right magnitude to approximately explain the CMB spectrum. As we will see and as is well known, the current observational limits require that the isocurvature contribution is subdominant.

Finally, it has been noted [23] that this class of models suffer from the boundary condition of $\langle \hat{X} \rangle = 0$ being an unnatural expansion point of the fluctuations of $X$ for $m_X \ll H$. Although it certainly is true that in this limit the $H$ dependent radiative corrections lift the flatness of the potential, there are no radiative tadpoles that are generated. Furthermore, although it is true that once the non-decaying mode decoheres as the wavelength is stretched outside the horizon, acting like a classical background, there is no strong tuning in choosing a patch that has $\langle \hat{X} \rangle = 0$ as long as the inflation did not last many orders of magnitude in efolds longer than what is needed to explain the CMB data. Indeed, because of the slight blue tilt, there is no infrared divergence in this class of models. As we will see, the blue tilt of significant isocurvature amplitudes is still compatible with observations.

For completeness, let us also explicitly state the cutoffs of our theory. Since the spectral index of the correlator (29) is related to $3 - 3\sqrt{1 - 4m_X^2/9H^2}$ and our scenario has $m_X^2 > 0$, the correlator has a blue spectrum (as we will see explicitly in Sec. III), and thus the loop integrals (see for example Eqs. (28) and (33)) for the two-point function of isocurvature is independent on IR cut-off. Therefore, an IR cut-off is not necessary unlike the correlators in the massless case [70]. The UV cutoff of our theory is set to be the horizon scale at the end of inflation: $k_{UV} = H_e a_e$.

Finally, we note that we use the scalar metric perturbation parameterization

$$ds^2 = (1 + E)dt^2 - 2a\partial_i F dt dx^i - a^2 [\delta_{ij} + A \delta_{ij} + \partial_i \partial_j B] dx^i dx^j.$$

We make the usual choice for the gauge invariant variable that describes the inflaton dynamical degree of freedom:

$$\zeta = A - H \frac{\delta \rho_\phi}{\dot{\varphi}}.$$

In terms of this variable, the nearly scale invariant slow-roll inflaton power spectrum is given as

$$\Delta_\zeta^2(k) = \frac{k^3}{2\pi^2} P_\zeta(k),$$

where to leading order in the slow-roll parameter

$$\epsilon(\phi_e) = \frac{M_p^2}{2} \left( \frac{U'(\phi_e)}{U(\phi_e)} \right)^2,$$
we have
\[ P_\gamma(k) = \frac{1}{12k^3 \epsilon(k)} \frac{U(\phi_k)}{M_p^4}. \] (13)

In the above, \( \phi_k \) denotes the field value when the mode \( k \) leaves the horizon. We will give more details about the \( X \) fluid variable when discussing the two-point function in the next section.

In summary, the class of dark matter models that is relevant for this paper corresponds to cases with gravitationally produced bosonic dark matter that never fully thermalizes with the reheating radiation produced from the inflaton decay. The slow-roll inflationary model produces the dominant curvature perturbation spectrum and couples to the dark matter sector only gravitationally.

### III. TWO-POINT FUNCTION

Although isocurvature perturbations have been previously computed for this class of models [50], here we redo the analysis with more careful attention to cross-correlations between the curvature perturbations and the isocurvature perturbations because the observational constraints have become increasingly stringent.

To begin, consider the energy momentum tensor of \( X \):
\[ T_{\mu\nu}^{(X)} = \partial_\mu X \partial_\nu X - g_{\mu\nu} \left[ \frac{1}{2} \partial_\alpha X \partial_\beta X - V(X) \right] \] (14)
where \( V(X) = m_X^2 X^2 / 2 \). Comparing this to
\[ T_{\mu\nu}^{(perfect \ fluid)} \equiv (\rho^{(p)}_X + P^{(p)}_X) u_\mu u_\nu - g_{\mu\nu} P^{(p)}_X, \] (15)
we see that if we define [71]
\[ u_\mu = \frac{\partial_\mu X}{\sqrt{g^\alpha\beta \partial_\alpha X \partial_\beta X}}, \] (16)
we can satisfy the equality
\[ T_{\mu\nu}^{(X)} = T_{\mu\nu}^{(perfect \ fluid)}, \] (17)
if
\[ \rho^{(p)}_X \equiv u_\mu u_\nu T_{\mu\nu}^{(perfect \ fluid)} = \frac{1}{2} \partial_\alpha X \partial_\alpha X + V(X) \] (18)
and
\[ P^{(p)}_X \equiv \frac{1}{2} \partial_\alpha X \partial_\alpha X - V(X). \] (19)

We note that to identify Eq. (16) with the fluid velocity, \( \partial_\mu X \) has to be timelike. This is consistent with the fact that any wave packet made of on-shell 1-particle states can be decomposed in terms of mode functions characterized by timelike 4-momenta. Unlike the coordinate dependent \( T_{00}^{(X)} \), \( \rho^{(p)}_X \) is a diffeomorphism scalar. We also note that even though Eq. (18) looks like it has the wrong sign between the \( (\partial_0 X)^2 \) and \(|\vec{\nabla} X|^2\), the sign is correct and \( \rho^{(p)}_X \) is positive definite whenever the fluid interpretation is valid (whenever \( \partial_\mu X \) is timelike).

We now quantize \( X \) by promoting it to an operator \( \hat{X} \). As explained in Appendix B, this allows us to identify
\[ \rho^{(p)}_0 = \langle : \hat{\rho}^{(p)}_X : \rangle, \] (20)
where the normal ordering is with respect to the operators defining the \( \hat{X} \) vacuum state during the quasi-dS era. After Bogoliubov transforming \( : \hat{\rho}^{(p)}_X : \) to operators of 1-particle states relevant for non-dS spacetime, \( \langle : \hat{\rho}^{(p)}_X : \rangle \) will develop a nonzero value at that later time. We note that \( \langle : \hat{\rho}^{(p)}_X : \rangle \) is homogeneous as long as the vacuum state is spatially translation invariant. (Here
the inflaton/scalar perturbations are treated as operators, which means that as long as the vacuum governing these are spatially translation invariant, \( \langle \delta \hat{\rho}_X^{(p)} \rangle \) will be spatially translation invariant as well.) Next, we consider the semi-classical variable
\[
\delta \rho_X^{(p)}(t) \equiv \rho_X^{(p)}(t) - \rho_0^{(p)}(t).
\]
(21)

We can then define the usual fluid variable associated with \( \rho_X \)
\[
\zeta_X = \frac{A}{2} - H \frac{\delta \rho_X^{(p)}}{\delta \rho_0^{(p)}(t)},
\]
(22)

where we parameterize the spatial scalar metric perturbation as be \( h_{ij} = -a^2(t) \left( A \delta_{ij} + \partial_i \partial_j B \right) \) with \( \delta_{\mu\nu} = \text{diag}\{1, -a^2, -a^2, -a^2\}. \) Under the diffeomorphism \( t \to t - \xi^0 \), we have
\[
A \to A + 2H \xi^0 \quad \text{(23)}
\]
\[
\delta \rho_X^{(p)} \to \delta \rho_X^{(p)} + \xi^0 \frac{d}{dt} \rho_0^{(p)}, \quad \text{(24)}
\]

which makes \( \zeta_X \) first order gauge invariant, as expected. Similarly, we can define gauge invariant variables \( \zeta, \zeta_\phi, \) and \( \zeta_R \) associated with total energy density \( \rho \), inflaton energy density \( \rho_\phi \), and radiation energy density \( \rho_R \), respectively. Since the gauge invariant isocurvature variable that describes the difference between the dark matter and the radiation, which is an inflaton descendant, is conventionally defined as [72]
\[
\delta_{\text{X}} = 3 (\zeta_X - \zeta_R). \quad \text{(25)}
\]

Note that the inflaton eventually decays into radiations and matters, while curvature perturbation \( \zeta \approx \zeta_\phi \) remains a conserved quantity on long wavelengths even after the inflaton decay and it is adiabatically matched to \( \zeta_R \approx \zeta \), since the dark matter is energetically subdominant at the primordial epoch.

