# Particle Production under External Fields and Its Applications

*by*

Hojin Yoo

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Physics)

*at the*

University of Wisconsin – Madison

2014

Defended on August 8th, 2014

Dissertation approved by the following members of the Final Oral Committee:

Daniel Chung · Associate Professor of Physics

Albert Stebbins · Scientist of Fermilab

Lisa Everett · Associate Professor of Physics

Yang Bai · Assistant Professor of Physics

Peter Timbie · Professor of Physics

# Abstract

The thesis presents studies of vacuum pair productions and its applications in early universe cosmology and high energy astrophysics. Vacuum often becomes unstable and spontaneously decays into pairs of particles in rapidly expanding universes or under strong external electromagnetic fields. Theoretically, spontaneous pair productions due to such non-trivial backgrounds of spacetimes or electromagnetic fields are well-understood. However, the effect of particle productions has not been observed so far because of experiemtal difficulties in obtaining large curvatures of space-times or strong electric fields. Although it may be impossible to observe the pair productions directly via laboratory experiments, there are still powerful sources of space-time curvatures or electric fields in cosmology and astrophysics, which result in observations. In Part I, we explore the inflationary models in early universe utilizing pair productions through gravity. We study observable signatures on the cosmic microwave background, such as isocurvature perturbations and non-Gaussianities, generated from the particle production of WIMPzillas and axions during or after inflation. In Part II, we investigate the electron-positron pair production in the magnetosphere of pulsars whose electromagnetic fields are expected to close to or even greater than the pair production threshold. In particular, we demonstrate that the pair production may be responsible for giant pulses from the Crab pulsar.

# Acknowledgements

My PhD study as well as my life in the US could not be more fruitful and pleasant without support from all the people around me. No matter how I wrote here, it may be not enough to describe my warm feelings and gratitude for these people.

Firstly, I would like to thank my advisor, Daniel Chung, for his expert guidance and tremendous support. He has been willing to share his knowledge and own experience and generous with his time. His answers to my questions are from his deep thought and far more detail than I expected, his tackling always clarifies my understanding, and our discussion never stops until we find an complete answer no matter how long it takes. Furthermore, he always encouraged me to attend conferences, workshops and summer schools.

I am also very grateful to Dr. Albert Stebbins at Fermilab. He generously accepted me and gave me an oppurtunities to work as a graduate fellow at Fermilab. His insight, intuition and interesting idea have inspired and guided me to broader areas of physics. Personally, because of his and his wife Sonya's help, Hyejung and I could safely settle in a new city and prepare a home for our baby.

I thank Professors Peter Timbie and Lisa Everett as well as Prof. Chung and Dr. Stebbins. They were willing to spend their precious time to help me find a job last winter. I thank Prof. Timbie again for sharing his teaching experience and resources while I was a teaching assistant in his course.

I also thank the high energy group members, and my friends Andrew Long and Peng Zhou. I have learnt many things from them about physics and life of graduate students. Especially, Peng Zhou was a great research collabotator.

Finally, I would like to thank my parents for their encouragement, and I am most thankful for my wife, Hyejung, for her constant support and patience, and also for

my son, Sihyung, for his smiles.

# Contents

# Part I

# Particle Production in the Early Universe

# Chapter 1

# Introduction

In this part of the thesis, we present a study of gravitational particle production and the isocurvature models of inflation. This introductory chapter provides a short review for the inflation (See also Refs. [1–4]) and the gravitational particle production (See Refs. [5–7]), which are the main themes of the study. In Chapter 2, we discuss isocurvature perturbations and non-Gaussianties originated from the particle production of a massive scalar field, which is stable and very weakly interacting. In Chapter 3, we investigate a model that the particle production of a spin-1/2 field generates isocurvature perturbations. In Chapter 4, we estimate the cross-correlation between curvature perturbations and isocurvature perturbations from QCD axions and WIMPzillas using a Ward identity of diffeomorphism. In this part, we will use the metric signature $(-, +, +, +)$ and the natural units $c = \hbar = 1$ if it is not specified.

## 1.1 Standard Big-Bang Model

One of the basic ideas of modern cosmology is that the universe is homogeneous and isotropic on large scales. The general line element of the homogeneous and isotropic

universe, called the Friedmann-Robertson- Walker (FRW) universe, takes the form

$$ds^2 = -dt^2 + a(t)^2 \left[ \frac{dr^2}{1 - kr^2} + r^2 \left( d\theta^2 + \sin^2\theta d\phi^2 \right) \right], \tag{1.1}$$

where $a(t)$ is the scale factor, and $k = -1, 0, -1$ denotes the global curvature of the spatial hyper-surfaces. We refer to $r$, $\theta$, and $\phi$ as the co-moving coordinates. An observer at rest remains at fixed co-moving coordinates $r$, $\theta$, and $\phi$, while proper distances between two co-moving coordinates scale as $a(t)$.

The Einstein equation describes how the space-time evolves by matter and energy in the universe

$$R_{\mu\nu} - \frac{1}{2} R\, g_{\mu\nu} = 8\pi G\, T_{\mu\nu}, \tag{1.2}$$

where $G$ is the Newton's constant, $R_{\mu\nu}$ is the Riemann tensor, $R$ is the Ricci scalar, and $T_{\mu\nu}$ is the energy-momentum tensor. For a general perfect fluid, the energy-momentum tensor can be written as

$$T_{\mu\nu} = (\rho + p)\, u_\mu u_\nu + p g_{\mu\nu}, \tag{1.3}$$

where $\rho$ and $p$ are the energy density and the pressure, and $u$ is the four velocity of the fluid. $u^\mu$ satisfies $u^\mu u_\mu = -1$. For the FRW space-time, the Einstein equation yields the Friedmann equation

$$H^2 \equiv \left( \frac{\dot{a}}{a} \right)^2 = \frac{8\pi G}{3} \rho - \frac{k}{a^2}, \tag{1.4}$$

and the conservation equation

$$\dot{\rho} + 3H(\rho + p) = 0. \tag{1.5}$$

If each era of the Big-Bang model is parametrized by the equation of state

$$p = w\rho, \tag{1.6}$$

we find

$$\rho \propto a^{-3(1+w)} = \begin{cases} a^{-3} & \text{for matter} \quad (w = 0), \\ a^{-4} & \text{for radiation} \quad (w = 1/3), \\ \text{constant} & \text{for cosmological constant} \quad (w = -1) \end{cases}, \qquad (1.7)$$

As it can be seen from the above equation, one finds that the universe was more energetic in the past. Extrapolating further back in time, the universe eventually reaches a "Big Bang" singularity, where all of space collapses to a single point at a finite time in the past. Although this singularity arises from a naive extrapolation to the limit that the classical equations of motion break down, the singularity provides a formal marker in the spacetime evolution. Starting from the discovery of the expansion of the universe in accordance with the Hubble's law in 1920s, numerous measurements including the cosmic microwave background (CMB) radiation, the relative abundances of light elements, and more recently the observation of large scale structures provide strong observational evidence of the Big Bang model.

## 1.2 Inflation

Although the Big-Bang model explains many aspects of the observation, rather special initial conditions seem to be required for a universe similar to ours at present. Some of the most important issues in the initial conditions are the followings:

### The Flatness Problem

Introducing the relative energy density

$$\Omega_{tot} \equiv \frac{\sum_i \rho_i}{\rho_c}, \qquad (1.8)$$

where $\rho_c$ is the critical energy density defined by $\rho_c \equiv 3H^2/8\pi G$, we can rewrite the Friedmann equation (1.4) as

$$\frac{k}{a^2} = H^2 \left[\Omega_{tot} - 1\right].$$

(1.9)

Taking the time-derivative, we find that

$$\frac{d}{dt}\left|\Omega_K\right| \equiv \frac{d}{dt}\left|\Omega_{tot} - 1\right| = -2\left|k\right|\frac{\ddot{a}}{\dot{a}^3}.$$

(1.10)

This implies that unless $k = 0$ the curvature tends to grow in time in a decelerating universe, such as matter or radiation dominated. The current observation indicates $\Omega_K = -0.042^{+4.3}_{-4.8}$, and the curvature should be even smaller at the early epoch, which requires an enormous fine-tuning of the initial flatness condition of the universe.

## The Horizon Problem

In the FRW universe, the particle horizon is defined by the maximum distance from which light could have reached an observer by a specific time:

$$R_H(t_0) = \int_0^{t_0} \frac{a(t_0)}{a(t)} dt.$$

(1.11)

For a matter or radiation dominated universe, the particle horizon is given by

$$R_H(t_0) = \begin{cases} 2t_0 & \text{for radiation} \\ 3t_0 & \text{for matter} \end{cases},$$

(1.12)

where the scale factor behaves as

$$a(t) \propto \begin{cases} t^{1/2} & \text{for radiation} \\ t^{2/3} & \text{for matter} \end{cases}.$$

This implies that any length scale inside the horizon today was outside the horizon in past. In other words, in a matter or radiation dominated universe the different regions

of the universe today were not in causal contact in the early universe. Nevertheless, the observed universe today looks highly isotropic. For example, the CMB experiments shows that causally disconnected regions on the last scattering surface have temperature anisotropies of the same order of magnitude $\delta T/T \sim 10^{-5}$. The horizon problem is about how causally disconnected regions of the early universe could have evolved to the nearly homogeneous and isotropic universe today.

## The topological defect problems

Suppose the standard model arises due to spontaneous symmetry breaking from a larger gauge theory. Phase transitions due to symmetry breaking may produce topological defects, such as monopoles, domain-walls, and cosmic strings, in the early universe. For example, nearly every sensible grand unified theory predicts the existence of such defects. Typically, monopoles are stable and very heavy $m \sim M_{GUT} \sim 10^{16} GeV$, and thus their energy density decays as $a^{-3}$. If they are produced at high temperature $T \sim M_{GUT}$, they should have persisted until now and their energy density should overwhelm that of ordinary particles. However, their existence is inconsistent with the evolution of the universe that we observe today.

A natural solution to these problems is provided by the inflationary paradigm [8–11]. It postulates a period in the early universe called *inflation* during which the universe undergoes an accelerated expansion $\ddot{a} > 0$. From (1.10), the curvature tends to decrease to zero. Furthermore, because the expansion of the universe during inflation grows faster than the horizon, causally connected regions would be stretched to super-horizon scales. Also, the universe cools down and its monopoles, if it has them, dilute away during inflation. However, whether inflation completely solves the problems, in particular, the topological defect problems, is still questionable [12,13], and so are the naturalness and the predictability of inflation (For example, see [14–16]

for recent criticisms).

Such an expansion is often achieved by a scalar field $\phi$ called the *inflaton.* As long as $\phi$ dominates the total energy density and it slowly rolls down a potential hill, the universe will expand almost like the de-Sitter universe. This accelerated expansion lasts long enough to resolve the initial condition problems until the inflaton $\phi$ reaches the minimum of the potential or close to it. Once inflation is over, the inflaton energy is eventually converted to ordinary standard model particles, which reheats the universe. After the *reheating,* the universe follows the standard Big Bang cosmology.

The simplest single field inflationary model has a Lagrangian

$$\mathcal{L} = -\frac{1}{2}\left(\partial\phi\right)^2 - V(\phi), \tag{1.13}$$

and the Friedmann equations (1.4) and the conservation equation (1.5) are written as

$$H^2 = \frac{8\pi G}{3}\left[\frac{1}{2}\left(\partial\phi\right)^2 + V(\phi)\right], \tag{1.14}$$

$$\ddot{\phi} + 3H\dot{\phi} + \frac{\partial}{\partial\phi}V(\phi) = 0. \tag{1.15}$$

As long as the slow-roll conditions

$$\varepsilon \equiv 4\pi G\left(\frac{\partial V/\partial\phi}{V}\right)^2 \ll 1, \tag{1.16}$$

$$\eta \equiv 8\pi G\left(\frac{\partial^2 V/\partial\phi^2}{V}\right) \ll 1, \tag{1.17}$$

are satisfied, the equations can be approximated as

$$H^2 \simeq \frac{8\pi G}{3}V(\phi), \tag{1.18}$$

$$3H\dot{\phi} \simeq -\frac{\partial}{\partial\phi}V(\phi). \tag{1.19}$$

Note that $\varepsilon$ and $\eta$ are called the slow-roll parameters, and inflation ends when $\varepsilon \simeq 1$.

One of the remarkable results of inflation models is that it provides the initial fluctuations responsible for the inhomogeneities of the universe, such as the CMB anisotropies. During inflation, quantum fluctuations of the inflaton are amplified and become classical once their wavelengths are stretched beyond the horizon. These fluctuations give rise to the primordial density perturbations, which eventually gives predictions for the CMB anisotropies. Single field inflationary models lead to almost Gaussian, adiabatic, and scale-invariant perturbations.

## 1.3   Primordial Density Perturbations

The inhomogeneities were generated from quantum fluctuations and stretched to the super-horizon scales due to the extremely rapid expansion of the universe during inflation. As the inflaton decays during the radiation era, these inhomogeneities provide the initial perturbations at this epoch, called primordial perturbations. The subsequent evolution of the perturbations is described by a set of differential equations from the Einstein equation combined with the Boltzmann equation. Integrating the equations one can relate the inflaton perturbations to the CMB anisotropies or the inhomogeneities of large scale structures.

### Adiabatic and Isocurvature Perturbations

One of interesting features of a single scalar field inflation is that it yields only adiabatic primordial perturbations. Once the inflaton decays to various components of the universe during the radiation era, the initial density fluctuations of these components of the inflaton decay naturally inherit that of the inflaton, and they follow the relation

$$\frac{\delta\rho_\phi}{\rho_\phi + p_\phi} = \frac{\delta\rho_\gamma}{\rho_\gamma + p_\gamma} = \frac{\delta\rho_\nu}{\rho_\nu + p_\nu} = \frac{\delta\rho_b}{\rho_b + p_b} = \frac{\delta\rho_c}{\rho_c + p_c}, \tag{1.20}$$

where the subscript $\phi, \gamma, \nu, b$ and $c$ denote the inflaton, photon, neutrino, baryon, and cold dark matter (CDM), respectively. This is called *adiabatic* initial condition. Although this is a quite restrictive condition, it is in good agreement with the observations.

However, the primordial perturbations do not necessarily satisfy the adiabatic condition a priori. For instance, if there are additional degrees of freedom, such as auxiliary scalar fields, during inflation, the perturbations in each field do not have to satisfy the adiabatic condition. Therefore the perturbations from the fields can lead to *isocurvature* perturbations. Intuitively, isocurvature perturbations correspond to setting nonzero the difference of the number over-densities of at least one pair of fluid element species while setting to zero the total energy density inhomogeneity on long wavelength scales, and they are defined by

$$S_{ij} = \frac{\delta\rho_i}{\rho_{i+}p_i} - \frac{\delta\rho_j}{\rho_{j+}p_j}, \tag{1.21}$$

where $i, j = \gamma, \nu, b$ and $c$. We often consider the CDM isocurvature, which corresponds to $S_{c\gamma}$.

From the CMB experiments, pure isocurvature is strongly ruled out. However, sub-dominant isocurvature perturbations mixed with adiabatic perturbations are still allowed, and the parameter bounds of isocurvature perturbations are sensitive to correlation with adiabatic perturbations. The detail discussion is given in 4.2.

## 1.4    Particle Production in an Expanding Universe

In this section, we develop the necessary tools for studying gravitational particle production in an expanding universe. To be concrete, we quantize matter fields, whereas we treat the space-time background purely classically. For demonstration, we use a scalar field. Most of discussion in this section follows Ref. [6].

## Quantization

We begin with the action

$$S = \int d^n x \, \frac{1}{2} \sqrt{-g} \left[ -(\partial\phi)^2 - \left( m^2 + \xi R \right) \phi^2 \right], \tag{1.22}$$

where $R$ is the Ricci scalar. Varying the action with respect to $\phi$ yields the field equation for $\phi$

$$\left[ -\Box_x + m^2 + \xi R \right] \phi(x) = 0, \tag{1.23}$$

where

$$\Box_x \phi(x) \equiv \frac{1}{\sqrt{-g}} \partial_\mu \left[ \sqrt{-g} g^{\mu\nu} \partial_\nu \right] \phi(x). \tag{1.24}$$

To decompose the field by modes, we define the scalar product of $\phi$

$$(\phi_1, \phi_2) \equiv -i \int_\Sigma \phi_1(x) \overleftrightarrow{\partial}_\mu \phi_2^*(x) \sqrt{-g} d\Sigma^\mu, \tag{1.25}$$

where $\Sigma$ denotes a space-like hyper-surface, and $\phi \overleftrightarrow{\partial}_\mu \psi = \phi \partial_\mu \psi - (\partial_\mu \phi) \psi$. Then using a complete set of mode solutions $\{u_i\}$ of (1.23) satisfying the orthonormal conditions

$$(u_i, u_j) = \left( u_i^*, u_j^* \right) = \delta_{ij}, \quad \left( u_i, u_j^* \right) = 0, \tag{1.26}$$

$\phi$ can be expanded as

$$\phi(x) = \sum_i \left[ a_i u_i(x) + a_i^\dagger u_i^*(x) \right], \tag{1.27}$$

and $\phi$ is quantized by promoting $a_i$ and $a_i^\dagger$ to the creation and the annihilation operators for a vacuum state $|0\rangle$ satisfying the commutation relation

$$\left[ a_i, a_j^\dagger \right] = \delta_{ij}, \quad [a_i, a_j] = \left[ a_i^\dagger, a_j^\dagger \right] = 0. \tag{1.28}$$

## Ambiguity of a Vacuum

For Minkowski space, $\partial_t$ is a time-like Killing vector, orthogonal to the space-like hyper-surfaces $t = constant$, and the modes are eigenfunctions of this Killing vector

with constant eigenvalues that allow to define positive and negative modes associated with $a$ and $a^\dagger$, respectively. However, unlike the Minkowski space-time, there is no natural positive and negative modes in a curved space-time because of absence of such a Killing vector. Therefore any set of modes satisfying the above basic requirements may be used in setting up the Fock space of the quantum field theory, and the notion of particles depends on the choice of a set of modes.

For example, suppose a second complete set of modes $\{\tilde{u}_i\}$ defining new creation and annihilation operators $\tilde{a}_i$, $\tilde{a}_i^\dagger$ and a vacuum state $|\tilde{0}\rangle$. As both sets are complete, they are related by

$$\tilde{u}_j = \sum_i \left( \alpha_{ji} u_i + \beta_{ji} u_i^* \right), \qquad (1.29)$$

$$a_i = \sum_i \left( \alpha_{ji} \tilde{a}_j + \beta_{ji}^* \tilde{a}_j^\dagger \right), \qquad (1.30)$$

where $\alpha_{ij}$ and $\beta_{ij}$ are called Bogoliubov coefficients. Then the expectation value of the operator $N_i \equiv a_i^\dagger a_i$ for the number of $i$ mode particles in the state $|\tilde{0}\rangle$ is

$$\langle \tilde{0} | N_i | \tilde{0} \rangle = \sum_j |\beta_{ji}|^2, \qquad (1.31)$$

which means the vacuum $|\tilde{0}\rangle$ contains $\sum_j |\beta_{ji}|^2$ particles defined on the vacuum $|\tilde{0}\rangle$.

Then a natural question to ask is which set of modes furnishes the best description of the physical vacuum. To answer the question, it is necessary to specify the quantum measurement process. Especially, the state of motion of the particle detecting device can affect the presence of quanta. For example, an inertial and an accelerated observer may not generally agree upon the presence of particles, even in Minkowski space. Although all inertial observers measure the same vacuum in Minkowski space, in general this will not hold in a curved space-time as there is no preferred coordinate system. Here we present ways to give a notion of particles in the FRW space-time.

## Hamiltonian Diagonalization

For convenience, we introduce the conformal time

$$\eta = \int \frac{1}{a} dt, \tag{1.32}$$

and the rescaled field

$$\phi = a\varphi, \tag{1.33}$$

and set $\xi = 0$. Substituting these into Eq. (1.23), we obtain

$$\varphi'' - \nabla^2\varphi + \left(a^2 m^2 - \frac{a''}{a}\right)\varphi = 0, \tag{1.34}$$

which has a similar form with the Klein-Gordon equation in Minkowski space with the time-dependent effective mass

$$m_{eff}(\eta) = a^2 m^2 - \frac{a''}{a}. \tag{1.35}$$

Following the procedure described above, we expand $\varphi$ using the Fourier modes

$$\varphi(x) = \int \frac{d^3 k}{(2\pi)^3} \left[a_{\vec{k}} u_k(t) e^{i\vec{k}\cdot\vec{x}} + a_{\vec{k}}^\dagger u_k^*(t) e^{-i\vec{k}\cdot\vec{x}}\right], \tag{1.36}$$

where

$$\left[a_{\vec{k}}, a_{\vec{k'}}^\dagger\right] = \delta^3(\vec{k} - \vec{k'}), \tag{1.37}$$

$$u_k' u_k^* - u_k u_k'^* = -i, \tag{1.38}$$

and $u_{\vec{k}}$ satisfies

$$u_k'' + \left(k^2 + a^2 m^2 - \frac{a''}{a}\right) u_k = 0. \tag{1.39}$$

Because the Hamiltonian for $\varphi$ is given by

$$H(\eta) = \frac{1}{2} \int d^3 x \left[\pi^2 + (\nabla\varphi)^2 + m_{eff}^2(\eta)\varphi^2\right], \tag{1.40}$$

where $\pi$ is the conjugate momentum of $\varphi$, $\pi \equiv \varphi'$, using Eqs. (1.36), (1.37), and (1.39), it is rewritten as

$$H(\eta) = \frac{1}{2} \int d^3k \left[ \epsilon_k \left( a_{\vec{k}} a_{\vec{k}}^\dagger + a_{\vec{k}}^\dagger a_{\vec{k}} \right) + f_k a_{\vec{k}}^\dagger a_{-\vec{k}}^\dagger + f_k^* a_{\vec{k}} a_{-\vec{k}} \right], \tag{1.41}$$

where

$$\epsilon_k(\eta) = |u_k'|^2 + \omega_k(\eta)^2 |u_k|^2, \tag{1.42}$$

$$f_k(\eta) = u_k'^2 + \omega_k(\eta)^2 u_k^2,$$

$$\omega_k(\eta)^2 \equiv k^2 + m_{eff}^2(t). \tag{1.43}$$

One can find that the non-diagonal terms in the Hamiltonian at $\eta_0$ vanish if

$$u_k'(\eta_0) = -i\omega_k u_k(\eta_0), \tag{1.44}$$

and the Wronskian condition (1.38) gives

$$|u_k(\eta_0)|^2 = \frac{1}{2\omega_k(\eta_0)}. \tag{1.45}$$

This allows to uniquely define an instantaneous vacuum at time $\eta = \eta_0$.

**Adiabatic Expansion**

Although the Hamiltonian diagonalization provides a way to define a unique vacuum, it has been strongly criticised by [17] because the procedure over-estimates produced particle numbers and it often does not lead to the cut-off in particle spetra at high energies. On the other hand, one wants to capture the notion of vacuum being a state devoid of particles, which results in a particle definition for which there is minimal particle production by the changing geometry. We expect that the geometry of space-time has little effect on the notion of particles and the particle number should remain nearly constant if the Compton wavelength of a particle is much smaller than the

scale of the variation of the geometry, i.e., if $k^2 + a^2 m^2$ is large with respect to the quantity associated with the space-time curvature, which corresponds to $(a''/a)$ or $(a'/a)^2$ in the FRW space-time . Thus the definition of particles can be consistently expanded in term of the Compton wavelength and the space-time curvature.

We can construct this idea based on the WKB-type solutions

$$u_k(\eta) = \frac{1}{\sqrt{2W_k(\eta)}} \exp\left(-i \int^{\eta} W_k(\eta')d\eta'\right), \qquad (1.46)$$

where $W_k$ satisfies

$$W_k^2(\eta) = \omega_k^2(\eta) - \frac{1}{2}\left(\frac{W_k''}{W_k} - \frac{3}{2}\frac{W_k'^2}{W_k^2}\right). \qquad (1.47)$$

Then we can iteratively approximate the solution of Eq. (1.47) using the slowness of the $W_k$ changes. For example, as a zeroth order approximation, we have

$$W_k^{(0)}(\eta) = \omega_k(\eta), \qquad (1.48)$$

and to the next order,

$$W_k^{(2)2}(\eta) = \omega_k^2(\eta) - \frac{1}{2}\left(\frac{\omega_k''}{\omega_k} - \frac{3}{2}\frac{\omega_k'^2}{\omega_k^2}\right). \qquad (1.49)$$

## Example

To complete the demonstration, we provide a toy model calculation for particle production. We consider the time dependent effective mass

$$m_{eff}(\eta)^2 = m^2\left(c^2 + d^2\eta^2\right), \qquad (1.50)$$

and calculate the Bogoliubov coefficient between two sets of modes defined by the adiabatic expansion in the asymptotic past and future, $\{u_k^{in}\}$ and $\{u_k^{out}\}$.

Although the space-time is not static in the asymptotic past and future, it is varying slowly

$$\frac{d^l}{d\eta^l}\frac{\omega_k'}{\omega_k} \to 0 \quad \text{as } \eta \to \pm\infty \quad \text{for } l \geq 0. \qquad (1.51)$$

Using the adiabatic expansion at the zero order, we have

$$u_k(\eta) \propto \exp\left[\mp i\left(\frac{1}{2}dm\eta^2\right) + \frac{i}{2dm}\left(k^2 + c^2m^2\right)\ln\eta\right] \quad \text{as } \eta \to \pm\infty. \tag{1.52}$$

Then we may construct the exact solution of (1.39) using this in the asymptotic limits:

$$u_k^{in}(\eta) = (2dm)^{-1/4}e^{-\pi\lambda/8}D_{-(1-i\lambda)/2}\left[(i-1)(dm)^{1/2}\eta\right], \tag{1.53}$$

$$u_k^{out}(\eta) = u_k^{in}(-\eta)^*, \tag{1.54}$$

where $\lambda = k^2/dm + c^2m/d$, and $D_\nu(z)$ is the parabolic cylinder function $D$.

From the result, we find

$$u_k^{in} = \frac{i(2\pi)^{1/2}e^{-\pi\lambda/4}}{\Gamma(\frac{1}{2}\text{-}\frac{i}{2}\lambda)}u_k^{out} - ie^{\pi\lambda/2}u_k^{out*}. \tag{1.55}$$

This indicates that if the state of the system is chosen to be $|0\rangle$, an observer at the asymptotic future would detect a particle number spectrum

$$|\beta_k|^2 = \exp\left\{-\pi\left[\left(k^2/dm\right) + mc^2/d\right]\right\}, \tag{1.56}$$

which falls exponentially fast at high energies. Note that this spectrum is the same as that for non-relativistic thermal particles.

# Chapter 2

# Isocurvature Perturbation from Boson Production

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.

This work was performed in collaboration with Daniel Chung. It was published in the journal Physical Review D in January of 2013 [18].

## 2.1 Introduction

Standard slow-roll inflationary models with a single dynamical field degree of freedom (e.g. see the review article [2]) cannot generate large local non-Gaussianities (NGs) [19–22], which have been widely discussed and speculated upon in the context of the cosmic microwave background (CMB) data (e.g. [23–29]) and large scale structure data (e.g. [30–39]). Many multifield mechanisms have been proposed to generate observably large local NGs (e.g. [40–64]). Most of these models utilize coherent condensate field degrees of freedom instead of incoherent many-particle states.

In this work, 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 [65,66] generate observably large NGs.[1] 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 $T_{\text{RH}}$. 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

---

[1]These particles are sometimes called gravitationally produced WIMPZILLAs.

of inflation. We also find that $f_{NL}$ will be suppressed if $T_{\mathrm{RH}} \gtrsim 10^6$ GeV if the dark matter is absolutely stable.[2] 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 [67]. We note that this was also briefly considered in [40], which arrived at a pessimistic conclusion. However, that paper did not consider the model as carefully as [67], which reached a more realistic conclusion regarding the viability of such scenarios. The purpose of this work 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. 2.3, 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. 2.4. In Sec. 2.5, we check our general analytic arguments by computing in detail numerically the observables in the context of an $m^2\phi^2/2$ inflationary model. We then close the main body of the work with a summary and conclusions in Sec. 2.6. In Appendix 2.7, 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 2.8, we justify how the effective classical background variable about which the classical perturbations are defined is given by the expectation value of the quantum operator. Throughout this work, our metric convention is $(+, -, -, -)$.

---

[2]This mass scale has part of its origins from the maximum dark matter abundance today.

## 2.2   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 work, 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:

$$
\begin{aligned}
S &= \int d^4x \sqrt{g} \frac{1}{2} \left[ (\partial\phi)^2 - 2U(\phi) + (\partial X)^2 - m_X^2 X^2 \right] \\
&= +S_{\mathrm{SM}}[g_{\mu\nu}, \{\psi_{\mathrm{SM}}\}] + S_{\mathrm{RH}}[g_{\mu\nu}, \phi, \{\psi_{\mathrm{SM}}\}],
\end{aligned}
\tag{2.1}
$$

where $S_{\mathrm{SM}}$ is the SM sector, $S_{\mathrm{RH}}$ 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 $T_{\mathrm{RH}} \gtrsim H_e/(2\pi)$, 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 work if

$$\left(\frac{200\mathrm{TeV}}{m_X}\right)^2 \left(\frac{T_{\mathrm{max}}}{m_X}\right) \lesssim 1, \tag{2.2}$$

where $T_{\mathrm{max}} \gtrsim T_{\mathrm{RH}}$ is the maximum effective temperature reached during reheating [68]. 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 [69–81].

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 [82] satisfies

$$\lim_{k\to\infty} \hat{\alpha}_k |BD, 0\rangle = 0, \tag{2.3}$$

where $\hat{\alpha}_k$ 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 [65,83–85]:

$$\hat{a}_k|BD,0\rangle = 0, \tag{2.4}$$

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

$$\langle BD,0|\hat{T}_{\mu\nu}^{(X,\text{ren})}|BD,0\rangle = 0, \tag{2.5}$$

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 [65,86]. 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 $\rho_x(t_e) \approx 10^{-2}H_e^4$, which leads to the relic abundance of $X$ particles today to be

$$\Omega_X h^2 \approx 10^{-1} \left(\frac{H_e}{10^{12} \text{ GeV}}\right)^2 \left(\frac{T_{\text{RH}}}{10^6 \text{ GeV}}\right). \tag{2.6}$$

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. (2.6) 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

$$\Omega_{CDM} = \Omega_{therm} + \Omega_X, \tag{2.7}$$

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

$$\omega_X \equiv \frac{\Omega_X}{\Omega_{CDM}}, \tag{2.8}$$

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 $H_*$ (the Hubble expansion rate when the modes of interest leave the horizon), $H_e$, and $T_{\mathrm{RH}}$ (the reheating temperature). As we will see, the primary role of $H_*$ is to determine the spectral index of the isocurvature spectrum, $H_e < H_*$ controls the particle production, and $T_{\mathrm{RH}}$ 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 [40] 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 with $\langle \hat{X} \rangle \neq 0$ over the patch of the size of that wavelength, 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 (2.29) is related to $3 - 3\sqrt{1 - 4m_X^4/9H^2}$ and our scenario has $m_X^2 > 0$, the correlator has a blue spectrum (as we will see explicitly in Sec. 2.3), and thus the loop integrals (see for example Eqs. (2.28) and (2.33)) 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 [87]. The UV cutoff of our theory is set to be the horizon scale at the end of inflation: $k_{\mathrm{UV}} = H_e a_e$.

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

$$ds^2 = (1 + E)dt^2 - 2a\partial_i F dt dx^i - a^2 \left[\delta_{ij} + A\delta_{ij} + \partial_i \partial_j B\right] dx^i dx^j. \qquad (2.9)$$

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

$$\zeta \equiv \frac{A}{2} - H\frac{\delta\rho_\phi}{\dot{\rho}_\phi}. \qquad (2.10)$$

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

$$\Delta_\zeta^2(k) \equiv \frac{k^3}{2\pi^2}P_\zeta(k), \qquad (2.11)$$

where to leading order in the slow-roll parameter

$$\epsilon(\phi_k) = \frac{M_p^2}{2}\left(\frac{U'(\phi_k)}{U(\phi_k)}\right)^2, \qquad (2.12)$$

we have

$$P_\zeta(k) = \frac{1}{12k^3\epsilon(\phi_k)}\frac{U(\phi_k)}{M_p^4}. \tag{2.13}$$

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 work 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.