In the comoving gauge defined by the coordinate system in which the inflaton fluctuations vanish (i.e., \( \delta \phi = 0 \)), we have \( \zeta_R = \zeta_\phi = A^{(c)}/2 \). Hence,
\[
\delta_{\text{X}} = \delta_X^{(c)} = \frac{\delta \rho_X^{(p)(c)}}{\rho_0^{(p)}(t)}, \quad \text{(26)}
\]

where the \((c)\) superscript refers to the comoving gauge quantity and we have used the fact that \( \rho_0 \) behaves as \( a^{-3} \) once the Hubble scale is sufficiently smaller compared to the mass \( m_X \ll H \). Therefore, correlator combinations involving \( \delta X^{(c)} \) and \( A^{(c)} \) are of primary physical interest. To compute them, we quantize \( \delta \rho_X^{(p)(c)} \) by promoting \( \rho_X^{(p)(c)} \to \hat{\rho}_X^{(p)(c)} \) (through the quantization of \( \hat{X} \)) and promoting \( \rho_0^{(p)}(t) \) to an identity operator (since this was already defined semi-classically as a matrix element according to Eq. (B23)) as follows:
\[
\delta \hat{\rho}_X^{(p)(c)} : = \hat{\rho}_X^{(p)(c)} - \hat{1} \rho_0^{(p)}(t). \quad \text{(27)}
\]

This can be used to compute \( \langle \delta \hat{\rho}_X^{(p)(c)}(t, \vec{r}) \delta \hat{\rho}_X^{(p)(c)}(t, 0) \rangle \) and then can be divided by \( \rho_0^2(t) \) after this quantity settles down to give an expression for \( \langle \delta X^{(c)}(t, \vec{r}) \delta X^{(c)}(t, 0) \rangle \).

Diagramatically, the 2-point function is as shown in Fig. 1. Hence, the 2-point power spectrum is
\[
\Delta_{\text{X}}^2(k) = \frac{k^3}{2\pi^2} \int d^3r e^{-i\vec{k} \cdot \vec{r}} \langle \delta X^{(c)}(t, \vec{r}) \delta X^{(c)}(t, 0) \rangle \approx \frac{k^3}{2\pi^2} \frac{2}{\Lambda_{\text{IR}}} \int_{\Lambda_{\text{IR}}}^a d^3k_1 \frac{1}{(2\pi)^3} P_X(k_1) P_X(|\vec{k} - \vec{k}_1|),
\]
(28)
in which
\[
P_X(k) = \frac{m_X^2}{2\rho_0^{(p)}} |X_k|^2, \quad \text{(29)}
\]
Figure 1: A diagrammatic representation of the 2-point function of a composite operator $\delta_X$. The open circle corresponds to external momentum flowing out, and the shaded circle to external momentum flowing in. The slashes on the legs indicate that the $X$ propagators are on-shell.

$\delta_X^{(c)}$ is approximated as $m_X^2\bar{X}^2/2\rho_1^{(p)}$, and $X_k$ is the solution to the mode equation in Appendix A. For $m_X/H < 3/2$ and $k/(aH) \ll 1$, we can express $P_X(k)$ approximately as

$$P_X(k) = A_X \left( \frac{k}{k_0} \right)^{\gamma_X(H_e)} k^{-3},$$

(30)

where

$$\gamma_X(H) = 3 - 3\sqrt{1 - \frac{4m_X^2}{9H^2}} > 0.$$  

(31)

In [50], $H_e$ is allowed a wave vector dependence; however, this effect is a subdominant correction to the already small $(m_X/H_e)^2$. It is important to note that for the parameter range of interest $\gamma_X \ll 1$ such that $k^3 P_X$ is nearly scale invariant, the amplitude $A_X$ is given by the approximate formula

$$A_X \sim \frac{10^2 \Gamma(\frac{3}{2} - \frac{\gamma_X}{2})}{2\pi^{3/2}} m_X H_e^2 \exp \left[ \frac{1}{\varepsilon} \left\{ \gamma_X(H_\ast) - \gamma_X(H_e) + 3\ln \left( \frac{1 - \frac{3}{2}\gamma_X(H_\ast) + \frac{H_e}{m_X}}{1 - \frac{3}{2}\gamma_X(H_e) + \frac{H_e}{m_X}} \right) \right\} \right],$$

(32)

where $\varepsilon = M_p^2 \left( \frac{U(\phi)}{U(0)} \right)^2$ is the usual inflaton slow-roll parameter. The complicated exponential factor arises from considering the time evolution of the mode function $X_k$ to a time beyond the time when $k/a < \varepsilon/H_\ast$. We see that with a tuning of the mass parameter to $O(0.1)$ precision, $A_X$ can be a small number despite the exponential factor containing $\varepsilon^{-1} \gg 1$. For the dark matter abundance to be compatible with cosmological observations, it is important that $H_\ast > H_e$ while $m_X/H_e \sim O(1)$. Using these approximations and Eq. (30), we find that Eq. (28) yields

$$\Delta^2_{\delta_X}(k) \approx \frac{k^3}{2\pi^2} A_X \left( \frac{k}{k_0} \right)^{2\gamma_X} \frac{1}{k^3} \times 2\int_{\Lambda_{IR}/k}^{aH_e/k} \frac{d^3 u}{(2\pi)^3} u^{\gamma_X - 3} |1 - i\varepsilon|^{-\gamma_X - 3}$$

$$\approx \frac{4}{\gamma_X} \left[ 1 - \frac{\Lambda_{IR}}{k} \right]^{\gamma_X} \frac{A_X}{2\pi^2} \left( \frac{k}{k_0} \right)^{\gamma_X/2}$$

$$\approx \frac{4}{\gamma_X} \left( \frac{k^3}{2\pi^2} P_X(k) \right)^2,$$

(33)

where we have used that $\gamma_X \ll 1$ in the second line. Note that IR cutoff $\Lambda_{IR}$ dependence does not appear because of the blueness of $P_X$ (i.e. $\gamma_X(H_\ast) > 0$), and $\Delta^2_{\delta_X}$ inherit a spectrum of $k^3 P_X(k)$. On the other hand, we will see in the next section that the key to obtaining large non-Gaussianities is that $(k^3 P_X(k))^2$ is much smaller than $k^3 P_X(k)$.