## 2.3   Two-Point Function

Although isocurvature perturbations have been previously computed for this class of models [67], 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 $X$:

$$T_{\mu\nu}^{(X)} = \partial_\mu X \partial_\nu X - g_{\mu\nu}\left[\frac{1}{2}\partial_\alpha X \partial^\alpha X - V(X)\right] \tag{2.14}$$

where $V(X) = m_X^2 X^2/2$. Comparing this to

$$T_{\mu\nu}^{(perfect\ fluid)} \equiv (\rho_X^{(p)} + P_X^{(p)})u_\mu u_\nu - g_{\mu\nu}P_X^{(p)}, \tag{2.15}$$

we see that if we define [88]

$$u^\mu \equiv \frac{\partial^\mu X}{\sqrt{g^{\alpha\beta}\partial_\alpha X \partial_\beta X}}, \tag{2.16}$$

we can satisfy the equality

$$T_{\mu\nu}^{(X)} = T_{\mu\nu}^{(perfect\ fluid)}, \tag{2.17}$$

if

$$\rho_X^{(p)} \equiv u^\mu u^\nu T_{\mu\nu}^{(perfect\ fluid)} = \frac{1}{2}\partial^\alpha X \partial_\alpha X + V(X) \tag{2.18}$$

and

$$P_X^{(p)} \equiv \frac{1}{2}\partial^\alpha X \partial_\alpha X - V(X). \tag{2.19}$$

We note that to identify Eq. (2.16) 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 $T_{00}^{(X)}$, $\rho_X^{(p)}$ is a diffeomorphism scalar. We also note that even though Eq. (2.18) looks like it has the wrong sign between the $(\partial_0 X)^2$ and $|\vec{\nabla}X|^2$, the sign is correct and $\rho_X^{(p)}$ 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 $\hat{X}$.explained in Appendix 2.8, this allows us to identify

$$\rho_0^{(p)} = \langle : \hat{\rho}_X^{(p)} : \rangle, \tag{2.20}$$

where the normal ordering is with respect to the operators defining the $\hat{X}$ vacuum state during the quasi-dS era. After Bogoliubov transforming $: \hat{\rho}_X^{(p)} :$ to operators of 1-particle states relevant for non-dS spacetime, $\langle : \hat{\rho}_X^{(p)} : \rangle$ will develop a nonzero value at that later time. We note that $\langle : \hat{\rho}_X^{(p)} : \rangle$ 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, $\langle : \hat{\rho}_X^{(p)} : \rangle$ will be spatially translation invariant as well.) Next, we consider the semi-classical variable

$$\delta\rho_X^{(p)} \equiv \rho_X^{(p)} - \rho_0^{(p)}(t). \tag{2.21}$$

We can then define the usual fluid variable associated with $\rho_X$

$$\zeta_X \equiv \frac{A}{2} - H \frac{\delta\rho_X^{(p)}}{\frac{d}{dt}\rho_0^{(p)}(t)}, \tag{2.22}$$

where we parameterize the spatial scalar metric perturbation as be $h_{ij}^S = -a^2(t)(A\delta_{ij} + \partial_i\partial_j B)$ with $\bar{g}_{\mu\nu} = \text{diag}\{1, -a^2, -a^2, -a^2\}$. Under the diffeomorphism $t \to t - \xi^0$, we have

$$A \rightarrow A + 2H\xi^0 \tag{2.23}$$

$$\delta\rho_X^{(p)} \rightarrow \delta\rho_X^{(p)} + \xi^0 \frac{d}{dt}\rho_0^{(p)}, \tag{2.24}$$

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 [89]

$$\delta_{S_X} \equiv 3\left(\zeta_X - \zeta_R\right). \tag{2.25}$$

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,

$$\delta_{S_X} = \delta_X^{(c)} = \frac{\delta\rho_X^{(p)(c)}}{\rho_0^{(p)}(t)}, \tag{2.26}$$

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
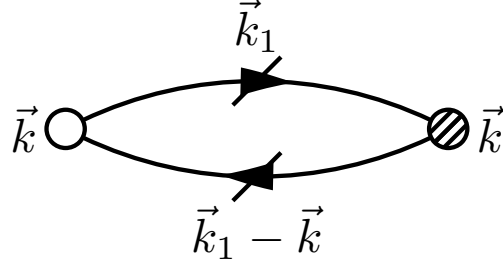
Figure 2.1: A diagrammatic representation of the 2-point function of a composite operator$\delta_X$. The open circle corresponds to external momentum flowing out, and the shaded circle to external momentum flowing in. The slashes on the legs indicate that the$X$ propagators are on-shell.

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. (2.89)) as follows:

$$\delta\hat{\rho}_X^{(p)(c)} =: \hat{\rho}_X^{(p)(c)} : -\hat{1}\rho_0^{(p)}(t). \tag{2.27}$$

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 $\langle \hat{\delta}_X^{(c)}(t,\vec{r})\hat{\delta}_X^{(c)}(t,0)\rangle$.

Diagramatically, the 2-point function is as shown in Fig. 2.1. Hence, the 2-point power spectrum is

$$\begin{aligned}\Delta_{\delta_{S_X}}^2(k) &= \frac{k^3}{2\pi^2}\int d^3r e^{-i\vec{k}\cdot\vec{r}}\langle \hat{\delta}_X^{(c)}(t,\vec{r})\hat{\delta}_X^{(c)}(t,0)\rangle \\ &\approx \frac{k^3}{2\pi^2}2\int_{\Lambda_{IR}}^{a_e H_e}\frac{d^3k_1}{(2\pi)^3}P_X(k_1)P_X(|\vec{k}-\vec{k}_1|),\end{aligned} \tag{2.28}$$

in which

$$P_X(k) \equiv \frac{m_X^2}{2\rho_0^{(p)}}|X_k|^2, \tag{2.29}$$

$\hat{\delta}_X^{(c)}$ is approximated as $m_X^2 \hat{X}^2/2\rho_0^{(p)}$, and $X_k$ is the solution to the mode equation in Appendix 2.7.$m_X/H < 3/2$ and $k/(aH) \ll 1$, we can express $P_X(k)$ approximately

as

$$P_X(k) = A_X \left(\frac{k}{k_0}\right)^{\gamma_X(H_*)} k^{-3}, \tag{2.30}$$

where

$$\gamma_X(H) = 3 - 3\sqrt{1 - \frac{4m_X^2}{9H^2}} > 0. \tag{2.31}$$

$H_*$ is allowed a wave vector dependence; however, this effect is a subdominant correction to the already small $(m_X/H_*)^2$. 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

$$
\begin{aligned}
A_X \quad &\sim \quad \frac{10^2 |\Gamma(\frac{3}{2} - \frac{\gamma_X(H_*)}{2})|^2}{2^{\gamma_X(H_*)}\pi} \frac{m_X H_*^2}{H_e^3} \\
&\times \exp\left[\frac{1}{\epsilon}\Re\left\{\gamma_X(H_*) - \gamma_X(H_e) + 3\ln\left(\frac{1 - \frac{1}{3}\gamma_X(H_*) + \frac{H_*}{m_X}}{1 - \frac{1}{3}\gamma_X(H_e) + \frac{H_e}{m_X}}\right)\right\}\right] \tag{2.32}
\end{aligned}
$$

where $\epsilon = \frac{M_p^2}{2}\left(\frac{U'(\phi)}{U(\phi)}\right)^2$ 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 $k/a < \epsilon/H_*$. 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 $\epsilon^{-1} \gg 1$. For the dark matter abundance to be compatible with cosmological observations, it is important that $H_* > H_e$ while $m_X/H_e \sim O(1)$. Using these approximations and Eq. (2.30), we find that Eq. (2.28) yields

$$
\begin{aligned}
\Delta_{\delta_{S_X}}^2(k) \quad &\approx \quad \frac{k^3}{2\pi^2}A_X^2\left(\frac{k}{k_0}\right)^{2\gamma_X}\frac{1}{k^3} \times 2\int_{\Lambda_{IR}/k}^{a_e H_e/k} \frac{d^3u}{(2\pi)^3}u^{\gamma_X-3}|1-\vec{u}|^{\gamma_X-3} \\
&\approx \quad \frac{4}{\gamma_X}\left[1 - \left(\frac{\Lambda_{IR}}{k}\right)^{\gamma_X}\right]\left[\frac{A_X}{2\pi^2}\left(\frac{k}{k_0}\right)^{\gamma_X}\right]^2 \\
&\approx \quad \frac{4}{\gamma_X}\left(\frac{k^3}{2\pi^2}P_X(k)\right)^2, \tag{2.33}
\end{aligned}
$$

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. $\gamma_X(H_*) > 0$), and $\Delta_{\delta_{S_X}}^2$
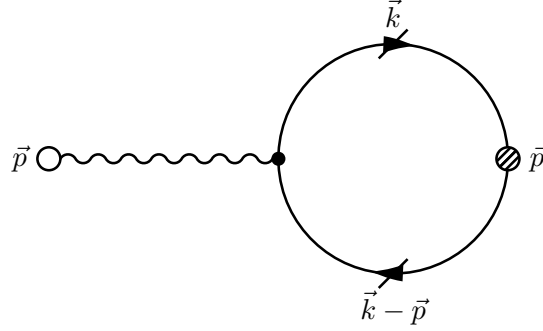
Figure 2.2: A diagrammatic representation of the cross-correlator $\langle \zeta \delta_S \rangle$, where the leading interaction vertex is proportional to $m_X^2$. The wavy propagator corresponds to the on-shell $\langle \zeta \zeta \rangle$ correlator. The open circle corresponds to external momentum flowing out, and the shaded circle to external momentum flowing in. The slashes on the legs indicate that the $X$ propagators are on-shell.

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 $\left(k^3 P_X(k)\right)^2$ 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. (2.7), we can rescale Eq. (2.33) to obtain the total CDM isocurvature perturbations as

$$\Delta_{\delta_S}^2(k) = \omega_X^2 \Delta_{\delta_{S_X}}^2(k), \tag{2.34}$$

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 $\Delta_{\delta_{S_X}}^2(k)$ is large. The observational bound on $\Delta_{\delta_S}^2$ is stringent unless the cross-correlation between curvature and isocurvature is negligible, i.e. [24, 90–92]

$$\Delta_{\zeta \delta_S}^2 \ll \Delta_\zeta \Delta_{\delta_S}, \tag{2.35}$$

where $\Delta_{\zeta \delta_S}^2$ is the power spectrum of the cross-correlation.left hand side $\langle \hat{\zeta} \hat{\delta}_S \rangle$ corresponding to the diagram shown in Fig. 2.2 can be computed using the in-in formal-

ism [19, 93] using the trilinear interaction Hamiltonian in the comoving gauge

$$H_I(t) \ni - \int d^3x \, a^3(t) \left[ \hat{T}_X^{ij}(t, \vec{x}) a^2(t) \delta_{ij} \hat{\zeta}(t, \vec{x}) \right], \tag{2.36}$$

where $T_X^{\mu\nu}$ is the stress energy tensor of $X$. Note that other interaction Hamiltonian contributions are derivatively suppressed. As will be shown elsewhere [94], the curvature-isocurvature cross correlation is

$$\beta \equiv \frac{\Delta_{\zeta\delta_S}^2}{\Delta_\zeta \Delta_{\delta_S}} \lesssim \frac{\Delta_\zeta}{2} \sim 2.5 \times 10^{-5}, \tag{2.37}$$

which shows that the cross-correlation is negligible. This fact is understood by the soft-$\zeta$ theorem [19, 95], which allows to factorize $\left\langle \hat{\zeta}\hat{\zeta} \right\rangle$ from $\left\langle \delta\hat{\rho}_X \hat{\zeta} \right\rangle$, i.e.

$$\left\langle \delta\hat{\rho}_X \hat{\zeta}_{\vec{p}} \right\rangle \sim \left\langle \hat{\zeta}_{-\vec{p}} \hat{\zeta}_{\vec{p}} \right\rangle \frac{\partial}{\partial \ln a} \left\langle \hat{\rho}_X \right\rangle \tag{2.38}$$

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 [24, 96] becomes

$$\alpha = \frac{\Delta_{\delta_S}(k_0)}{\Delta_\zeta(k_0) + \Delta_{\delta_S}(k_0)} < 0.064, \tag{2.39}$$

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-tiled 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 [91, 92] show
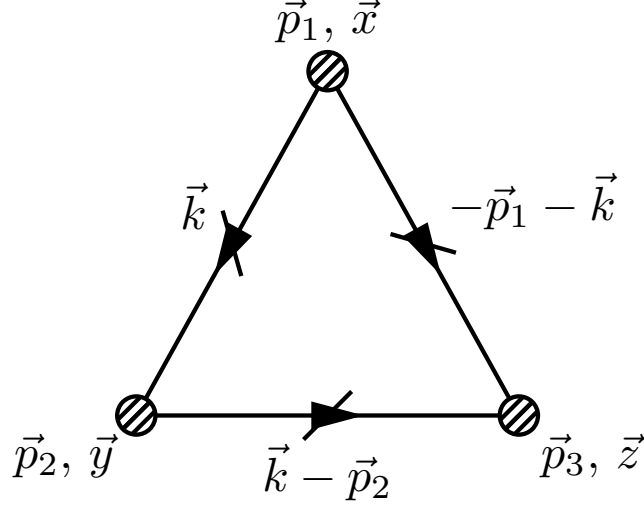
Figure 2.3: A diagrammatic representation of the 3-point function of a composite operator$\delta_X$. The slashes indicate that the $X$ propagators are on-shell. The spatial variables $\{\vec{x}, \vec{y}, \vec{z}\}$ indicate the insertion points of the external momenta$\vec{p}_i$.

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

## 2.4  Bispectrum

In this section, we now compute the bispectrum $B_{\delta_S}(\vec{p}_1, \vec{p}_2, \vec{p}_3)$ defined by

$$
\begin{aligned}
(2\pi)^3 \delta^{(3)}(\sum_i \vec{p}_i) B_{\delta_S}(\vec{p}_1, \vec{p}_2, \vec{p}_3) &= \omega_X^3 \int d^3 x_1 d^3 x_2 d^3 x_3 e^{-i\sum_n \vec{p}_n \cdot \vec{x}_n} \\
&\times \langle \hat{\delta}_X^{(c)}(t, \vec{x}_1) \hat{\delta}_X^{(c)}(t, \vec{x}_2) \hat{\delta}_X^{(c)}(t, \vec{x}_3) \rangle, \quad (2.40)
\end{aligned}
$$

where we recall from the previous section that $\delta_X^{(c)}$ in the comoving gauge can be identified with the gauge invariant quantity $\delta_{S_X}$. With the diagrammatic representation given by Fig. 2.3, we find

$$
B_{\delta_S}(\vec{p}_1, \vec{p}_2, -(\vec{p}_1 + \vec{p}_2)) \approx 8\omega_X^3 \int \frac{d^3 k}{(2\pi)^3} P_X(|\vec{k}|) P_X(|\vec{p}_1 + \vec{k}|) P_X(|\vec{p}_2 - \vec{k}|), \quad (2.41)
$$

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:

$$
\begin{aligned}
B_{\delta_S}(\vec{p}_1, \vec{p}_2, \vec{p}_3) &\approx \frac{8}{\gamma_X(H_*)} \frac{p_{\min}^3}{2\pi^2} P_X(p_{\min}) \omega_X^3 \\
&\times [P_X(p_1)P_X(p_2) + P_X(p_2)P_X(p_3) + P_X(p_3)P_X(p_1)] \quad (2.42)
\end{aligned}
$$

where $p_{\min} = \min\{|\vec{p}_i|\}$. 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 $\gamma_X(H_*)$ is large, and hence $m_X$ has to be smaller than $H_*$ for a non-negligible bispectrum. On the other hand, if $m_X$ is too small compared to $H_*$, 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 $f_{NL} = 1$

$$
f_{NL}^S \equiv \frac{B_S}{B_\zeta|_{f_{NL}=1}} = \frac{5}{6} \frac{B_{\delta_S}(\vec{p}_1, \vec{p}_2, \vec{p}_3)}{P_\zeta(p_1)P_\zeta(p_2) + P_\zeta(p_2)P_\zeta(p_3) + P_\zeta(p_3)P_\zeta(p_1)}, \quad (2.43)
$$

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.

$$\frac{\Delta T}{T} = -\frac{1}{5}\zeta - \frac{2}{5}\delta_S. \tag{2.44}$$

However, on scales smaller than $1/k_{eq}$, 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.adiabatic initial conditions, the gravitational potential on superhorizon scales are frozen.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,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 leads to an approximate relationship on large scales [59] which can be summarized as[3]

$$f_{NL} \approx 4f_{NL}^{S} \tag{2.45}$$

---

[3]As mentioned in Ref. [59], 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 $b_{NL,l}^{(iso)} \equiv \frac{2}{\pi} \int dk \, k^2 g_{Tl}^{(iso)}(k) j_l(kr)$, which is independent of the spectra index of isocurvature.

where $f_{NL}$ approximately coincides with the usual local $f_{NL}$ definition.[4] Thus, in the squeezed triangle limit, $f_{NL}$ is analytically estimated to be

$$f_{NL} = \frac{80}{3} \frac{\omega_X^3}{\gamma_X} \left( \frac{p_{\min}^3}{2\pi^2} P_X(p_{\min}) \right) \frac{P_X(p_1)P_X(p_2) + P_X(p_2)P_X(p_3) + P_X(p_3)P_X(p_1)}{P_\zeta(p_1)P_\zeta(p_2) + P_\zeta(p_2)P_\zeta(p_3) + P_\zeta(p_3)P_\zeta(p_1)}. \tag{2.46}$$

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 $\delta_S$ contributions to $\Delta T/T$, and the transfer function. As $p_{\min}^3 P_X(p_{\min})$ is suppressed partly from the smallness of the $P_X$ amplitude as well as the blue tilt (causing $p_{\min}^3$ 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 $P_X/P_\zeta$ will arise from the fact that $\delta_S$ 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. (2.33)), we have

$$P_X \approx \pi^2 k^{-3} \sqrt{\gamma_X \Delta_{\delta_S}^2} \frac{1}{\omega_X}. \tag{2.47}$$

If we define [97]

$$\frac{\alpha}{1-\alpha} \equiv \frac{\Delta_{\delta_S}^2}{\Delta_\zeta^2}, \tag{2.48}$$

we find for $\alpha \ll 1$ that

$$\frac{P_X}{P_\zeta} \sim \frac{1}{2} \sqrt{\frac{\alpha \gamma_X}{\Delta_\zeta^2}} \frac{1}{\omega_X} \sim 10^4 \sqrt{\alpha \gamma_X} \frac{1}{\omega_X}. \tag{2.49}$$

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 $\Delta_\zeta^2 \ll \sqrt{\Delta_\zeta^2}$ and $\Delta_{\delta_S}^2 \propto P_X^2$. Combining Eq. (2.46) and Eq. (2.49), we see

$$f_{NL} \sim 6 \times 10^3 \alpha^{3/2} \sqrt{\gamma_X}. \tag{2.50}$$

Hence, if we can achieve $\alpha \approx 0.07$ [24] and $\gamma_X \sim 0.1$ in our model, we can achieve

$$f_{NL} \approx O(35) \tag{2.51}$$

---

[4]This relationship is obtained by comparison between the reduced bispectrums $b_{l_1 l_2 l_3}^{(adi)}$ and $b_{l_1 l_2 l_3}^{(iso)}$ only at large angular scales $(l_1, l_2, l_3 \lesssim 10)$. Thus, $f_{NL}^S$ should not be interpreted as $f_{NL}^{local}$ in Refs. [24, 96], which is obtained by the full analysis involving large and small scales.

on large scales. Although this $f_{NL}$ cannot be directly interpreted as $f_{NL}^{local}$ of Ref. [24, 96] (as discussed in footnote4), a similar result which is obtained through a full numerical analysis in Ref. [90]

$$f_{NL}^{local} \approx 30 \left(\frac{\alpha}{0.07}\right)^{3/2} \tag{2.52}$$

gives a consistency check to our analytic argument leading to Eq. (2.51). This level of non-Gaussianity is clearly observable according to the forecasts of Planck and large scale structure experiments (see e.g. [98,99]). 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. (2.33) and (2.34) show perturbation theory would break down if$\alpha$ needs to be fixed as $\omega_X \to 0$ limit is literally taken. For example,have large local non-Gaussianities, the CDM isocurvature should be Hence in the case $\omega_X \lesssim 10^{-5}$, the perturbativity bound of $\delta_X < 1$ is violated.the next section, we will compute the relevant quantities more precisely in the context of a simple $U(\phi) = m_\phi^2 \phi^2 / 2$ inflationary model.

## 2.5   Numerical Results

As shown in Sec. 2.4, 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 2.7, 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 $U(\phi) = m_\phi^2 \phi^2/2$ model.

inflation, the energy density $\rho_X$ (as investigated analytically and numerically in [66, 100]) is estimated to be

$$\frac{\rho_X}{\rho_\phi} \approx 10^{-10} \frac{m_X}{m_\phi} \left(\frac{m_\phi}{10^{13} \text{ GeV}}\right)^2 \exp(-2\pi m_X/m_\phi). \qquad (2.53)$$

To obtain the $k$ that appears in the mode function $X_k$, we use the pivot scale $k_0 = 0.002$ Mpc$^{-1}$ and the standard reheating relationships [101]

$$k = \frac{a_k}{a_0} k_0 \qquad (2.54)$$

$$\frac{a_k}{a_0} = \frac{a_k}{a_e} \frac{a_e}{a_0} \qquad (2.55)$$

$$\approx 2 \times 10^{-31} \left(\frac{a_k}{a_e}\right) \left(\frac{H(t_e)}{10^{13} \text{GeV}}\right)^{-2/3} \left(\frac{T_{\text{RH}}}{10^9 \text{GeV}}\right)^{1/3}. \qquad (2.56)$$

The scale factor ratio $a_k/a_e$ is computed directly from the solution of $\phi$ with the potential $U(\phi)$. The mode function $X_k$ is then obtained by solving the equation of motion

$$\ddot{X}_k + 3H\dot{X}_k + \frac{k^2}{a^2} X_k + m_X^2 X_k = 0 \qquad (2.57)$$

with the Bunch-Davies initial condition.

In Fig. 2.5, 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
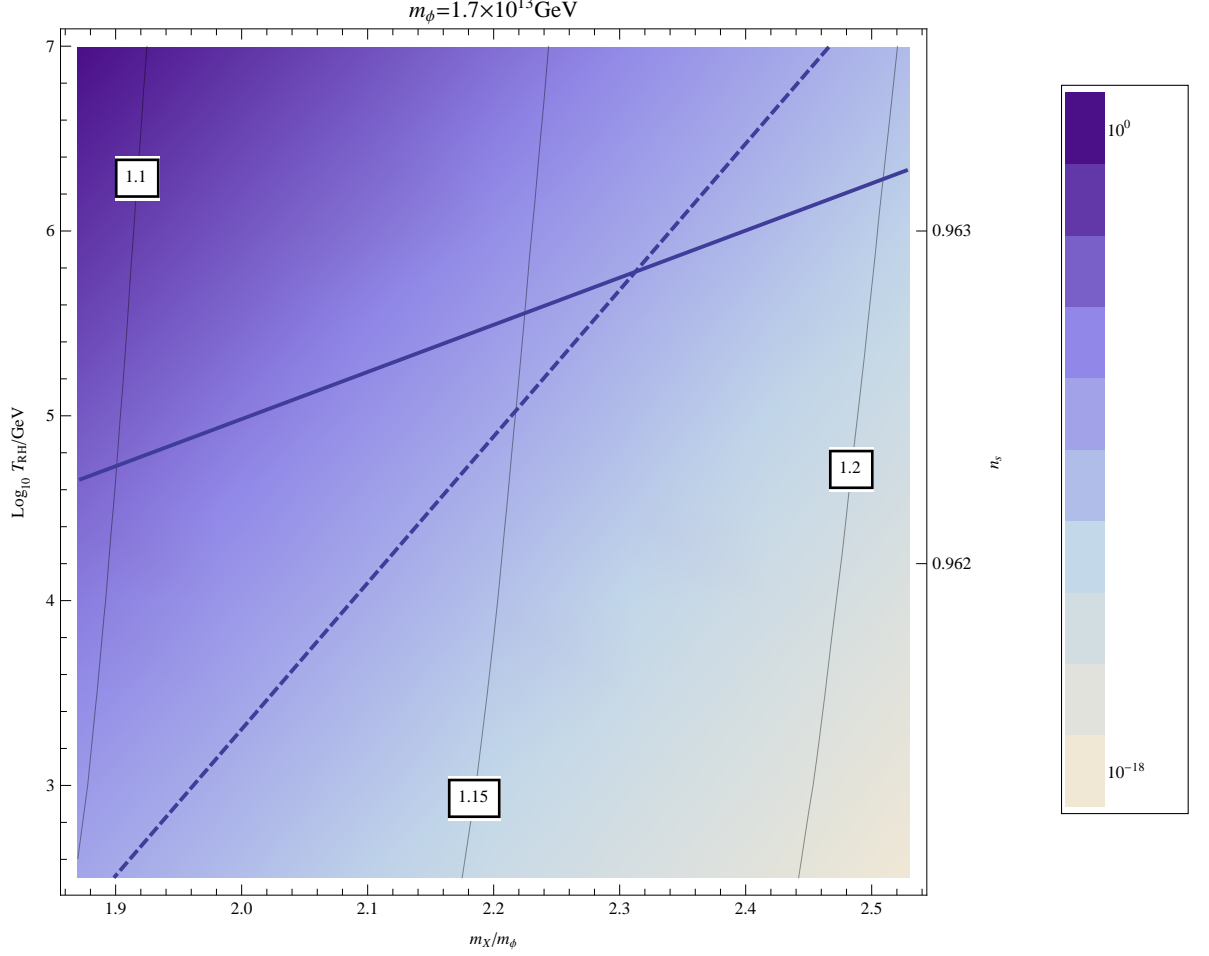
Figure 2.4: The right bottom of the parameter space under the thick lines is allowed by constraints: the current matter density (the solid line) and the upper limit ($\alpha \sim 0.07$) of the isocurvature power spectrum (the dashed line). The background color shows the power spectrum amplitude and the 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 $n_s$.

a robust bound on the reheating temperature of

$$T_{\mathrm{RH}} \lesssim 10^6 \text{ GeV}, \tag{2.58}$$

as well as a bound on the mass of the superheavy dark matter of

$$\frac{m_X}{m_\phi} \lesssim 2.3, \tag{2.59}$$

which corresponds t o$m_X \lesssim 4H_e$. 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. 2.5, 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. 2.4, the dashed curve in Fig. 2.5 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. 2.5, 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 $T_{\mathrm{RH}} \ll H_e$ implies a long period of matter domination that may lead to nontrivial dark matter and inflaton condensate clustering phenomenology (see e.g. [102–104]). As demonstrated in Sec. 2.4, 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.

## 2.6 Conclusions

In this work, 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-Gaussianities characterized by $f_{NL} \approx 30$ can result for a particular set of masses

and reheating temperatures without violating isocurvature bounds. The conditions that result in large $f_{NL}$ yield a bound on the reheating temperature of around $10^6$ GeV, as well as a bound on the dark matter mass of around $3H_e \lesssim m_X \lesssim 4H_e$. For lower allowed values of $m_X$ masses, the $X$ bosons can generate a large $f_{NL}$ 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 $m_\phi^2\phi^2/2$ 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 work 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 [105, 106], 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.

# 2.7   Appendix: Mode functions beyond the dS approximation

To calculate the correlation function or the energy density, it is necessary to solve the following mode equation (2.57). In dS space, where $H \equiv \dot{a}/a$ is constant, this has a well known solution with the Bunch-Davies boundary condition

$$X_k^{\mathrm{dS}}(t) \;\; = \;\; \frac{\sqrt{\pi}}{2a^{3/2}\sqrt{H}} e^{i\frac{\pi}{2}(\nu+1/2)} H_\nu^{(1)} \left( \frac{k}{aH} \right), \tag{2.60}$$

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

$$X_k^{\mathrm{dS}}(t) \;\; \approx \;\; -\frac{1}{\sqrt{\pi}}(-1)^{3/4} 2^{-1+\nu} e^{i\nu\pi/2} \left( \frac{k}{aH} \right)^{-\nu} \frac{\Gamma(\nu)}{\sqrt{Ha^3}}, \tag{2.61}$$

in which the decaying solution has been dropped. For $m_X/H > 3/2$, $|X_k^{\mathrm{dS}}|^2 \propto a^{-3}$ 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: $\Delta t \equiv t - t_* \lesssim 1/\epsilon H$, where $\epsilon \equiv -\dot{H}/H^2$. Conversely, Eq. (2.60) is a good approximation up to the time $t_*$ just after the horizon crossing. (2.57) can be approximated for $t > t_*$ as

$$\ddot{X}_k(t) + 3H(t)X_k(t) + m_X^2 X_k(t) \;\; \approx \;\; 0 \tag{2.62}$$

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

$$X_k(t) \;\; = \;\; X_k^{\mathrm{dS}}(t_*) \left( \frac{a(t_*)}{a(t)} \right)^{3/2} \exp \left[ \int_{t_*}^{t} dt_1\, H(t_1)\nu(t_1) \right]. \tag{2.63}$$

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

$$\frac{1}{H(t)} - \frac{1}{H(t_*)} \;\; = \;\; \epsilon(t - t_*), \tag{2.64}$$

$$a(t) \;\; = \;\; a(t_*)(1 + \epsilon H_* t)^{1/\epsilon}. \tag{2.65}$$

Hence, we find

$$\int_{t_*}^{t} dt' \, H(t')\nu(t') \;\; = \;\; \frac{1}{\epsilon}\left[\nu(t) - \frac{3}{2}\tanh^{-1}\left(\frac{3}{2\nu(t)}\right) - \nu_* + \frac{3}{2}\tanh^{-1}\left(\frac{3}{2\nu_*}\right)\right] \quad (2.66)$$

We note that the mode function $X_k$ is oscillatory for imaginary $\nu$, but that is not reflected in Eq. (2.66) because we have kept only the non-decaying mode in Eq. (2.61). 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

$$\begin{aligned}
|X_k(t)|^2 \;\; &\approx \;\; |X_k^{\mathrm{dS}}(t_*)|^2 \left(\frac{a_*}{a(t)}\right)^3 \\
&\quad \times \exp\left[\frac{3}{\epsilon}\Re\left\{\frac{2\nu(t)}{3} - \tanh^{-1}\left(\frac{3}{2\nu(t)}\right) - \frac{2\nu_*}{3} + \tanh^{-1}\left(\frac{3}{2\nu_*}\right)\right\}\right] \\
&= \;\; \frac{2^{-2+2\nu_*}}{\pi}\left(\frac{k}{a_* H_*}\right)^{-2\nu_*}\frac{|\Gamma(\nu_*)|^2}{H_* a^3(t)} \\
&\quad \times \exp\left[\frac{1}{\epsilon}\Re\left\{2\nu(t) - 2\nu_* + 3\ln\left[\frac{\left(\nu_* + \frac{3}{2}\right)\frac{H_*}{m_X}}{\left(\nu(t) + \frac{3}{2}\right)\frac{H(t)}{m_X}}\right]\right\}\right],
\end{aligned} \qquad (2.67)$$

in which the subscript $*$ indicates the variable is evaluated at $t_*$ and $\nu_*$ is a positive real number.

## 2.8  Appendix: Justification of the background

In situations in which quantum operators $\hat{\mathcal{O}}_\epsilon^{(q)}$ do not have classical expansion field values $\mathcal{O}_0^{(q)}$, it is necessary to justify how the $\mathcal{O}_0^{(q)}$ are to be identified. In such cases, the typical procedure is to solve for $\hat{\mathcal{O}}_\epsilon^{(q)}$ 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 $\mathcal{O}_0^{(q)}$ from matrix elements of $\hat{\mathcal{O}}_\epsilon^{(q)}$. Here we outline how to extract $\mathcal{O}_0^{(q)}$ 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 $\delta\mathcal{O}^{(q)}$ through $\mathcal{O}_{\epsilon}^{(q)} = \mathcal{O}_0^{(q)} + \delta\mathcal{O}^{(q)}$. If $\mathcal{O}_0^{(q)}$ is independent of spatial coordinates for some given fixed coordinate choice and $\delta\mathcal{O}^{(q)}$ satisfies the condition

$$\lim_{L\to\infty} \frac{1}{V_L} \int_{V_L} d^3x \delta\mathcal{O}^{(q)} = 0, \tag{2.68}$$

then we have

$$\mathcal{O}_0^{(q)} \approx \lim_{L\to\infty} \frac{1}{V_L} \int_{V_L} d^3x \mathcal{O}_{\epsilon}^{(q)}. \tag{2.69}$$

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 $\hat{\mathcal{O}}_0^{(q)}$.