Thus far, we have focused on only the $X$ isocurvature perturbations. In the mixed scenario described near Eq. (7), we can rescale Eq. (33) to obtain the total CDM isocurvature perturbations as

$$\Delta^2_{\delta}(k) = \omega_X^2 \Delta^2_{\delta_X}(k),$$

(34)

since the rest of the dark matter contribution has no isocurvature perturbations.
As we will see in the next section, the bispectrum is maximized in the parameter region in which \( \Delta_\delta^2 (k) \) is large. The observational bound on \( \Delta_\delta^2 \) is stringent unless the cross-correlation between curvature and isocurvature is negligible, i.e. \[ \Delta_2^{\zeta \delta S} \ll \Delta_\zeta \Delta_\delta S, \] (35)

where \( \Delta_\delta S \) is the power spectrum of the cross-correlation. The left hand side \( \langle \hat{\zeta} \hat{\delta} S \rangle \) corresponding to the diagram shown in Fig. 2 can be computed using the in-in formalism \([2, 76]\) using the trilinear interaction Hamiltonian in the comoving gauge

\[
H_I(t) \equiv - \int d^3x a^3(t) \left[ T^{ij}_X(t, \vec{x}) a^2(t) \delta_{ij} \hat{\zeta}(t, \vec{x}) \right],
\] (36)

where \( T^{\mu \nu}_X \) is the stress energy tensor of \( X \). Note that other interaction Hamiltonian contributions are derivatively suppressed.

As will be shown elsewhere \([77]\), the curvature-isocurvature cross correlation is

\[ \beta \equiv \Delta_\zeta^2 \Delta_\delta S \sim 2.5 \times 10^{-5}, \] (37)

which shows that the cross-correlation is negligible. This fact is understood by the soft-\( \zeta \) theorem\([2, 78]\), which allows to factorize \( \langle \hat{\delta} \hat{\rho}_X \hat{\zeta} \rangle \) from \( \langle \hat{\delta} \hat{\rho}_X \hat{\zeta} \rangle \), i.e.

\[
\langle \delta \hat{\rho}_X \hat{\zeta} \rangle \sim \langle \hat{\xi} \hat{\zeta} \rangle \frac{\partial}{\partial \ln a} \langle \hat{\rho}_X \rangle,
\]

up to a momentum conserving delta function. Physically, the curvature perturbation \( \zeta \) can affect the energy density \( \rho_X \) and generate correlation only at its horizon crossing, because after the perturbation \( \zeta \) crosses the horizon and then freezes, it can be effectively treated as a gauge mode, which corresponds to the spatial dilation in the general coordinate transform.

Because the isocurvature is of the uncorrelated type in this scenario, the current observational bound on the adiabaticity in terms of the parameter \( \alpha \) from WMAP, BAO, and SN combined \([7, 79]\) becomes

\[ \alpha = \frac{\Delta_\delta(k_0)}{\Delta_\zeta(k_0) + \Delta_\delta(k_0)} < 0.064, \] (39)

where \( k_0 = 0.002 \text{Mpc}^{-1} \). Note the contraint is considered under the assumption that the isocurvature perturbation is scale-invariant, while our model predicts the blue-tilted spectrum. However, we expect that the bound should not be either altered significantly (or more severely constrained) since the spectral index is less than 1.2 within the parameter range of interest. Furthermore, recent analyses \([74, 75]\) show that the best CMB fit favors a blue-tilted isocurvature perturbation, which relaxes the isocurvature constraints. Therefore, we will use the contraint as a conservative bound. In the next section, we will see that this implies the maximum \( f_{NL} \approx O(35) \).
In this section, we now compute the bispectrum \( B_{\delta_X}(\vec{p}_1, \vec{p}_2, \vec{p}_3) \) defined by

\[
(2\pi)^3 \delta^{(3)}(\sum \vec{p}_i) B_{\delta_X}(\vec{p}_1, \vec{p}_2, \vec{p}_3) = \omega_X^3 \int d^3 x_1 d^3 x_2 d^3 x_3 e^{-i \sum \vec{p}_i \cdot \vec{x}_i} \delta_X^{(c)}(t, \vec{x}_1) \delta_X^{(c)}(t, \vec{x}_2) \delta_X^{(c)}(t, \vec{x}_3),
\]

where we recall from the previous section that \( \delta_X^{(c)} \) in the comoving gauge can be identified with the gauge invariant quantity \( \delta_{\delta_X} \). With the diagrammatic representation given by Fig. 3, we find

\[
B_{\delta_X}(\vec{p}_1, \vec{p}_2, - (\vec{p}_1 + \vec{p}_2)) \approx 8 \omega_X^3 \int \frac{d^3 k}{(2\pi)^3} P_X(|\vec{k}|) P_X(|\vec{p}_1 - \vec{k}|) P_X(|\vec{p}_2 - \vec{k}|),
\]

which gives the analytic estimate of the primordial bispectrum for different triangle configurations fixed by \( \vec{p}_1 \) and \( \vec{p}_2 \).

As large non-Gaussianities are difficult to obtain in slow-roll inflation in the squeezed triangle limit, we will focus on that limit in this work. Using the one-pole approximation, we estimate the isocurvature bispectrum (written in a symmetrized form in the wave vectors) in the limit that one of the \(|\vec{p}_i|\) is much smaller than the others as follows:

\[
B_{\delta_X}(\vec{p}_1, \vec{p}_2, \vec{p}_3) \approx \frac{8}{\gamma_H(H_s)} \frac{p_{\min}^3}{2\pi^2} P_X(p_{\min}) \omega_X^3 \left[ P_X(p_1) P_X(p_2) + P_X(p_2) P_X(p_3) + P_X(p_3) P_X(p_1) \right]
\]

where \( p_{\min} = \min(\left|\vec{p}_i\right|) \). Since the energy density of \( X \) is quadratic in \( X \), the density correlator scales as \( P_X^2 \) and not just \( P_X \), which means that the coefficient \( p_{\min}^3 P_X(p_{\min}) \) is not quite the power spectrum. Nonetheless, because of the blueness of \( P_X \), \( p_{\min}^3 P_X(p_{\min}) \) can be strongly suppressed if \( \gamma_H(H_s) \) is large, and hence \( m_X \) has to be smaller than \( H_s \) for a non-negligible bispectrum. On the other hand, if \( m_X \) is too small compared to \( H_s \), we saw in the previous section that the isocurvature perturbations are larger than what is allowed by current data. Hence, there is a window for which the non-Gaussianities can be large and the isocurvature perturbations are consistent with the existing data.

To see why the bispectrum composed of quadratic fields is larger than the bispectrum composed of linear fields (such as for ordinary inflatons), consider the following ratio of isocurvature bispectrum to a fiducial local bispectrum defined with \( f_{NL} = 1 \)

\[
f_{NL}^S = \frac{B_S}{B_{\zeta}} |_{f_{NL}=1} = \frac{5}{6} \frac{p_{\min}^3}{2\pi^2} P_{\zeta}(p_1) P_{\zeta}(p_2) P_{\zeta}(p_3) P_{\zeta}(p_1) P_{\zeta}(p_2) P_{\zeta}(p_3),
\]

where \( P_{\zeta} \) is a two-point function of adiabatic perturbation. On large scales, the \( \delta_S \) contribution to the temperature perturbation \( \Delta T/T \) compared to the \( \zeta \) contribution is different by factor 2 due to the Sachs-Wolfe effect, i.e.

\[
\frac{\Delta T}{T} = -\frac{1}{5} \frac{\gamma}{\zeta} - \frac{2}{5} \delta_S.
\]
However, on scales smaller than $1/k_{eq}$, the transfer function of the isocurvature perturbations during radiation domination is suppressed by an additional factor $k_{eq}/k$ compared to that of adiabatic perturbations. One can understand this intuitively in the Newtonian gauge in terms of how isocurvature perturbations source the gravitational potential which in turn is proportional to the temperature perturbations. For adiabatic initial conditions, the gravitational potential on superhorizon scales are frozen. In contrast, isocurvature initial conditions effectively fix the superhorizon gravitational potential to be zero during the early radiation domination period when the dark matter energy density is negligible. As the fraction of dark matter energy density grows during the radiation domination period, the dark matter perturbations carrying the isocurvature information source the gravitational potential until the dark matter becomes the dominant energy component or until the modes enter the horizon. Thus, contrary to the temperature perturbations sourced by the adiabatic perturbations, those sourced by the isocurvature perturbations are proportional to the matter fraction at the horizon entry, which yields the additional suppression factor of $k_{eq}/k$ in the transfer function.

Because the isocurvature transfer function has different features from the adiabatic one, the isocurvature bispectrum leads to CMB temperature imprints distinct from that of the adiabatic bispectrum. A careful treatment of the transfer function incorporating the effects just discussed allows us to approximate a relationship on large scales [42] which can be summarized as

$$f_{NL} \approx 4f_{NL}^{\text{local}}$$

where $f_{NL}$ approximately coincides with the usual local $f_{NL}$ definition. Thus, in the squeezed triangle limit, $f_{NL}$ is analytically estimated to be

$$f_{NL} = \frac{80}{3} \frac{\alpha_X^3}{\gamma_X} \left( \frac{p_{min}^2}{2\pi^2} P_X(p_{min}) \right) \left( P_X(p_1)P_X(p_2) + P_X(p_2)P_X(p_3) + P_X(p_3)P_X(p_1) \right)$$

(46)

The large numerical factor of $80/3$ can be traced to the factor of $8 \times 8/2$ that arises from the product of the contraction permutations, the relative weight of the $\zeta$ and $\delta_X$ contributions to $\Delta T/T$, and the transfer function. As $p_{min}^2 P_X(p_{min})$ is suppressed partly from the smallness of the $P_X$ amplitude as well as the blue tilt (causing $p^3_{min}$ to be a suppressor), a large ratio of $P_X/P_\zeta$ is required for an unsuppressed $f_{NL}$. As we will now argue, a $P_X/P_\zeta$ ratio will arise from the fact that $\delta_X$ is quadratic in $X$.

To understand how to obtain a large $P_X/P_\zeta$, we first note that since $\delta_X^{(c)}$ is a quadratic functional of $X$ (see Eq. (33)), we have

$$P_X \approx \pi^2 k^{-3} \sqrt{\gamma_X \delta_X \frac{1}{\omega_X}}.$$  

(47)

If we define [80]

$$\frac{\alpha}{1 - \alpha} = \frac{\Delta_X^2}{\Delta_\zeta^2},$$

(48)

we find for $\alpha \ll 1$ that

$$\frac{P_X}{P_\zeta} \sim \frac{1}{2} \sqrt{\frac{\omega_X}{\Delta_\zeta}} \frac{1}{\omega_X} \sim 10^4 \sqrt{\gamma_X} \frac{1}{\omega_X}.$$  

(49)

Hence, if $\omega_X = 1$, as long as $\alpha \gg 10^{-8}/\gamma_X$, there is a large ratio of $P_X/P_\zeta$ because $\Delta_\zeta^2 \ll \Delta_X^2$ and $\delta_X \approx P_X^2$. Combining Eq. (46) and Eq. (49), we see

$$f_{NL} \sim 6 \times 10^3 \alpha^{3/2} \sqrt{\gamma_X}.$$  

(50)

Hence, if we can achieve $\alpha \approx 0.07$ [7] and $\gamma_X \sim 0.1$ in our model, we can achieve

$$f_{NL} \approx O(35).$$  

(51)

---

3 As mentioned in Ref. [42], the isocurvature bispectrum is enhanced by factor 4 instead of 8 due to the destructive contribution of the small scale modes. Although the scale-invariant isocurvature power spectrum has been used in their numerical analysis, their argument still applies to the slightly blue-tilted spectrum because the transfer function effect of small scales arises from $b^{(iso)}_{1l_1l_2l_3} \approx \frac{1}{2} f dk k^2 g^{(iso)}_{l_1l_2l_3}(k) f_{l_1l_2l_3}$, which is independent of the spectra index of isocurvature.

4 This relationship is obtained by comparison between the reduced bispectra $b^{(ad)}_{1l_1l_2l_3}$ and $b^{(iso)}_{1l_1l_2l_3}$ only at large angular scales ($l_1, l_2, l_3 \lesssim 10$). Thus, $f_{NL}^{\text{iso}}$ should not be interpreted as $f_{NL}^{\text{iso}}$ in Refs. [7, 79], which is obtained by the full analysis involving large and small scales.
on large scales. Although this $f_{NL}$ cannot be directly interpreted as $f_{NL}^{\text{local}}$ of Ref. [7, 79] (as discussed in footnote 4), a similar result which is obtained through a full numerical analysis in Ref. [73]

$$f_{NL}^{\text{local}} \approx 30 \left( \frac{\alpha}{0.07} \right)^{3/2}$$

(52)
gives a consistency check to our analytic argument leading to Eq. (51). This level of non-Gaussianity is clearly observable according to the forecasts of Planck and large scale structure experiments (see e.g. [81, 82]).Remarkably, this result is independent of $\omega_X$. Hence, even when the dark matter composition of the $X$ particles is small, non-Gaussianities associated with $X$ may be observable. We note that hidden in this analytic estimate is the implicit assumption that $\alpha$ can remain fixed as $\omega_X \to 0$. However, Eqs. (33) and (34) show perturbation theory would break down if $\omega_X \to 0$ limit is literally taken. For example, to have large local non-Gaussianities, the CDM isocurvature should be

$$\alpha \approx \frac{m_X}{m_\phi} \phi^2/2$$

model.