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 $p_\Sigma(\{\mathcal{O}_{\epsilon}^{(q)}\})$ such that

$$\langle \hat{\mathcal{O}}_{\epsilon}^{(q_1)}...\hat{\mathcal{O}}_{\epsilon}^{(q_r)} \rangle = \int \prod_q D\mathcal{O}_{\epsilon}^{(q)} p_\Sigma(\{\mathcal{O}_{\epsilon}^{(q)}\}) \mathcal{O}_{\epsilon}^{(q_1)}...\mathcal{O}_{\epsilon}^{(q_r)}, \tag{2.70}$$

where the left hand side is computed in a fixed vacuum.[5] Since we are computing this at a fixed time, the right hand side functional measure $D\mathcal{O}_{\epsilon}^{(q)}$ is only discretized over the 3-space of $\Sigma$. It is important to think of an $\{\mathcal{O}_{\epsilon}^{(q)}\}$ element in the ensemble (governed by $p_\Sigma$) as a classical configuration only for a fixed time slice. The time evolution of $\{\mathcal{O}_{\epsilon}^{(q)}\}$ 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

---

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

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 $\mathcal{O}_\epsilon^{(q)}$ in Eq. (2.69)) 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 $p_\Sigma$ 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. (2.69) given $N$ samples in the ensemble governed by $p_\Sigma$. The number of configurations with a given characteristic $\Gamma$ is given by

$$N_\Gamma = N \int \prod_q D\mathcal{O}_\epsilon^{(q)} p_\Sigma(\{\mathcal{O}_\epsilon^{(q)}\})\delta(\Gamma)\det\frac{\delta\Gamma}{\delta\mathcal{O}_\epsilon}, \qquad (2.71)$$

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

$$\Gamma_{\mathcal{O}_0^{(q)}, V_L} \equiv \frac{1}{V_L(\Sigma)} \int_{V_L(\Sigma)} d^3x \mathcal{O}_\epsilon^{(q)} - \mathcal{O}_0^{(q)}, \qquad (2.72)$$

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

$$\frac{\delta\Gamma_{\mathcal{O}_0^{(q)}, V_L}}{\delta\mathcal{O}_\epsilon} = \frac{1}{V_L(\Sigma)}, \qquad (2.73)$$

which is independent of $\mathcal{O}_\epsilon$. The functional derivatives are taken only with respect

to functions of 3-spatial variable. Incorporating this into Eq. (2.71), we find

$$N_{\Gamma_{\mathcal{O}_0^{(q)}, V_L}} = \frac{N}{V_L(\Sigma)} \int \prod_q D\mathcal{O}_\epsilon^{(q)} p_\Sigma(\{\mathcal{O}_\epsilon^{(q)}\}) \delta\left(\frac{1}{V_L(\Sigma)} \int_{V_L(\Sigma)} d^3x \mathcal{O}_\epsilon^{(q)} - \mathcal{O}_0^{(q)}\right). \quad (2.74)$$

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

$$\begin{aligned} 0 &= \frac{\partial}{\partial \mathcal{O}_0^{(s)}(t)} N_{\Gamma_{\mathcal{O}_0^{(q)}, V_L}} \\ &= -\frac{N}{V_L(\Sigma)} \int \prod_q D\mathcal{O}_\epsilon^{(q)} p_\Sigma(\{\mathcal{O}_\epsilon^{(q)}\}) \delta^{qs} \\ &\quad \times \delta'\left(\frac{1}{V_L(\Sigma)} \int_{V_L(\Sigma)} d^3x \mathcal{O}_\epsilon^{(q)} - \mathcal{O}_0^{(q)}(t)\right), \end{aligned} \quad (2.75)$$

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

$$\delta\left[\frac{1}{V_L(\Sigma)} \int_{V_L(\Sigma)} d^3x \mathcal{O}_\epsilon^{(q)}\right] = \frac{\delta\mathcal{O}_\epsilon^{(q)}}{V_L(\Sigma)}. \quad (2.76)$$

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

$$\int \prod_q D\mathcal{O}_\epsilon^{(q)} \frac{\delta}{\delta\mathcal{O}_\epsilon^{(q)}} p_\Sigma(\{\mathcal{O}_\epsilon^{(q)}\}) \delta\left(\frac{1}{V_L(\Sigma)} \int_{V_L(\Sigma)} d^3x \mathcal{O}_\epsilon^{(q)} - \mathcal{O}_0^{(q)}\right) = 0. \quad (2.77)$$

If $\mathcal{O}_0^{(q)}$ is chosen to be $\mathcal{O}_*^{(q)}$ such that th e$\mathcal{O}_\epsilon^{(q)}$ configurations that satisfy

$$\mathcal{O}_*^{(q)} = \frac{1}{V_L(\Sigma)} \int_{V_L(\Sigma)} d^3x \mathcal{O}_\epsilon^{(q)} \quad (2.78)$$

also satisfy

$$\frac{\delta}{\delta\mathcal{O}_\epsilon^{(q)}} p_\Sigma(\{\mathcal{O}_\epsilon^{(q)}\}) = 0, \quad (2.79)$$

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

Consider

$$\langle \hat{\mathcal{O}}_\epsilon^{(s)} \rangle = \int \prod_q D\mathcal{O}_\epsilon^{(q)} p_\Sigma(\{\mathcal{O}_\epsilon^{(q)}\}) \mathcal{O}_\epsilon^{(s)}. \tag{2.80}$$

Assuming $p_\Sigma(\{\mathcal{O}_\epsilon^{(q)}\})$ is sharply peaked, consider

$$f \equiv \ln p_\Sigma(\{\mathcal{O}_\epsilon^{(q)}\}). \tag{2.81}$$

The peak is located at $\frac{\delta f}{\delta \mathcal{O}_\epsilon^{(q)}} = 0$ corresponding to the field configuration that satisfies

$$\left. \frac{\delta p_\Sigma(\{\mathcal{O}_\epsilon^{(q)}\})}{\delta \mathcal{O}_\epsilon^{(s)}} \right|_{\mathcal{O}_P^{(q)}} = 0. \tag{2.82}$$

Hence, the quadratic expansion of $f$ about the peak configuration is

$$
\begin{aligned}
f = & f(\mathcal{O}_P^{(q)}) + \frac{1}{2} \int dx_1^3 dx_2^3 \left. \frac{\delta^2 f}{\delta \mathcal{O}_\epsilon^{(q_1)}(x_1) \delta \mathcal{O}_\epsilon^{(q_2)}(x_2)} \right|_{\mathcal{O}_P^{(q)}} \\
& \times \left( \mathcal{O}_\epsilon^{(q_1)}(x_1) - \mathcal{O}_P^{(q_1)}(x_1) \right) \left( \mathcal{O}_\epsilon^{(q_2)}(x_2) - \mathcal{O}_P^{(q_2)}(x_2) \right),
\end{aligned} \tag{2.83}
$$

where the repeated $q_i$ indices are summed. $\mathcal{O}_\epsilon^{(s)}$ can be raised in Eq. (2.80) using the usual trick of introducing a source

$$\mathcal{O}_\epsilon^{(s)} \to \frac{\delta}{\delta J} \exp\left[ \int d^3x J \mathcal{O}_\epsilon^{(s)} \right]_{J=0} \tag{2.84}$$

and carrying out the leading saddle-point approximation integral to obtain

$$\langle \hat{\mathcal{O}}_\epsilon^{(s)}(x) \rangle = \mathcal{O}_P^{(s)}(x), \tag{2.85}$$

where we used

$$
\begin{aligned}
1 = & \int \prod_q D\mathcal{O}_\epsilon^{(q)} p_\Sigma(\{\mathcal{O}_\epsilon^{(q)}\}) \tag{2.86} \\
\approx & \ e^{f(\mathcal{O}_*^{(q)})} \int \prod_q D\mathcal{O}_\epsilon^{(q)} \times \\
& \exp\left[ \frac{1}{2} \int dx_1^3 dx_2^3 \left. \frac{\delta^2 f}{\delta \mathcal{O}_\epsilon^{(q_1)}(x_1) \delta \mathcal{O}_\epsilon^{(q_2)}(x_2)} \right|_{\mathcal{O}_P^{(q)}} \right. \\
& \left. \times \left( \mathcal{O}_\epsilon^{(q_1)}(x_1) - \mathcal{O}_P^{(q_1)}(x_1) \right) \left( \mathcal{O}_\epsilon^{(q_2)}(x_2) - \mathcal{O}_P^{(q_2)}(x_2) \right) \right]
\end{aligned} \tag{2.87}
$$

We note that if we use a spatially translation invariant state to take the expectation value, $\mathcal{O}_P^{(q)}(x)$ is automatically spatially translation invariant and thus

$$\mathcal{O}_*^{(q)} = \mathcal{O}_P^{(q)} \tag{2.88}$$

in Eq. (2.78). Hence, Eqs. (2.85) and (2.88) combine to give the most probable spatially averaged configuration to be

$$\mathcal{O}_*^{(s)} = \langle \hat{\mathcal{O}}_\epsilon^{(s)} \rangle \tag{2.89}$$

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 $\langle \hat{\mathcal{O}}_\epsilon^{(s)} \rangle$. It is important to remember that since $\mathcal{O}_*^{(s)}$ (in Eq. (2.78)) is being identified with $\mathcal{O}_0^{(q)}$ (in Eq. (2.69)), we are now assuming that the $\epsilon = 0$ solution of perturbation theory is governed by equations that depend only on time for the coordinate system that we have chosen.

# Chapter 3

# Isocurvature from Fermion Production

Isocurvature perturbations in the inflationary literature typically involve quantum fluctuations of bosonic field degrees of freedom. In this work, we consider isocurvature perturbations from fermionic quantum fluctuations during inflation. When a stable massive fermion is coupled to a non-conformal sector different from the scalar metric perturbations, observably large amplitude scale invariant isocurvature perturbations can be generated. In addition to the computation of the isocurvature two-point function, an estimate of the local non-Gaussianities is also given and found to be promising for observations in a corner of the parameter space. The results provide a new class of cosmological probes for theories with stable massive fermions. On the technical side, we explicitly renormalize the composite operator in curved spacetime and show that gravitational Ward identities play an important role in suppressing certain contributions to the fermionic isocurvature perturbations.

This work was performed in collaboration with Daniel Chung and Peng Zhou.

## 3.1   Introduction

The Cosmic Microwave Background (CMB) measurements [96,107–116] and the Large Scale Structure (LSS) observations [117,118] are consistent with single field inflationary models which can seed approximately adiabatic, scale-invariant, and Gaussian primordial density perturbations [8, 10, 11, 119–124]. However, from the multi-field nature of the Standard Model of particle physics, one may naturally guess that there would be more than one light degrees of freedom during inflation which may be responsible for generating isocurvature primordial perturbation initial conditions. Indeed, in any slow-roll inflationary scenario, non-inflaton degrees of freedom must eventually turn on in order to reheat successfully.[1] Hence, isocurvature scenarios are theoretically well motivated.

Isocurvature perturbations have been studied in various scenarios, such as double inflation [125–128], curvaton scenario [41, 56, 129–141], axions [142–151] and gravitationally produced superheavy dark matter [18, 40, 65, 67]. Isocurvature perturbations also can generate rich density perturbation phenomenology. For example, unlike standard single field inflationary scenarios, isocurvature perturbations are able to generate large primordial local non-Gaussianities [18, 40, 47, 51, 59, 132, 152–174]. However, most previous studies of isocurvature perturbations focused on bosonic degrees of freedom such as axions and curvatons. Fermionic isocurvature degrees of freedom such as gravitinos were only discussed in the literature associated with the decay products of the inflaton or other scalars [57, 175–181]. Furthermore, these fermions discussed in the literature were characterized only by their dependence on the entropy temperature fluctuation $\delta T$ which was assumed to be directly linked to the curvature perturbation $\zeta$, in a manner consistent with the "separate universe" picture of

---

[1]Even though the reheat degrees of freedom do not need to be dynamically important during the quasi-dS era, multiple fields are certainly lurking in the scenario.

$\delta N$ formalism [182–184]. Such previously discussed fermionic isocurvature scenarios lead to strong correlation or anticorrelation with the curvature perturbation $\zeta$. One can intuitively characterize these previous fermionic isocurvature works as having no fermionic quantum fluctuation information from the inflationary era.

In contrast, we examine in this work a fermionic isocurvature scenario that is not (significantly) correlated with $\zeta$ and has fermionic quantum fluctuation information during inflation encoded in the isocurvature correlator. In our scenario, the horizon length scale interaction dynamics of the fermion particles is important, in sharp contrast with the "separate universe" picture of $\delta N$ formalism. As we will show, although classical gravitational field interactions alone are sufficient to generate enough fermions during the exit process of inflation to saturate the phenomenologically required cold dark matter abundance [66, 185, 186], fermion propagators in the classical FRW background is insufficient to produce any observable isocurvature perturbations because of the fact that massless fermions enjoy a classical conformal symmetry.[2] Hence, any large fermion isocurvature correlator must involve couplings to a conformal symmetry breaking sector.

For illustrating the existence of such fermionic isocurvature perturbations, we minimally extend the single field inflation by adding a stable massive fermion field coupled through a Yukawa coupling to a light non-inflaton scalar field whose mass is much lighter than the fermion field (hence, there are no decays of the scalars to the fermions). The light non-inflaton scalar field (which is minimally coupled to gravity) serves as a conformal symmetry breaking sector through which the fermions will attain appreciable correlations. We compute the isocurvature two-point function of fermions that are gravitationally produced during inflation and identify the phenomenologically viable parameter space. We also estimate the local non-Gaussianity and show that

---

[2]Even with the massive fermions, we will be naturally concerned with light fermions where $m_\psi/H \ll 1$.

it may be observationally large in a particular parametric regime.

At the technical level, treating fermionic isocurvature fluctuations during inflation requires composite operator renormalization in quasi-dS spacetime because the fermionic energy-momentum tensor is a composite bilinear operator (i.e. fermions cannot get VEVs) and the leading two-point function contribution involves a one loop 1PI diagram. To our knowledge, this work is the first one to apply composite fermion operator renormalization in inflationary spacetime to treat isocurvature perturbations. Indeed, an improper treatment of the operator renormalization can in principle lead to answers that are many orders of magnitude off as we pointed out with bosonic composite operators [94]. We also show that a gravitational Ward identity plays an important role in suppressing the scalar metric perturbation interaction contribution to the isocurvature two-point function (thereby justifying our introduction of another scalar sector).

This work is presented in the following order. In Sec. 3.2, we motivate and discuss the fermion isocurvature model. Next, we review the definition of the gauge-invariant variables and the quantum operator associated with the cold dark matter (CDM) isocurvature in Sec. 3.3. In subsection 3.3, we present the regulator and the renormalization conditions for our isocurvature operator. We explain the constraints on the Yukawa coupling coming from the self-consistency of our simplified scenario in Sec. 3.4. In Sec. 3.5, we compute the isocurvature 2-point function. The leading order and the next leading order results are given in subsection 3.5 and 3.5, and the power spectrum is presented in subsection 3.5. In Sec. 3.6, we discuss the numerical implications of our results and non-Gaussianities. Afterwards in Sec. 3.7, we discuss the explicit computation of how a diffeomorphism Ward identity plays a role in suppressing the scalar metric perturbation contribution to the isocurvature two-point function. Finally, in Sec. 3.8 we summarize and conclude. Some technical details of

the computations are given in the Appendices.

## 3.2  Fermion Isocurvature Model

As is well known, if any small mass fermion field degrees of freedom exist during infla-
tion which is usually assumed to be a Bunch-Davies vacuum state, fermion particles
will be produced gravitationally (see e.g. [6,66,185–187]). The inhomogeneities of the
gravitationally produced fermions will generically not align with the inhomogeneities
of the inflaton, depending on its interactions. If most of the radiation in the universe
comes from the inflaton decay, then the misalignment of the inhomogeneities of the
fermions and the inflaton will lead to isocurvature perturbations [3, 188, 189].

Now, to motivate our fermion model with Yukawa interactions, it is important to
understand why interactions to conformal symmetry breaking sector is required. It is
also well known that massless fermion classical action enjoys a conformal symmetry:

$$g_{\mu\nu} \to e^{2\sigma(x)} g_{\mu\nu} \tag{3.1}$$

$$\psi \to e^{-3\sigma(x)/2} \psi. \tag{3.2}$$

Since FRW spacetime can be written as a conformal transformation of the Minkowski
space (i.e. $a = \exp(\sigma)$), we would expect for a tree level fermion propagating on an
FRW spacetime without any interactions with a conformal symmetry breaking sector

$$\langle \bar{\psi}\psi(t, \vec{x})\bar{\psi}\psi(t, \vec{y}) \rangle_{conn} = \langle \bar{\psi}_M \psi_M(t, \vec{x})\bar{\psi}_M \psi_M(t, \vec{y}) \rangle_{conn} a^{-6} \tag{3.3}$$

where $\psi_M$ is the Minkowski fermion. At leading order, there are no other scales in
this function except $|\vec{x} - \vec{y}|$. Hence, we conclude

$$\langle \bar{\psi}\psi(t, \vec{x})\bar{\psi}\psi(t, \vec{y}) \rangle_{conn} \sim \frac{1}{a^6 |\vec{x} - \vec{y}|^6} \tag{3.4}$$

in the massless limit.[3] We expect this to be the dominant contribution in the limit that $m_\psi/H \ll 1$. When $m_\psi/H \gg 1$, we also expect there can be factors multiplying this that vanishes exponentially fast as $m_\psi/H \to \infty$ (we show this explicitly in Sec. 3.5). Hence, we expect Eq. (3.4) to be the leading order of magnitude composite correlator if the theory is approximately conformally invariant. As we will show below, the comoving gauge isocurvature perturbations is proportional to

$$\langle \frac{\delta\rho_\psi^{(C)}}{\bar{\rho}_\psi} \frac{\delta\rho_\psi^{(C)}}{\bar{\rho}_\psi} \rangle \sim \frac{\langle \bar{\psi}\psi(t,\vec{x})\bar{\psi}\psi(t,\vec{y}) \rangle_{conn}}{\langle \bar{\psi}\psi \rangle^2}. \tag{3.5}$$

where one sees the appearance of the suppressed correlator computed in Eq. (3.4). Because of this suppression, fermionic isocurvature perturbations require nontrivial interactions with a conformal symmetry breaking sector.