V. NUMERICAL RESULTS

As shown in Sec. IV, a local non-Gaussianity value of $f_{NL} \sim 30$ is achievable as long as the parameters of the model result in $\alpha \approx 0.07$ and $\gamma_X \approx 0.1$. The identification of these parameters requires the computation of $|X_k|^2$ associated with the mode function. Although an analytic approximation exists in Appendix A, it is still difficult to identify the parametric dependence because of the fact that the massive field $X$, unlike the variable $\zeta$, evolves during inflation even when the mode wavelength is superhorizon in magnitude. Its evolution depends on the variable $\gamma_X$, whose slow time dependence is difficult to account for analytically. Hence, to check the phenomenological viability of this isocurvature model, we now compute the necessary mode functions numerically within a $U(\phi) = m^2_\phi \theta^2/2$

In Fig. 4, we plot the allowed parameter space given the isocurvature perturbation and relic abundance constraints. We see that large local non-Gaussianities by the superheavy dark matter with a small isocurvature power spectrum are attainable in the vicinity of the thick dashed line of $\alpha = 0.07$. Thus, the upper left parameter region of the dashed line is ruled out due to overproduction of isocurvature perturbations. Furthermore, the region above of the solid line is excluded by the relic abundance condition $\Omega_X - \Omega_M \lesssim 0.2$. These conditions for large local non-Gaussianities yields a robust bound on the reheating temperature of

$$T_{RH} \lesssim 10^6 \text{ GeV},$$

(58)
as well as a bound on the mass of the superheavy dark matter of

$$\frac{m_X}{m_\phi} \lesssim 2.3,$$

(59)
Figure 4: The right bottom of the parameter space under the thick lines is allowed by constraints: the current matter density (the solid line) and
the upper limit ($\alpha \sim 0.07$) of the isocurvature power spectrum (the dashed line). The background color shows the power spectrum amplitude
and the labels on the thin lines represents the spectral index of the isocurvature $n_X$. The right y-axis shows the spectral index of the curvature
perturbation $n_s$.

which corresponds to $m_X \lesssim 4H_e$. We see once again that large observable non-Gaussianities can come from dark matter particles
that only compose a small fraction of the total dark matter. In Fig. 4, this region corresponds to the region along the dashed
curve that is far below the solid curve. This means that the CMB non-Gaussianities are sensitive to bosonic stable particles that
are negligible as far as their contribution to the gravitational energy budget today. As discussed in Sec. IV, the dashed curve in
Fig. 4 does not continue indefinitely as $m_X/m_\phi \to 0$ since perturbation theory for the primordial spectrum breaks down when
$\omega_X$ is less than around $10^{-5}$. Parametrically, this breakdown point occurs for $m_X/m_\phi \approx 1.4$, which corresponds to $m_X \approx 3H_e$.

From Fig. 4, we can also see that $\gamma = n_X - 1 \approx 0.1$ from the isocurvature spectral index $n_X$. This yields the advertised result
that the isocurvature non-Gaussianities can reach $f_{NL} \approx 30$.

In this chaotic inflationary model that realizes large non-Gaussianities, the low reheating temperature $T_{RH} \ll H_e$ implies a
long period of matter domination that may lead to nontrivial dark matter and inflaton condensate clustering phenomenology (see
e.g. [85–87]). As demonstrated in Sec. IV, a lower inflationary scale model that still realizes $\alpha \sim O(1)$ and $\gamma \sim O(0.1)$ would
allow for an evasion of any phenomenologically undesirable features that may arise from the long duration of clustering.
VI. CONCLUSIONS

In this paper, we have analyzed the effect of nonthermal dark matter consisting of weakly interacting gravitationally produced X bosons on the two and three point functions of the primordial CMB spectrum. We have demonstrated that large local non-Gaussians characterized by $f_{\text{NL}} \approx 30$ can result for a particular set of masses and reheating temperatures without violating isocurvature bounds. The conditions that result in large $f_{\text{NL}}$ yield a bound on the reheating temperature of around $10^6$ GeV, as well as a bound on the dark matter mass of around $3H_0 \lesssim m_X \lesssim 4H_0$. For lower allowed values of $m_X$ masses, the X bosons can generate a large $f_{\text{NL}}$ value despite the fact that they are an essentially negligible fraction of the dark matter in the universe.

Although explicit numerical computations were carried out to find the phenomenologically viable region only in the chaotic $m_X^2 \phi^2/2$ inflationary model, we have presented analytic arguments to demonstrate that a similar semi-quantitative behavior is expected for most slow-roll inflationary models, including models with lower inflationary scales. The agreement between the analytic considerations and the numerical results represents a nontrivial self-consistency check.

The mechanism presented in this paper connects nonthermal dark matter phenomenology to inflationary phenomenology. The possibility that a future discovery of large local non-Gaussianities may provide support for the existence of a nonthermal dark matter component is indeed intriguing. This may even be considered one of the several generic string phenomenological signatures associated with nonthermal dark matter components such as those considered in [88, 89], although a careful generalization of our work is required to apply the results in that context. We look forward to studying further hints nature may be willing to reveal in this direction through observations at the on-going Planck mission and other experiments that probe large scale structure.

Appendix A: Mode functions beyond the dS approximation

To calculate the correlation function or the energy density, it is necessary to solve the following mode equation (57). In dS space, where $H \equiv \dot{a}/a$ is constant, this has a well known solution with the Bunch-Davies boundary condition

$$X^\text{dS}_k(t) = \frac{\sqrt{\pi}}{2a^{3/2}\sqrt{H}} e^{\frac{\nu}{2}(\nu+1/2)} H^{(1)}_{\nu}(\frac{k}{aH}), \quad (A1)$$

in which $\nu = \sqrt{9/4 - m_X^2/H^2}$ and $H^{(1)}_{\nu}$ is a Hankel function. In the long-wavelength limit, $k/a \ll H$, this solution behaves as

$$X^\text{dS}_k(t) \approx -\frac{1}{\sqrt{\pi}} (-1)^{3/4} 2^{-1+\nu} e^{\nu \pi/2} \left(\frac{k}{aH}\right)^{-\nu} \frac{\Gamma(\nu)}{\sqrt{H a^3}}, \quad (A2)$$

in which the decaying solution has been dropped. For $m_X/H > 3/2$, $|X^\text{dS}_k|^2 \sim a^{-3}$ dilutes like pressureless dust.

However, in a quasi-dS spacetime in which $H$ varies slowly, this solution fails to be a good approximation after many efolds past the horizon crossing: $\Delta \equiv t - t_\star \lesssim 1/\epsilon H$, where $\epsilon = -H/H^2$. Conversely, Eq. (A1) is a good approximation up to the time $t_\star$ just after the horizon crossing. (57) can be approximated for $t > t_\star$ as

$$X_k(t) + 3H(t)X_k(t) + m_X^2X_k(t) \approx 0, \quad (A3)$$

and hence we can use the following ansatz for the non-decaying mode for small $\epsilon \ll 1$.

$$X_k(t) = X^\text{dS}_k(t_\star) \left(\frac{a(t_\star)}{a(t)}\right)^{3/2} \exp \left[\int^{t_\star}_{t_\star} dt_1 H(t_1) \nu(t_1)\right]. \quad (A4)$$

Now consider a quasi-dS spacetime with the constant slow-roll parameter $\epsilon, \dot{\epsilon} = 0$, for which

$$\frac{1}{H(t)} - \frac{1}{H(t_\star)} = \epsilon (t-t_\star), \quad (A5)$$

$$a(t) = a(t_\star)(1 + \epsilon H, t)^{1/\epsilon}. \quad (A6)$$

Hence, we find

$$\int^{t_\star}_{t_\star} dt' H(t') \nu(t') = \frac{1}{\epsilon} \left[\nu(t) - \frac{3}{2} \tanh^{-1} \left(\frac{3}{2\nu(t)}\right) - \nu_\star + \frac{3}{2} \tanh^{-1} \left(\frac{3}{2\nu_\star}\right)\right]. \quad (A7)$$
We note that the mode function $\mathcal{X}_k$ is oscillatory for imaginary $\nu$, but that is not reflected in Eq. (A7) because we have kept only the non-decaying mode in Eq. (A2). To see the oscillatory behavior of the mode, the decaying mode should be taken into account when the real to imaginary transition of $\nu$ occurs. We thus arrive at

$$
|\mathcal{X}_k(t)|^2 \approx |\mathcal{X}_k^{\text{dis}}(t_s)|^2 \left( \frac{a_s}{a(t)} \right)^3 \exp \left[ \frac{2\pi}{\epsilon} \left\{ \nu(t) - \frac{3}{2} \tanh^{-1} \left( \frac{3}{2\nu(t)} \right) - \nu_s + \frac{3}{2} \tanh^{-1} \left( \frac{3}{2\nu_s} \right) \right\} \right]
$$

$$
= \frac{2^{-2+2\nu_s}}{\pi} \left( \frac{k}{a_s H_s} \right)^{-2\nu_s} |\Gamma(\nu_s)|^2 \exp \left[ -\frac{2\pi}{\epsilon} \nu(t) - 2\nu_s + 3\ln \left( \frac{(\nu_s + \frac{3}{2}) H_s}{(\nu(t) + \frac{3}{2}) H(t)} \right) \right],
$$

(A8)
in which the subscript * indicates the variable is evaluated at $t_s$ and $\nu_s$ is a positive real number.

**Appendix B: Justification of the background**

In situations in which quantum operators $\hat{O}^{(q)}_k$ do not have classical expansion field values $\hat{O}^{(0)}_k$, it is necessary to justify how the $\hat{O}^{(q)}_k$ are to be identified. In such cases, the typical procedure is to solve for $\hat{O}^{(q)}_k$ directly in the presence of linear metric fluctuations because quantum operators are generically not spatially homogeneous even in Minkowski space while their matrix elements can be. Hence, it is convenient to construct $\hat{O}^{(0)}_k$ from matrix elements of $\hat{O}^{(q)}_k$. Here we outline how to extract $\hat{O}^{(0)}_k$ through a spatial average in both the classical and the quantum case. In the quantum case, the procedure is to construct a semiclassical quantity that corresponds to the most probable semi-classical configuration computed from a quantum expectation value. As explained below, this semiclassical construction is meaningful when the quantity has a classical interpretation.

Let us define the perturbation $\delta \hat{O}^{(q)}_k$ through $\hat{O}^{(q)}_k = \hat{O}^{(0)}_k + \delta \hat{O}^{(q)}_k$. If $\hat{O}^{(0)}_k$ is independent of spatial coordinates for some given fixed coordinate choice and $\delta \hat{O}^{(q)}_k$ satisfies the condition

$$
\lim_{L \to \infty} \frac{1}{V_L} \int_{V_L} d^3 x \delta \hat{O}^{(q)}_k = 0,
$$

(B1)

then we have

$$
\hat{O}^{(0)}_k \approx \lim_{L \to \infty} \frac{1}{V_L} \int_{V_L} d^3 x \hat{O}^{(q)}_k.
$$

(B2)

Here we do not use a 3-diffeomorphism invariant measure such that metric perturbations need not be included in the spatial average. Since unperturbed quantum operators are intrinsically inhomogeneous, spatial averaging cannot be used to extract $\hat{O}^{(0)}_k$.

On the other hand, for the quantum case, we are only interested in using it to obtain stochastic boundary conditions. For a large class of operators, for any fixed time slice $\Sigma$, there exists a probability functional $p_{\Sigma}(\{ \hat{O}^{(q)}_k \})$ such that

$$
\langle \hat{O}^{(q_1)}_k \ldots \hat{O}^{(q_n)}_k \rangle = \int \prod_q D\hat{O}^{(q)}_k p_{\Sigma}(\{ \hat{O}^{(q)}_k \}) \hat{O}^{(q_1)}_k \ldots \hat{O}^{(q_n)}_k,
$$

(B3)

where the left hand side is computed in a fixed vacuum. Since we are computing this at a fixed time, the right hand side functional measure $D\hat{O}^{(q)}_k$ is only discretized over the 3-space of $\Sigma$. It is important to think of an $\{ \hat{O}^{(q)}_k \}$ element in the ensemble (governed by $p_{\Sigma}$) as a classical configuration only for a fixed time slice. The time evolution of $\{ \hat{O}^{(q)}_k \}$ may not be governed by classical equations starting from that initial condition. Furthermore, if $p_{\Sigma}$ is sharply peaked, then the system is essentially in a single configuration (i.e., most elements in the ensemble are similar). This is one way of reaching classicality. Suppose there exists a such very probable field configuration on a fixed time slice. It is convenient to express this probable configuration (to use for $\hat{O}^{(q)}_k$ in Eq. (B2)) using matrix elements.

To construct this probable configuration, we start with the following question: if there are $N$ samples drawn from a quantum governed ensemble, how many would have exactly homogeneous configurations compared to those with inhomogeneous configurations. The set of exactly homogeneous configurations form a set of measure zero even though it can be the peak of the $p_{\Sigma}$.

\[5\] This argument requires generalization when the operators involve time derivatives. This argument will apply for the case of our interest for correlators of $\hat{X}^2$. 

functional (unless this functional diverges at that field configuration point). Most of the $N$ samples would be inhomogeneous. Having oriented ourselves, the next question that is relevant for us is what is the most probable value of Eq. (B2) given $N$ samples in the ensemble governed by $p_{\Sigma}$. The number of configurations with a given characteristic $\Gamma$ is given by

$$
N_{\Gamma} = N \int \prod_q D\sigma^{(q)}_\gamma p_{\Sigma}(\{\sigma^{(q)}_\gamma\}) \delta(\Gamma) \det \frac{\delta \Gamma}{\delta \sigma^{(q)}_\gamma},
$$

(B4)

where $\delta(\Gamma)$ represents an appropriately generalized delta-function in the functional space and the determinant is there for the appropriate normalization. For a fixed spatial average we have the regularized constraint

$$
\Gamma_{\sigma^{(q)}_0, V_L} = \frac{1}{V_L(\Sigma)} \int_{V_L(\Sigma)} d^3x \sigma^{(q)}_\gamma - \sigma^{(q)}_0,
$$

(B5)

in which case $N_{\Gamma_{\sigma^{(q)}_0, V_L}}$ gives the number of elements in the $N$ sample that realize the homogeneous value of $\sigma^{(q)}_0$ within the volume $V_L$. We will in the end take $L \to \infty$ if the infrared regulator removal is meaningful. Hence, we find

$$
\frac{\delta \Gamma}{\delta \sigma^{(q)}_\gamma} = \frac{1}{V_L(\Sigma)},
$$

(B6)

which is independent of $\sigma^{(q)}_\gamma$. The functional derivatives are taken only with respect to functions of 3-spatial variable. Incorporating this into Eq. (B4), we find

$$
N_{\Gamma_{\sigma^{(q)}_0, V_L}} = N \int \prod_q D\sigma^{(q)}_\gamma p_{\Sigma}(\{\sigma^{(q)}_\gamma\}) \delta \left( \frac{1}{V_L(\Sigma)} \int_{V_L(\Sigma)} d^3x \sigma^{(q)}_\gamma - \sigma^{(q)}_0 \right) \frac{1}{V_L(\Sigma)}. 
$$

(B7)