If the conformal symmetry breaking sector is just the $\zeta$ sector of the inflaton, then its effective coupling to the fermions is suppressed because there is an infinitesimal shift symmetry of the $\zeta$ coming from a residual diffeomorphism symmetry in the comoving gauge. (We will explain this explicitly in Sec. 3.7 in terms of a Ward identity.) Hence, to generate an observable fermionic correlator during the horizon exit, another conformal symmetry breaking sector must be introduced which does not suffer from derivative coupling suppression similar to $\zeta$.[4] We thus introduce a Yukawa coupling to a light non-inflaton scalar and demonstrate that this interaction can induce observable isocurvature fluctuations.[5]

Given this motivation, let us now specify the model studied in this work. We use one real scalar $\phi$ slow-roll inflaton degree of freedom that dominates the energy

---

[3] The scaling behavior of the two-point correlator is similar to that of correlators considered in Ref. [190] in the context of conformal field theories.

[4] Although we have not investigated the suppression for the tensor perturbation interactions with a full computation, we expect a similar suppression of the tensor perturbation interactions.

[5] Note that this introduction of a light scalar is not particularly attractive from a model building perspective since we provide no explicit mechanism to protect its light mass. We defer the challenge of building an attractive model to a future work since the purpose of this work is to demonstrate the basic physics mechanism.

density during inflation and then perturbatively decays to the SM sector to reheat the universe. We also introduce another minimally coupled light real scalar degree of freedom $\sigma$ which has no coupling to $\phi$ or the SM sector (necessary for reheating) stronger than gravity.[6] As we explained, the main role of $\sigma$ is to provide a conformal symmetry breaking sector which can couple to the Dirac fermions $\psi$ through a Yukawa coupling. We assume $\psi$ is charged under a conserved discrete charge such that the one particle states are stable and can act as dark matter. Note that since we do not require all of the dark matter to come from the fermions, this system is consistent with the existence of the weakly interacting massive particle (WIMP) dark matter. Because $\psi$ is too weakly interacting with the SM to be produced directly, gravitational production of $\psi$ during and after inflation is significant and gives rise to non-thermal cold dark matter (CDM) and its isocurvature perturbations.

Such a model is described by the action[7]

$$
\begin{aligned}
S \;=\; & \int (dx) \left\{ \mathcal{L}_{inf}\left[g_{\mu\nu}, \phi\right] + \mathcal{L}_{SM+CDM}\left[g_{\mu\nu}, \{\Psi\}\right] + \mathcal{L}_{RH}\left[g_{\mu\nu}, \phi, \{\Psi\}\right] \right. \\
& \left. + -\frac{1}{2}g^{\mu\nu}\partial_\mu\sigma\partial_\nu\sigma - \frac{1}{2}m_\sigma^2\sigma^2 - \frac{y}{4!}\sigma^4 + \bar{\psi}(i\gamma^a\nabla_{e_a} - m_\psi)\psi - \lambda\sigma\bar{\psi}\psi \right\}, \quad (3.6)
\end{aligned}
$$

where $M_p^2 = \frac{1}{8\pi G} = 1$, $(dx) \equiv \sqrt{-g}d^4x$, and $\mathcal{L}_{inf}$ and $\mathcal{L}_{SM+CDM}$ are the Lagrangians for the inflaton and the SM+CDM sectors, and $\mathcal{L}_{RH}$ describes the sector responsible for reheating. Because an interesting parameter region exists for our scenario in which the $\psi$ constitute a tiny fraction of the total dark matter content, the Lagrangian $\mathcal{L}_{SM+CDM}$ describes the CDM sector different from $\psi$ to make the scenario phenomenologically viable. Note that natural heavy dark matter candidates for $\psi$ exist in the context of string phenomenology [69, 70]. Furthermore, many extensions of the Standard Model also possess superheavy dark matter candidates (see,

---

[6]For now, we will consider this as a tuning and will not address serious model building issues in this work. It is plausible that this kind of scenario can be realized in the context of SUSY hidden sector.

[7]Our metric convention is $(-, +, +, +)$.

e.g., [71–81]). Since there are many scalar field degrees of freedom in typical BSMs, the possibility of identifying one of these scalars with $\sigma$ is also plausible. Although the cosmological phenomenology of weakly interacting dark matter on large scales have been investigated already in literature (see, e.g., [18,40,41,67,129,149–151,164,191]), our work is the first to describe fermionic fluctuation correlations during inflation. Note that although Eq. (3.6) has a quartic term $\sigma^4$, we will focus on the parametric region in which the quartic coupling $y$ will be small and tuned against radiative generated quartic couplings from the Yukawa interaction to keep the effects of the $\sigma$ interactions to a minimum. Hence, our effective parametric domain will be controlled by $\{\lambda, m_\sigma, m_\psi\}$.

We focus on a particular parametric region of $\{\lambda, m_\sigma, m_\psi\}$ such that $\sigma$ only assists in generating large scale density perturbations of $\psi$, and the density perturbations and the relic abundance from the $\sigma$ particles vanish or are suppressed compared to those from the $\psi$ particles. For example, requiring the correlator $\langle \sigma\sigma \rangle|_{t_*}$ relevant for the isocurvature perturbations not be suppressed gives the condition $m_\sigma/H(t_*) < 1$ where $t_*$ is the time at which the fermion production ends. This implies $m_\sigma < m_\psi$ is the relevant parameter region. Furthermore, in order to prevent any large isocurvature perturbations and relic abundance of $\sigma$, we assume that the $\sigma$ particles decay before $\sigma$ becomes an important fluid component of the evolution of the universe (e.g. before matter-radiation equality). Note however that this restriction is a matter of simplicity. In general, we note that a weakly interacting and stable $\sigma$ may also be phenomenologically allowed without problems regarding the relic abundance and the isocurvature from $\sigma$. Moreover, for simplicity, we restrict $\lambda$ such that 1) $\sigma\sigma \to \bar{\psi}\psi$ via the Yukawa interactions is suppressed compared to the gravitational process in producing $\bar{\psi}\psi$ 2) any $\sigma + \text{gravity} \to \bar{\psi}\psi$ processes are estimated to be unimportant. This restriction is approximately equivalent to being in a parametric region where

tree-level propagator neglecting resumption of the Yukawa interactions is valid.

In addition, in order to detach our model from the details of the inflationary model of $\phi$, we focus on the light fermion $\psi$, such that $m_\psi < H_e$, where $H_e$ is the Hubble scale at the end of inflation. This is because the gravitational particles production is generally sensitive to how the inflation ends in a such way that an extra suppression factor $\exp\left(-cm_\psi^2/H_e^2\right)$ (where $c$ is a number depending on how the inflation connected with the post inflationary era) appears in the estimation of the gravitationally produced particle number density $n_\psi$. (Throughout the work, we will sometime distinguish $H_e$ from $H_{inf}$ which is defined to be the expansion rate during inflation.) On the other hands, if $m_\psi < H_e$, the factor becomes simply an $O(1)$ number, and particularly, for fermions we can estimate the number density $n_\psi(t_*)$ as $O(0.1)m_\psi^3$ at $H(t_*) \sim m_\psi$ regardless of how the inflation ends [185]. The physics of this universality is tied to the conformal symmetry of the fermions in the massless limit.

At this point, we emphasize that our model is different from other fermionic (e.g., gravitino) isocurvature models in literature (e.g. [57, 181, 192]). We explicitly predict the amplitudes of fermion density perturbations from a joint effect of the gravitational particle production and $\sigma$ modulation on $m_\psi$ via the matter loop diagrams. In contrast, in Refs. [57, 181, 192] the fermions are produced from the on shell inflatons and/or curvatons (the latter has the closest identification in our model with $\sigma$) after the end of inflation. A sharp observable contrast of our model with these other models is that our scenario predicts an uncorrelated type of isocurvature (i.e. curvature-isocurvature cross correlation is negligible) while these other models purportedly generate correlated type of isocurvature. This is a consequence of the fact that these other models do not describe any fermionic fluctuations during inflation while in our model, the expansion during inflation imparts work to virtual fermionic

fluctuations to put them on shell.

## 3.3  Operator for Isocurvature Perturbation

Recall that the scalar perturbation of the metric is parametrized as

$$
\delta g_{\mu\nu}^{(S)} = \begin{pmatrix} -E & aF_{,i} \\ aF_{,i} & a^2[A\delta_{ij} + B_{,ij}] \end{pmatrix}.
\tag{3.7}
$$

The gauge-invariant variables are constructed by combining metric perturbations and other perturbations, such as density perturbations. For example, the conventional first-order gauge-invariant perturbation associated with the energy density of a fluid $a$ is defined (see, e.g., [1,3] and references therein) by

$$
\zeta_a \equiv \frac{A}{2} - H\frac{\delta\rho_a}{\dot{\bar{\rho}}_a}.
\tag{3.8}
$$

In particular, we define the conventional curvature perturbation as

$$
\zeta \equiv \frac{A}{2} - H\frac{\delta\rho_{tot}}{\dot{\bar{\rho}}_{tot}},
\tag{3.9}
$$

where

$$
\delta\rho_{tot} = \sum_i \delta\rho_i, \quad \bar{\rho}_{tot} = \sum_i \bar{\rho}_i.
\tag{3.10}
$$

This quantity $\zeta$ is conserved when modes are stretched out of the horizon even through the reheating era as long as it is set by the adiabatic initial condition, i.e., $\zeta = \zeta_a$ for any fluid $a$. Furthermore, if perturbations are generated solely by inflaton during inflation, such as the single field inflation, superhorizon perturbations automatically satisfy the adiabatic initial condition and the perturbations are conserved so that we can match them with those during the early radiation dominated (RD) era, $\zeta_\phi(t_{inf}) = \zeta_\gamma(t_{RD}) = \zeta_m(t_{RD}) = \cdots$.

On the other hand, an isocurvature perturbation is defined by a relative density perturbation between two different fluids

$$\delta_{Sij} \equiv 3\left(\zeta_i - \zeta_j\right) = -3H\left(\frac{\delta\rho_i}{\dot{\bar{\rho}}_i} - \frac{\delta\rho_j}{\dot{\bar{\rho}}_j}\right). \tag{3.11}$$

In general, it may arise during inflation if there are more than one degree of freedom. Although their mixing with perturbations of different fluids can lead to the failure of the conservation of the curvature perturbation $\zeta$, such effects are negligible as for any species $i$ whose $\bar{\rho}_i + \bar{P}_i$ is sufficiently smaller than $\bar{\rho}_{\text{tot}} + \bar{P}_{\text{tot}}$ until the Universe reaches radiation domination. Particularly, for gravitationally produced fermions, we have

$$\left.\frac{\bar{\rho}_\psi + \bar{P}_\psi}{\bar{\rho}_{\text{tot}} + \bar{P}_{\text{tot}}}\right|_{t_*} \sim \left.\frac{\bar{\rho}_\psi}{\bar{\rho}_{tot}}\right|_{t_*} \sim \frac{m_\psi^2}{M_p^2} \ll \Delta_\zeta^2, \tag{3.12}$$

where $t_*$ is the time that the gravitational fermion production ends, $H(t_*) \sim m_\psi$. Hence, we expect the superhorizon curvature perturbation to be approximately conserved through the reheating, $\zeta(t_{RD}) \approx \zeta_\phi(t_{inf})$.

The dominant fraction of the produced fermions are non-relativistic.[8] Then the fermion energy density behaves as [9]

$$\frac{d}{dt}\bar{\rho}_\psi(t) \approx -3H\bar{\rho}_\psi \quad \text{for } t > t_*, \tag{3.13}$$

and from Eq. (3.11) a general CDM isocurvature is written as

$$\delta_S = \frac{\delta\rho_{CDM}}{\bar{\rho}_{CDM}} - \frac{3}{4}\frac{\delta\rho_\gamma}{\bar{\rho}_\gamma}. \tag{3.14}$$

As discussed in Sec. 3.2, the CDM may include decay products of the inflaton $\phi$. Thus the CDM density perturbation is generally expressed as

$$\frac{\delta\rho_{CDM}}{\bar{\rho}_{CDM}} = \omega_\psi\frac{\delta\rho_\psi}{\bar{\rho}_\psi} + (1 - \omega_\psi)\frac{\delta\rho_m}{\bar{\rho}_m}, \tag{3.15}$$

---

[8]This is a valid assumption because gravitationally excited fermion modes that contributions to the energy density are less than the fermion mass, i.e., $|\beta_k|^2$ for $k/a \lesssim m_\psi$, where $\beta_k$ is the Bogoliubov coefficient. See Appendix 3.10 for the detail.

[9]One can find that $\bar{\rho}_\psi \propto a^{-3}(t)$ for $t > t_*$ if $\bar{\rho}_\psi$ is renormalized by the adiabatic subtraction. See Appendix 3.10 and Ref. [6]. Then we can treat $\psi$ as a pressure less matter.

where the subscript $m$ denotes the CDM component associated with the inflaton decay products (such as WIMPs of minimal supersymmetric models), and

$$\omega_\psi \equiv \bar{\rho}_\psi / \left( \bar{\rho}_\psi + \bar{\rho}_m \right). \tag{3.16}$$

In particular, in the comoving gauge ($\delta\rho_\phi/\dot{\bar{\rho}}_\phi = \delta\rho_m/\dot{\bar{\rho}}_m = \delta\rho_\gamma/\dot{\bar{\rho}}_\gamma = 0$), the CDM isocurvature becomes

$$\delta_S^{(C)} \approx \omega_\psi \frac{\delta\rho_\psi^{(C)}}{\bar{\rho}_\psi}, \tag{3.17}$$

where the superscript denotes the gauge choice.

Under the non-relativistic assumption, we also approximate the fermion mass term $m_\psi \bar{\psi}\psi$ as its energy density[10]

$$\rho_\psi \approx m_\psi \bar{\psi}\psi, \tag{3.18}$$

and then the the fermion isocurvature perturbation becomes

$$\delta_S^{(C)} \approx \omega_\psi \frac{\rho_\psi - \langle \rho_\psi \rangle}{\langle \rho_\psi \rangle} = \omega_\psi \frac{\bar{\psi}\psi - \langle \bar{\psi}\psi \rangle}{\langle \bar{\psi}\psi \rangle}. \tag{3.19}$$

Notice that as it is a quantum composite operator, we renormalize it with regulators and counter-terms invariant under the underlying gauge symmetry, diffeomorphism in this case. In the following subsection, we present the technical detail of the composite operator renormalization. From now on, we will use the comoving gauge in calculating the correlation function and drop the superscript $(C)$ for convenience.

---

[10]Using the adiabatic vacuum prescription, the renormalized energy density is approximated in the non-relativistic case as

$$\langle (\rho_\psi)_r \rangle \approx m_\psi \langle N_\psi \rangle = 2m_\psi \int \frac{d^3k}{(2\pi^3)} \frac{1}{a^3} |\beta_k|^2,$$

where $N_\psi$ is a fermion number operator, and the subscript r denotes that the operator is a renormalized composite operator. This quantity is in accord with

$$m_\psi \langle (\bar{\psi}\psi)_r \rangle = 2m_\psi \int \frac{d^3k}{(2\pi^3)} \frac{m_\psi}{\omega_p} |\beta_k|^2 \approx 2m_\psi \int \frac{d^3k}{(2\pi^3)} |\beta_k|^2.$$

In particular, $(\bar{\psi}\psi)$ has an advantage in constructing gauge-invariant variables because it is manifestly 4-scalar, but $N_\psi$.