Now, the maximum of $N_{\Gamma_{\sigma^{(q)}_0, V_L}}$ is obtained by taking a derivative with respect to $\sigma^{(q)}_0(t)$ (since we are working on a single time slice, this derivative need not be functional):

$$
0 = \frac{\partial}{\partial \sigma^{(q)}_0(t)} N_{\Gamma_{\sigma^{(q)}_0, V_L}} = -N \int \prod_q D\sigma^{(q)}_\gamma p_{\Sigma}(\{\sigma^{(q)}_\gamma\}) \delta(\sigma^{(q)}_\gamma - \sigma^{(q)}_0) \left( \frac{1}{V_L(\Sigma)} \int_{V_L(\Sigma)} d^3x \sigma^{(q)}_\gamma - \sigma^{(q)}_0(t) \right) \frac{1}{V_L(\Sigma)},
$$

(B8)

where the prime on the delta-function corresponds to the derivative with respect to the functional argument of the delta function. The functional argument of the delta function can be considered to be a function of the variation

$$
\delta \left[ \frac{1}{V_L(\Sigma)} \int_{V_L(\Sigma)} d^3x \sigma^{(q)}_\gamma \right] = \frac{\delta \sigma^{(q)}_\gamma}{V_L(\Sigma)}.
$$

(B10)

Hence, an integration by parts will yield a solvable equation to the problem of maximizing $N_{\Gamma_{\sigma^{(q)}_0, V_L}}$ as

$$
\int \prod_q D\sigma^{(q)}_\gamma \frac{\delta}{\delta \sigma^{(q)}_\gamma} p_{\Sigma}(\{\sigma^{(q)}_\gamma\}) \delta \left( \frac{1}{V_L(\Sigma)} \int_{V_L(\Sigma)} d^3x \sigma^{(q)}_\gamma - \sigma^{(q)}_0 \right) = 0.
$$

(B11)

If $\sigma^{(q)}_0$ is chosen to be $\sigma^{(q)}_\gamma$ such that the $\sigma^{(q)}_\gamma$ configurations that satisfy

$$
\sigma^{(q)}_\gamma = \frac{1}{V_L(\Sigma)} \int_{V_L(\Sigma)} d^3x \sigma^{(q)}_\gamma
$$

(B12)

also satisfy

$$
\frac{\delta}{\delta \sigma^{(q)}_\gamma} p_{\Sigma}(\{\sigma^{(q)}_\gamma\}) = 0,
$$

(B13)

we have a solution $\sigma^{(q)}_\gamma$. Hence, we must look for the peak of $p_{\Sigma}(\{\sigma^{(q)}_\gamma\})$. 

Consider
\[
\langle \hat{\mathcal{O}}^{(s)}_\varepsilon \rangle = \int \prod_q D\hat{\mathcal{O}}^{(q)}_\varepsilon \ p_\Sigma(\{\hat{\mathcal{O}}^{(q)}_\varepsilon \}) \hat{\mathcal{O}}^{(s)}_\varepsilon .
\] (B14)

Assuming \( p_\Sigma(\{\hat{\mathcal{O}}^{(q)}_\varepsilon \}) \) is sharply peaked, consider
\[ f \equiv \ln p_\Sigma(\{\hat{\mathcal{O}}^{(q)}_\varepsilon \}). \] (B15)

The peak is located at \( \frac{\delta f}{\delta \hat{\mathcal{O}}^{(q)}_\varepsilon} = 0 \) corresponding to the field configuration that satisfies
\[
\frac{\delta p_\Sigma(\{\hat{\mathcal{O}}^{(q)}_\varepsilon \})}{\delta \hat{\mathcal{O}}^{(q)}_\varepsilon}\bigg|_{\hat{\mathcal{O}}^{(q)}_\varepsilon = 0} = 0. \] (B16)

Hence, the quadratic expansion of \( f \) about the peak configuration is
\[
f = f(\hat{\mathcal{O}}^{(q)}_P) + \frac{1}{2} \int dx_1 dx_2 \frac{\delta^2 f}{\delta \hat{\mathcal{O}}^{(q_1)}_\varepsilon (x_1) \delta \hat{\mathcal{O}}^{(q_2)}_\varepsilon (x_2)} \hat{\mathcal{O}}^{(q_1)}_\varepsilon (x_1) - \hat{\mathcal{O}}^{(q_1)}_P (x_1) \hat{\mathcal{O}}^{(q_2)}_\varepsilon (x_2) - \hat{\mathcal{O}}^{(q_2)}_P (x_2), \] (B17)

where the repeated \( q_i \) indices are summed. \( \hat{\mathcal{O}}^{(s)}_\varepsilon \) can be raised in Eq. (B14) using the usual trick of introducing a source
\[
\hat{\mathcal{O}}^{(s)}_\varepsilon \rightarrow \frac{\delta}{\delta J} \exp \left[ \int d^3 x J \hat{\mathcal{O}}^{(s)}_\varepsilon \right] \bigg|_{J = 0} \] (B18)
and carrying out the leading saddle-point approximation integral to obtain
\[
\langle \hat{\mathcal{O}}^{(s)}_\varepsilon (x) \rangle = \hat{\mathcal{O}}^{(s)}_P (x), \] (B19)

where we used
\[
1 = \int \prod_q D\hat{\mathcal{O}}^{(q)}_\varepsilon \ p_\Sigma(\{\hat{\mathcal{O}}^{(q)}_\varepsilon \}) \] (B20)
\[
\approx \exp \left[ \frac{1}{2} \int dx_1 dx_2 \frac{\delta^2 f}{\delta \hat{\mathcal{O}}^{(q_1)}_\varepsilon (x_1) \delta \hat{\mathcal{O}}^{(q_2)}_\varepsilon (x_2)} \hat{\mathcal{O}}^{(q_1)}_\varepsilon (x_1) - \hat{\mathcal{O}}^{(q_1)}_P (x_1) \hat{\mathcal{O}}^{(q_2)}_\varepsilon (x_2) - \hat{\mathcal{O}}^{(q_2)}_P (x_2) \right] . \] (B21)

We note that if we use a spatially translation invariant state to take the expectation value, \( \hat{\mathcal{O}}^{(q)}_P (x) \) is automatically spatially translation invariant and thus
\[
\hat{\mathcal{O}}^{(q)}_s = \hat{\mathcal{O}}^{(q)}_P \] (B22)
in Eq. (B12). Hence, Eqs. (B19) and (B22) combine to give the most probable spatially averaged configuration to be
\[
\hat{\mathcal{O}}^{(s)}_s = \langle \hat{\mathcal{O}}^{(s)}_\varepsilon \rangle \] (B23)
to leading order in saddle-point approximation (an expansion in the peakedness of the distribution function \( p_\Sigma \)).

Hence, when matching to the classical fluid, it is appropriate to consider the homogeneous background quantity associated with the quantum operator to be \( \langle \hat{\mathcal{O}}^{(s)}_\varepsilon \rangle \). It is important to remember that since \( \hat{\mathcal{O}}^{(s)}_s \) (in Eq. (B12)) is being identified with \( \hat{\mathcal{O}}^{(q)}_0 \) (in Eq. (B2)), we are now assuming that the \( \varepsilon = 0 \) solution of perturbation theory is governed by equations that depend only on time for the coordinate system that we have chosen.
Acknowledgments

We thank Peng Zhou and Toni Riotto for useful conversations. DJHC also thanks Xingang Chen and Eichiro Komatsu for discussions following a preliminary presentation of this work at the “Cosmological Non-Gaussianity Workshop” which was held on May 13-15, 2011 at the University of Michigan. This work was supported in part by the DOE grant DE-FG02-95ER40896.