# Regularization and Renormalization for Isocurvature Perturbation

In this subsection, we explain our regularization procedure and renormalization scheme that determines the counter-terms. The most crucial renormalization condition that the isocurvature perturbations are sensitive to is Eq. (3.38).

For the convenience of preserving covariance and incorporating the adiabatic vacuum boundary condition, we use Pauli-Villars (PV) regularization [193, 194]. This involves the replacements

$$\psi \to \psi + \sum_n \psi_n, \qquad \sigma \to \sigma + \sum_n \sigma_n, \tag{3.20}$$

and the addition of the Pauli-Villars part in the free Lagrangian

$$\mathcal{L}_{PV} = \sum_{n=1} C_n\left(-\frac{1}{2}g^{\mu\nu}\partial_\nu\sigma_n\partial_\nu\sigma_n - \frac{1}{2}M_n^2\sigma_n^2\right) \tag{3.21}$$

$$+ \sum_{n=1} D_n\bar{\psi}_n(i\gamma^a\nabla_a - m_n)\psi_n. \tag{3.22}$$

For notational simplicity, we let $C_0 = 1$, $M_0 = m_\sigma$ and $D_0 = 1$, $m_0 = m_\psi$, and let index $N = 0, 1, \cdots$ and $n = 1, 2, \cdots$. We require the following constraints for scalar regulators

$$\sum_N C_N^{-1} = 0, \quad \sum_N C_N^{-1}M_N^2 = 0, \quad \sum_N C_N^{-1}M_N^4 = 0, \cdots \tag{3.23}$$

and the following constraints for fermion regulators

$$\sum_N D_N^{-1} = 0, \quad \sum_N D_N^{-1}m_N = 0, \quad \sum_N D_N^{-1}m_N^2 = 0, \cdots \tag{3.24}$$

where we need to introduce sufficient numbers of PV fields and constraints to cancel all the divergences. Notice that the additional constraints in the fermions with odd powers of $m_N$.
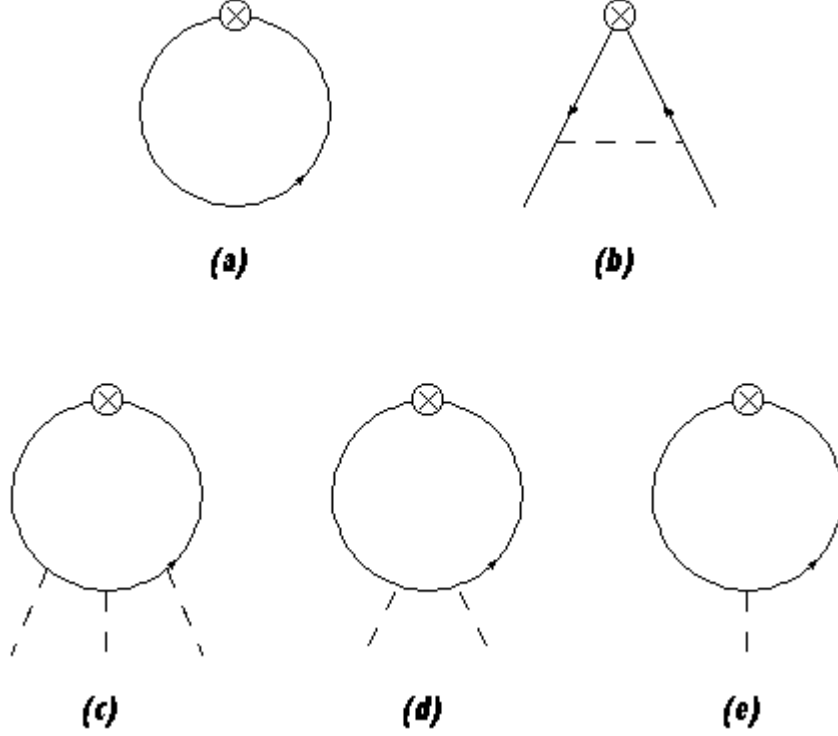
Figure 3.1: Diagrams determining counter-terms where the solid line corresponds to the fermion line and the dashed lines corresponds to $\sigma$ lines. It is convenient to truncate the external $\sigma$ legs on diagrams c), d), and e) with zero momentum insertion, making these mass insertions.

With the operator dimension and the symmetry considered, the renormalized operator is written as

$$
\begin{aligned}
(\bar{\psi}\psi)_{x,r} &= (\bar{\psi}_x)_r(\psi_x)_r(1+\delta Z_1) + \delta Z_2(\sigma_{x,r})^3 + \delta Z_3(\sigma_{x,r})^2 \\
&\quad + \delta Z_4 \sigma_{x,r} + \delta Z_5 + \delta Z_6 \Box \sigma_{x,r} + \delta Z_7 R + \delta Z_8 R \sigma_{x,r}
\end{aligned} \tag{3.25}
$$

where each field operator should be understood as including a sum of the PV fields as in Eq. (3.20). Then we give the renormalization conditions to determine the counter terms. For $\delta Z_i$ which are not coupled to $R, R_{\mu\nu}, R^{\alpha}_{\beta\mu\nu}$ and their derivatives, we can go to the Minkowski space and impose the renormalization conditions there. (Of course, we do not need to separate the curved space contribution and the flat space contribution with two computations, but we present this here this way here for clarity

in the physical partition.) We define the renormalized operator $\bar{\psi}\psi$ at one-loop order, such that it measures the number density of the fermion particles. First, we require its expectation value in the flat space vacuum to vanish:

$$\langle vac|\bar{\psi}\psi(x)|vac\rangle_{flat} + \sum_{n=1}\langle vac|\bar{\psi}_n\psi_n(x)|vac\rangle_{flat} + \delta Z_5 \;=\; 0 \qquad (3.26)$$

$$\Rightarrow -\int \frac{d^4p}{(2\pi)^4}\sum_N D_N^{-1}\text{Tr}\left\{\frac{1}{i}\frac{-\not{p}+m_N}{p^2+m_N^2-i\epsilon}\right\} + \delta Z_5 \;=\; 0. \qquad (3.27)$$

This corresponds to the evaluation of diagram (a) in Fig. 3.1.

Next, we impose the renormalization condition consistent with the fact that as far as the fermion sector is concerned, a shift of $\sigma$ by a constant in the tree-level action is equivalent to a shift in the mass of the fermion. More explicitly, we demand that if $\sigma$ is shifted as $\sigma \to \sigma + c$, the one-point function satisfies

$$\langle vac|(\bar{\psi}\psi)_{x,r}|vac\rangle_{flat} = \langle vac|\left[(\bar{\psi}\psi)_{x,r} + \Delta(\bar{\psi}\psi)_{x,r}\right]|vac\rangle_{flat,\mathcal{L}_I=-\lambda c\bar{\psi}_y\psi_y} \qquad (3.28)$$

where $\Delta(\bar{\psi}\psi)_{x,r}$ corresponds to a shift in the $\sigma$ dependent composite operator counterterms and $\mathcal{L}_I$ corresponds to $c$ dependent mass shift Lagrangian term. This leads to diagrams (c)-(e) in Fig. 3.1 with the external $\sigma$ propagators truncated and fixes $\delta Z_2, \delta Z_3, \delta Z_4$:

$$-(-i\lambda)^3\int\frac{d^4k}{(2\pi)^4}\text{Tr}\left\{\left(\sum_M D_M^{-1}\frac{1}{i}\frac{-\not{k}+m_M}{k^2+m_M^2-i\epsilon}\right)^4\right\} + \delta Z_2 \;=\; 0 \qquad (3.29)$$

$$-(-i\lambda)^2\int\frac{d^4k}{(2\pi)^4}\text{Tr}\left\{\left(\sum_M D_M^{-1}\frac{1}{i}\frac{-\not{k}+m_M}{k^2+m_M^2-i\epsilon}\right)^3\right\} + \delta Z_3 \;=\; 0, \qquad (3.30)$$

and

$$-i\lambda\int d^4y\langle(\bar{\psi}\psi)_x(\bar{\psi}\psi)_y\rangle + \delta Z_4 \;=\; 0 \qquad (3.31)$$

$$\Rightarrow -i\lambda\int\frac{d^4k}{(2\pi)^4}(-)\text{Tr}\left\{\left(\sum_M D_M^{-1}\frac{1}{i}\frac{-\not{k}+m_M}{k^2+m_M^2-i\epsilon}\right)^2\right\} + \delta Z_4 \;=\; 0. \qquad (3.32)$$

Furthermore, we require $\bar{\psi}\psi$ to have no loop corrections when contracted with on-shell fermion. This leads to the diagram (b) of Fig. 3.1 (where we have set the composite operator momentum to be 0 for convenience) and fixes $\delta Z_1$:

$$\delta Z_1 + (i\lambda)^2 \int \frac{d^4k}{(2\pi)^4} \sum_{L,M,N} C_L^{-1} D_M^{-1} D_N^{-1} \frac{1}{i} \frac{1}{k^2 + M_L^2 - i\epsilon}$$

$$\times \frac{1}{i} \frac{[-\slashed{k} - \slashed{p} + m_M]}{(k+p)^2 + m_M^2 - i\epsilon} \times \frac{1}{i} \frac{[-\slashed{k} - \slashed{p} + m_N]}{(k+p)^2 + m_N^2 - i\epsilon} = 0. \tag{3.33}$$

Similarly, we demand $\bar{\psi}\psi$ to have no loop corrections when contracted with on-shell scalar line. Explicitly, the diagram corresponds to the diagram (e) of Fig. 3.1 determining $\delta Z_6$ :

$$-i\lambda \int d^4y \langle (\bar{\psi}\psi)_x (\bar{\psi}\psi)_y \rangle e^{ip\cdot y} + \delta Z_4 - p^2 \delta Z_6 = 0 \tag{3.34}$$

$$\Rightarrow i\lambda \int \frac{d^4k}{(2\pi)^4} \text{Tr} \left\{ \sum_M D_M^{-1} \frac{1}{i} \frac{-\slashed{k} + m_M}{k^2 + m_M^2 - i\epsilon} \right.$$

$$\left. \times \sum_N D_N^{-1} \frac{1}{i} \frac{-\slashed{k} - \slashed{p} + m_N}{(k+p)^2 + m_N^2 - i\epsilon} \right\} + \delta Z_4 - p^2 \delta Z_6 = 0, \tag{3.35}$$

where $p^2 = -m_\sigma^2$.

For $\delta Z_i$ that depend on curved spacetime nature, we match the renormalized result to that from the adiabatic subtraction. In order to fix $\delta Z_7$, we impose the number density $\langle in|(\bar{\psi}\psi)_{r,x}|in\rangle$ to be the density defined by the adiabatic prescription (See, e.g., [6, 65, 66, 86, 100, 185, 186]):

$$n_\psi \equiv \langle in|\bar{\psi}\psi(x)|in\rangle + \sum_{n=1} \langle in|\bar{\psi}_n\psi(x)_n|in\rangle + \delta Z_5 + \delta Z_7 R(x) \tag{3.36}$$

$$= \langle in|\bar{\psi}\psi(x)|in\rangle - \langle WKB, vac, t_x|\bar{\psi}\psi(x)|WKB, vac, t_x\rangle, \tag{3.37}$$

where $|WKB, vac, t_x\rangle$ is the WKB vacuum defined at $t_x$ by the adiabatic prescription. The diagram of interest is diagram (a) of Fig. 3.1, and the divergent part of $\delta Z_7$ determined this way is linear in the fermion mass.

In order to determine $\delta Z_8$, we repeat the consideration analogous to Eq. (3.32) on a background field $\sigma(x) = c$, where $c$ is an infinitesimal constant. Since a constant $\sigma$ shift is equivalent to a shift of the fermion mass, we want to choose $\delta Z_8$ to get

$$
\begin{aligned}
\lambda \partial_m n_\psi(x) &= -i\lambda \int_{CTP} (dy) \sum_{N,M} \langle in|P\{\bar{\psi}_M(x)\psi_N(x)\bar{\psi}_N(y)\psi_M(y)\}|in\rangle_{conn} \\
&\quad +\delta Z_4 + \delta Z_8 R(x),
\end{aligned}
\tag{3.38}
$$

where the subscript CTP denotes closed-time-path, and P is the path-ordering operator for a "in-in" exception value. (For example, see Refs. [195, 196]). Note that diagram of interest corresponds to (e) of Fig. 3.1. As we will see later, this renormalization condition plays a crucial role in determining the isocurvature correlator. The solution for all the $\delta Z_i$ can be expressed in terms of Feynman parameter integrals. However, such explicit expressions are not relevant to determine the isocurvature correlation function. In contrast the left hand side of Eq. (3.38) is important.

To summarize, we have given a prescription to regularize and renormalize the composite operator $\bar{\psi}\psi$. The renormalization conditions ensure that $\langle in|(\bar{\psi}\psi)_{r,x}|in\rangle$ agrees with that defined by the adiabatic prescription in curved spacetime, and they also ensure that a constant shift in $\sigma$ is equivalent to a constant shift in the fermion mass. Note that because the gravitational production of fermions are still in flux when $m_\psi < H$ , we evaluate the number density $n_\psi$ later than $t_*$, where $H(t_*) \sim m_\psi$, as far as the renormalization conditions are concerned.

## 3.4   Scenario Constraints on Scalar Field $\sigma$

In this section, we explain the constraints on the Yukawa coupling $\lambda$ that comes from requiring $\sigma$ to behave as an unscreened long range force carrier whose on-shell particle states do not significantly participate in $\psi$ production.

We will find that $\langle\sigma\sigma\rangle|_{t_*}$ power spectrum relevant for the isocurvature perturbations is not suppressed if $m_\sigma/H(t_*) < 1$ where $t_*$ is the time at which $H(t_*) = m_\psi$ (i.e. $t_*$ is the time at which the fermion + anti-fermion number freezes [185]). This implies $m_\sigma < m_\psi$ is the relevant parameter region for the scenario of this work. Furthermore, in order to prevent any large isocurvature perturbations and relic abundance of $\sigma$, we assume that $\langle\sigma\rangle = 0$ and the $\sigma$ particles decay before $\sigma$ becomes an important fluid component of the evolution of the universe (e.g. before matter-radiation equality). Note however that this restriction is a matter of simplicity. There exist parameter regions in $(m_\sigma, \lambda)$ such that $\sigma$ survives as a long-lived weakly interacting particle (i.e. a dark matter). However, in such cases, the constraints from the relic abundance and the isocurvature of $\sigma$ restrict the $\sigma$ mass to be very small, e.g., $m_\sigma \lesssim 10^{-6}\text{eV}$ for $H_{inf} \sim 10^{13}\text{GeV}$. (See, e.g, [148, 149, 151, 197] for the parametric bounds for the QCD axion produced by inflation.) In principle, it is possible to build a model that has such small $m_\sigma$ with help of some underlying symmetry, such as a shift symmetry.