[1] D. H. Lyth and A. Riotto, Phys.Rept. 314, 1 (1999), hep-ph/9807278.
[2] J. M. Maldacena, JHEP 0305, 013 (2003), astro-ph/0210603.
[3] D. Seery and J. E. Lidsey, JCAP 0506, 003 (2005), astro-ph/0503692.
[4] V. Acquaviva, N. Bartolo, S. Matarrese, and A. Riotto, Nucl.Phys. B667, 119 (2003), astro-ph/0209156.
[5] P. Creminelli, JCAP 0310, 003 (2003), astro-ph/0306122.
[6] A. P. Yadav and B. D. Wandelt, Phys.Rev.Lett. 100, 181301 (2008), 0712.1148.
[7] E. Komatsu et al. (WMAP Collaboration), Astrophys.J.Suppl. 180, 330 (2009), 0803.0547.
[8] J. Fergusson and E. Shellard, Phys.Rev. D80, 043510 (2009), 0812.3413.
[9] K. M. Smith, L. Senatore, and M. Zaldarriaga, JCAP 0909, 006 (2009), * Brief entry *, 0901.2572.
[10] A. P. Yadav and B. D. Wandelt, Adv.Astron. 2010, 565248 (2010), * Temporary entry *, 1006.0275.
[11] E. Komatsu, Class.Quant.Grav. 27, 124010 (2010), 1003.6097.
[12] A. Curto, E. Martinez-Gonzalez, R. Barreiro, and M. Hobson (2011), 1105.6106.
[13] M. LoVerde, A. Miller, S. Shandera, and L. Verde, JCAP 0804, 014 (2008), * Brief entry *, 0711.4126.
[14] A. Slosar, C. Hirata, U. Seljak, S. Ho, and N. Padmanabhan, JCAP 0808, 031 (2008), * Brief entry *, 0805.3580.
[15] A. Afshordi and A. J. Tolley, Phys.Rev. D78, 123507 (2008), 0806.1046.
[16] S. Matarrese and L. Verde, Astrophys.J. 677, L77 (2008), * Brief entry *, 0801.4826.
[17] E. Sefusatti and E. Komatsu, Phys.Rev. D76, 083004 (2007), 0705.0343.
[18] N. Dalal, O. Dore, D. Huterer, and A. Shirokov, Phys.Rev. D77, 123514 (2008), 0710.4560.
[19] S. Shandera, N. Dalal, and D. Huterer, JCAP 1103, 017 (2011), 1010.3722.
[20] K. M. Smith and M. LoVerde (2010), * Temporary entry *, 1010.0055.
[21] G. D’Amico, M. Musso, J. Norena, and A. Paranjape, JCAP 1102, 001 (2011), * Temporary entry *, 1005.1203.
[22] D. Tseliakhovich, C. Hirata, and A. Slosar, Phys.Rev. D82, 043531 (2010), 1004.3302.
[23] A. D. Linde and V. F. Mukhanov, Phys. Rev. D56, 535 (1997), astro-ph/9610219.
[24] D. H. Lyth, C. Ungarelli, and D. Wands, Phys. Rev. D67, 023503 (2003), astro-ph/0208055.
[25] M. Zaldarriaga, Phys. Rev. D69, 043508 (2004), astro-ph/0306006.
[26] O.-G. Huang, JCAP 0811, 005 (2008), 0808.1793.
[27] J. Kumar, L. Leblond, and A. Rajaraman, JCAP 1004, 024 (2010), 0909.2040.
[28] C. T. Byrnes and K.-Y. Choi, Adv.Astron. 2010, 724525 (2010), * Temporary entry *, 1002.3110.
[29] N. Barnaby, Adv.Astron. 2010, 156180 (2010), * Temporary entry *, 1010.5507.
[30] J.-O. Gong and H. M. Lee (2011), 1105.0073.
[31] V. Demozzzi, A. Linde, and V. Mukhanov, JCAP 1104, 013 (2011), 1012.0549.
[32] C. M. Peterson and M. Tegmark, Phys.Rev. D84, 023520 (2011), 1011.6675.
[33] T. Suyama, T. Takahashi, M. Yamaguchi, and S. Yokoyama, JCAP 1012, 030 (2010), 1009.1979.
[34] D. Langlois and A. Lepidi, JCAP 1101, 008 (2011), 1007.5498.
[35] C. T. Byrnes, K. Enqvist, and T. Takahashi, JCAP 1009, 026 (2010), 1007.5148.
[36] C. Burgess, M. Cicoli, M. Gomez-Reino, F. Quevedo, G. Tasinato, et al., JHEP 1008, 045 (2010), 1005.4840.
[37] A. Mazumdar and J. Rocher, Phys.Rept. 497, 85 (2011), 1001.0993.
[38] X. Chen and Y. Wang, JCAP 1004, 027 (2010), 0911.3380.
[39] L. Alabidi, K. Malik, C. T. Byrnes, and K.-Y. Choi, JCAP 1011, 037 (2010), 1002.1700.
[40] T. Takahashi, M. Yamaguchi, J. Yokoyama, and S. Yokoyama, Phys.Lett. B678, 15 (2009), * Brief entry *, 0905.0240.
[41] Q.-G. Huang and Y. Wang, JCAP 0809, 025 (2008), 0808.1168.
[42] M. Kawasaki, K. Nakayama, T. Sekiguchi, T. Suyama, and F. Takahashi, JCAP 0811, 019 (2008), * Brief entry *, 0808.0009.
[43] A. Chambers and A. Rajantie, JCAP 0808, 002 (2008), * Brief entry *, 0805.4795.
[44] S. Yokoyama, T. Suyama, and T. Tanaka, JCAP 0707, 013 (2007), * Brief entry *, 0705.3178.
[45] M. Sasaki, J. Valiviita, and D. Wands, Phys.Rev. D74, 103003 (2006), astro-ph/0607627.
[46] N. Barnaby and J. M. Cline, Phys.Rev. D73, 106012 (2006), astro-ph/0601481.
[47] H. R. Cogollo, Y. Rodriguez, and C. A. Valenzuela-Toledo, JCAP 0808, 029 (2008), 0806.1546.
[48] D. J. Chung, E. W. Kolb, and A. Riotto, Phys.Rev. D59, 023501 (1999), in *Venice 1999, Neutrino telescopes, vol. 2* 217-237, hep-ph/9802238.
[49] V. Kuzmin and I. Tkachev, Phys. Rev. D59, 123006 (1999), hep-ph/9809547.
[50] D. J. Chung, E. W. Kolb, A. Riotto, and L. Senatore, Phys.Rev. D72, 023511 (2005), astro-ph/0411468.
[51] D. J. Chung, E. W. Kolb, and A. Riotto, Phys.Rev.Lett. 81, 4048 (1998), hep-ph/9805473.
[52] J. R. Ellis, J. L. Lopez, and D. V. Nanopoulos, Phys.Lett. B247, 257 (1990).
[53] K. Benakli, J. R. Ellis, and D. V. Nanopoulos, Phys. Rev. D59, 047301 (1999), hep-ph/9803333.