Although we assume that $m_\sigma < m_\psi$, $\sigma$ would generally acquire a plasma mass correction through interactions with an ensemble of $\psi$ particles. Thus we consider the effect of the produced $\psi$ on the $\sigma$ correlator and show that the effect is negligible. We expect the fermions do not affect scalar modes before horizon exit because the mass correction by the fermion is still small compared to the Hubble friction during inflation. After the scalar mode exits the horizon, the fermions exert a tiny computable

drag on $\sigma$. The equation of motion of $\sigma$ from the action $(3.6)$[11] is written as

$$
\begin{aligned}
0 &= \left\langle \left[ (\Box - m_\sigma^2)\sigma_x - \lambda\bar\psi\psi_x \right.\right. \\
&\quad \left.\left. + \delta Z_0 + \delta Z_R R_x + \delta Z_\sigma \Box\sigma_x - \delta m_\sigma^2 \sigma_x + \delta Z_\xi R_x \sigma_x \right] [\cdots] \right\rangle \qquad (3.39) \\
&= (\Box_x - m_\sigma^2)\left\langle \sigma_x [\cdots] \right\rangle \\
&\quad + i\lambda^2 \int^x (dz)\langle [\bar\psi\psi_x, \bar\psi\psi_z]\rangle \left\langle \sigma_z [\cdots] \right\rangle + \left( \delta Z_\sigma \Box_x - \delta m_\sigma^2 + \delta Z_\xi R_x \right) \left\langle \sigma_x [\cdots] \right\rangle \\
&\quad + \left( \delta Z_0 + \delta Z_R R_x - \lambda\langle\bar\psi\psi_x\rangle \right) \left\langle [\cdots] \right\rangle + O(\lambda^3, y), \qquad (3.40)
\end{aligned}
$$

where $[\cdots]$ denotes any quantum operators in the correlation function. We choose the counter term $\delta Z_0$ and $\delta Z_R$ such that the tadpole $\langle\sigma\rangle$ vanishes, i.e., $(\delta Z_0 + \delta Z_R R - \lambda\langle\bar\psi\psi\rangle) = 0$, where the PV regulator is assumed. Moreover, when $\sigma$ varies very slowly outside the horizon, we factor $\langle\sigma_z [\cdots]\rangle$ out of the integral in Eq. (3.40), and we renormalize the integral using the counter terms $(\delta Z_\sigma \Box_x - \delta m_\sigma^2 + \delta Z_\xi R_x)\langle\sigma_x [\cdots]\rangle$ such that the result is consistent with that obtained by the adiabatic subtraction[12]:

$$
i\lambda^2 \int^x (dz)\langle[\bar\psi\psi_x, \bar\psi\psi_z]\rangle + \left( -\delta m_\sigma^2 + \delta Z_\xi R_x \right) = -\lambda^2 \left( \frac{\partial n_\psi}{\partial m_\psi} \right), \qquad (3.41)
$$

where $n_\psi$ is the renormalized fermion number density defined by Eq. (3.37), and we have used Eq. (3.38) in the derivation. Therefore, we find the effective mass of $\sigma$ when it slowly varies (i.e., $k/a \ll H$ and $m_\sigma \ll H$)

$$
m_\sigma^{eff} = m_\sigma^2 + \Delta m_\sigma^2(t) \approx m_\sigma^2 + \lambda^2 \frac{\partial n_\psi(t)}{\partial m_\psi}. \qquad (3.42)
$$

---

[11]The counter-terms appearing in the action includes

$$
S_{c.t.} \ni \int (dx) \left[ -\frac{1}{2}\delta Z_\sigma (\partial\sigma)^2 - \frac{1}{2}\delta m_\sigma^2 \sigma^2 + \delta Z_0 \sigma + \delta Z_R R\sigma + \delta Z_\xi R\sigma^2 \right].
$$

Note that the the linear $\sigma$ terms exist in the action because the action does not preserve the $Z_2$ symmetry due to the Yukawa coupling.

[12]In other words, we identify $-\delta m_\sigma^2$ and $\delta Z_\xi$ with $\delta Z_4$ and $\delta Z_8$ in Eq. (3.38), and $\delta Z_\sigma\Box$ is neglected since $\sigma$ is slowly varying.

Because we estimate $n_\psi \lesssim O(0.1)\,(m_\psi H)^{3/2}$ when $m_\psi \lesssim H$,[13] based on dimensional analysis, we expect that the mass correction by the $\psi$ loop is

$$\Delta m_\sigma^2(t) \approx \lambda^2 \frac{\partial n_\psi(t)}{\partial m_\psi} \sim \begin{cases} O(0.1 \text{ or less})\lambda^2 m_\psi^{1/2} H^{3/2} & \text{for } m_\psi < H(t) \\ O(0.1)\lambda^2 m_\psi^2 & \text{for } m_\psi > H(t) \end{cases}. \qquad (3.43)$$

Therefore, in general, before the fermion production ends $m_\psi < H$, this scalar mass correction $\Delta m_\sigma^2$ does not ruin the stability of our scenario $m_\sigma^2 + \Delta m_\sigma^2(t) < m_\psi^2 < H^2(t)$ as long as $m_\sigma^2 < m_\psi^2$.

Next, we ask the question of which parametric region would be consistent with the simplifying assumption that $\psi$ particles are primarily produced gravitationally and not by $\sigma$. To this end, we first consider the annihilation $\sigma\sigma \to \bar{\psi}\psi$. The annihilation is the most significant at the end of inflation because $\psi$ particles produced from $\sigma$ before the end of inflation are diluted, and $\sigma\sigma \to \bar{\psi}\psi$ after the end of inflation is also limited because the allowed kinematic phase space is redshifted. Thus we compare the number density of the produced $\psi$ from $\sigma$ at the end of inflation, $n_{\sigma\to\psi}$ with that of gravitationally produced $\psi$, $n_\psi(t_*) \sim m_\psi^3$, and we require their ratio to be small:

$$\left(\frac{a_e}{a(t_*)}\right)^3 \frac{n_{\sigma\to\psi}(t_e)}{n_\psi(t_*)} \quad \sim \quad \left(\frac{a_e}{a(t_*)}\right)^3 \frac{n_\sigma \Gamma(\sigma\sigma \to \psi\psi)\Delta t|_{t_e}}{n_\psi(t_*)} \qquad (3.44)$$

$$\sim \quad \left(\frac{H(t_*)}{H_e}\right)^2 \frac{H_e^3 \cdot \frac{\lambda^4}{16\pi^2} H_e \cdot \frac{1}{H_e}}{H^3(t_*)} \sim \frac{\lambda^4}{16\pi^2} \frac{H_e}{m_\psi} \lesssim 1, \quad (3.45)$$

where the subscript $e$ means a variable is evaluated at the end of inflation $t_e$.

Even though $m_\sigma < m_\psi$, the decay production of $\psi$ through $\sigma \to \bar{\psi}\psi$ may still be possible if $\sigma$ is sufficiently off shell due to its interactions with finite density of $\psi$ in

---

[13]Note that the adiabatic prescription to determine the number density $n_\psi$ does not apply for modes $m_\psi < k/a < \sqrt{m_\psi H}$ when $m_\psi < H$ because vacuum varies non-adiabatically, i.e., the adiabaticity parameter $\epsilon_k \equiv \frac{m_\psi k_p H}{\omega_k} \gtrsim 1$, where $k_p = k/a$ and $\omega_k = \sqrt{k_p^2 + m_\psi^2}$. See Appendix 3.10 for detail. However, we can estimate the upper bound of the number density as

$$n_\psi(t) = \int \frac{d^3 k_p}{(2\pi)^3} |\beta_k|^2 \lesssim \int^{\sqrt{m_\psi H}} \frac{d^3 k_p}{(2\pi)^3} \frac{1}{2} \sim O(0.1)\,(m_\psi H)^{3/2} \quad \text{for } t < t_*.$$

the subhorizon region (the subhorizon physics here is different from the superhorizon physics considered in Eq. (3.42)). To turn off this channel, we require that the $\sigma$ mass corrections from the fermion number density at the time of end of inflation be small. This requires

$$\lambda^\kappa \frac{H_e/(2\pi)}{m_\psi} \lesssim 1 \tag{3.46}$$

where $\kappa \gtrsim O(1)$. To see how $\kappa \gtrsim O(1)$ can come about, consider the following estimate of subhorizon thermal effect. The maximum effective number density of fermions at the end of inflation is

$$n_\psi(t_e) \lesssim 4m_\psi \left(\frac{H_e}{2\pi}\right)^2. \tag{3.47}$$

The energy density associated with these fermions is

$$\Delta V \sim n_\psi(t_e) \sqrt{\left(\frac{H_e}{2\pi}\right)^2 + \lambda^2\sigma^2} \tag{3.48}$$

where we neglected $m_\sigma \ll H_e/(2\pi)$. This leads to an effective $m_\sigma$ correction of

$$\Delta m_\sigma^2 \sim n_\psi(t_e) \frac{\lambda^2}{H_e/(2\pi)} \lesssim 4\lambda^2 m_\psi \left(\frac{H_e}{2\pi}\right). \tag{3.49}$$

Kinematically blocking the $\sigma$ decay into $\psi$, we find

$$4\lambda^2 \left(\frac{H_e}{2\pi}\right) < m_\psi \tag{3.50}$$

which corresponds to $\kappa = 2$. Note that this condition is more restrictive than Eq. (3.45).

In sum, requiring $\sigma$ to behave as an unscreened long range force carrier whose on-shell particle states do not significantly participate in $\psi$ production gives a constraint on $\lambda$. The strongest condition is given by Eq. (3.46) with $\kappa \gtrsim O(1)$.

## 3.5 Isocurvature two-point function

In this section, we evaluate the two-point function of the renormalized isocurvature operator $\delta_S$, given by Eq. (3.19). The average number density was computed in [185], the result is summarized in Appendix 3.10. We only need to evaluate $\langle(\bar{\psi}\psi)_{x,r}(\bar{\psi}\psi)_{y,r}\rangle_c$. Since we want to use the quantum computation to set the initial condition for the subsequent classical fluid evolution, we will choose the time of the evaluation such that both the quantum and the classical fluid descriptions apply. We take $x^0 = y^0 = \eta_f$ at time after the particle production ends, since the fluid description cannot describe the particle production process. We will take the separation $|\vec{x} - \vec{y}|$ to be large enough such that the intersection of their past light-cone $I^-(x) \cap I^-(y)$ lives deep within the inflationary era. This ensures that the contributions from late-time short distance physics (e.g. reheating, phase transition) are minimized. The relevant diagrams for $\langle(\bar{\psi}\psi)_{x,r}(\bar{\psi}\psi)_{y,r}\rangle_c$ are given in Fig. (3.2). The crossed dot represent $(\bar{\psi}\psi)_{x,r}$ insertion, the solid dot represent the Yukawa interaction vertex, the dashed line represent the scalar $\sigma$ propagator, and the solid line represent the fermion propagator.

### Leading order result

We first consider the leading order diagram (a) in Fig (3.2). The diagram is explicitly written as

$$\langle\bar{\psi}\psi_x\bar{\psi}\psi_y\rangle_{(a)} = -\mathrm{Tr}\left[\langle\psi_x\bar{\psi}_y\rangle\langle\psi_y\bar{\psi}_x\rangle\right] = \sum_{i,j}\bar{V}_{i,x}U_{j,x}\bar{U}_{j,y}V_{i,y} \qquad (3.51)$$
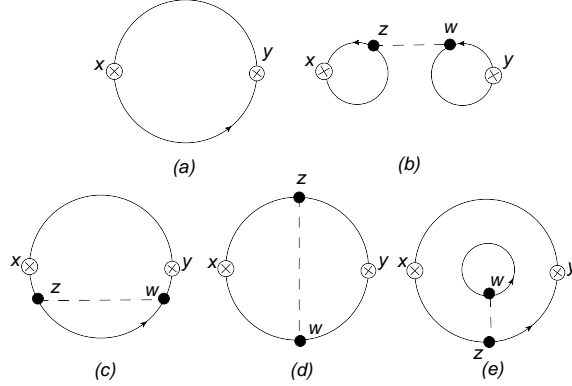
Figure 3.2: The leading order and the next leading order contribution to $\langle \bar\psi\psi_x \bar\psi\psi_y \rangle$ , the cross-dot vertices corresponds to $\bar\psi\psi$ insertion. By comparing the large $r$ ($r \equiv |\vec{x} - \vec{y}|$) behavior of the equal-time correlator of the fermion and the scalar field, we want to show that diagram (b) dominates in the limit $r \to \infty$.

Using a contour integration technique, we can evaluate the mode-sum analytically. The details are in given Appendix 3.11. The result[14] is

$$\langle \bar\psi\psi_x \bar\psi\psi_y \rangle_{LO} = \begin{cases} \frac{1}{\pi^4 a_x^6 |\vec{x}-\vec{y}|^6} \left( 1 + O\left[ \left( \frac{m_\psi}{H_{inf}} \right)^2 \right] \right) & (m_\psi \ll H_{inf}) \\ \frac{1}{\pi^4 a_x^6 |\vec{x}-\vec{y}|^6} (4\pi) \left( \frac{m_\psi}{H_{inf}} \right)^3 \exp(-2\pi \frac{m_\psi}{H_{inf}}) & (m_\psi \gg H_{inf}) \end{cases} \tag{3.52}$$

where $H_{inf}$ is the expansion rate during inflation. We can understand this result by backtracking the two points $x, y$ to the time when they were deep inside the horizon, and see what happened as they grow apart.

In the heavy mass case ($m_\psi \gg H_{inf}$), the Compton radius $m_\psi^{-1}$ is smaller than the horizon radius $H_{inf}^{-1}$. The physical separation $r_{\mathrm{phys}}$ will first grow to the Compton wavelength, and trigger the exponential suppression factor $\exp(-2m_\psi r_{phys})$ in the correlator.

$$\langle \bar\psi\psi_x \bar\psi\psi_y \rangle_{flat, m_\psi r_{phys} > 1} \sim \frac{m_\psi^3}{4\pi^3 r_{phys}^3} \exp(-2m_\psi r_{phys}) \tag{3.53}$$

---

[14]Note that we do not consider the the heavy mass case, $m_\psi \gg H_{inf}$ where $H_{inf}$ is the expansion rate during inflation, for the isocurvature because the estimation of the particle production depends on how the inflation ends as described in Section 3.2. However, we provide the leading order of the two-point function to develop better intuition for the behavior of super horizon modes of $\psi$.

As the physical separation $r_{\text{phys}}$ grows further to exceed the horizon radius $H_{inf}^{-1}$, the correlator would freeze and start decreasing as $(a_r/a_\eta)^6$, where $a_r = 1/(H_{inf}r)$ denote the scale factor at the horizon crossing. Substituting $a_r = \frac{1}{H_{inf}r}$ and $r_{phys} = H_{inf}^{-1}$, we recover the heavy mass formula:

$$\left(\frac{a_r}{a_\eta}\right)^6 \frac{m_\psi^3}{4\pi^3 r_{phys}^3} \exp(-2m_\psi r_{phys}) \sim \frac{1}{a_x^6 r^6} \left(\frac{m_\psi}{H_{inf}}\right)^3 \exp(-2\frac{m_\psi}{H_{inf}}). \tag{3.54}$$

In the light mass case ($m_\psi \ll H_{inf}$), the physical distance will cross the horizon radius first, without the exponential suppression of $\exp(-2m_\psi r_{phys})$. From the flat space UV limit result $\frac{1}{r_{phys}^6}$,

$$\langle \bar{\psi}\psi_x \bar{\psi}\psi_y \rangle_{flat, mr_{phys}<1} \sim \frac{1}{r_{phys}^6} \tag{3.55}$$

we use $a_r = \frac{1}{H_{inf}r}$ and $r_{phys} = H_{inf}^{-1}$ to obtain

$$\left(\frac{a_r}{a_\eta}\right)^6 \frac{1}{r_{phys}^6} \sim \frac{1}{a_x^6 r^6} \tag{3.56}$$

Thus we recover the light mass result.

Unfortunately, the fractional relic density fluctuation at CMB scale[15] is too small

$$\frac{\langle \delta\rho_x \delta\rho_y \rangle}{\langle \bar{\rho}_\psi \rangle^2} \sim \frac{m_\psi^2/(\pi^4 a^6 r_{CMB}^6)}{m_\psi^2 m_\psi^6 (a_*^6/a^6)} \sim \left(\frac{1}{a_* m_\psi r_{CMB}}\right)^6. \tag{3.57}$$

where $r_{CMB}$ is the comoving distance for typical CMB observation scale and the subscript $*$ denotes the time when fermion production ends. Let $a_{CMB}$ denotes the scale factor when CMB scale exits the horizon then we have

$$r_{CMB}^{-1} \sim a_{CMB} H_{inf} \tag{3.58}$$

Assuming the fermion production ends during reheating when $m_\psi = H(t_*)$, and $H \propto a^{-\alpha}$ during reheating, then we have

$$\frac{a_e H_{inf}}{a_* m_\psi} \sim \frac{a_e H_e}{a_* H_*} \sim \left(\frac{a_e}{a_*}\right)^{1-\alpha} \sim \left(\frac{H_e}{H_*}\right)^{1-\frac{1}{\alpha}} \tag{3.59}$$

---

[15]Since $\langle \delta\rho\delta\rho \rangle$ is frozen as long as the two points are outside of horizon, we can extrapolate this large spatial separation result obtained at the end of inflation to the recombination time.

Assuming that inflation ends 50 efolds after the CMB scale exits horizon and a MD-like reheating, i.e., $\alpha = 3/2$, then we have

$$\frac{\langle \delta\rho_x \delta\rho_y \rangle}{\langle \bar{\rho}_\psi \rangle^2} \sim \left( \frac{a_{CMB} H_{inf}}{a_* m_\psi} \right)^6 \sim \left( \frac{a_{CMB}}{a_e} \frac{a_e H_{inf}}{a_* m_\psi} \right)^6 \sim e^{-300} \left( \frac{H_e}{m_\psi} \right)^2 \qquad (3.60)$$

Using the fermion relic abundance formula (for $T_{RH} = 10^9 \text{GeV}$ and $g_* = 100$ case) $\omega_\psi \sim (m_\psi/10^{10}\text{GeV})^2$, we obtain

$$\frac{\langle \delta\rho_x \delta\rho_y \rangle}{\rho_{tot}^2} \sim \omega_\psi^2 \frac{\langle \delta\rho_x \delta\rho_y \rangle}{\langle \bar{\rho}_\psi \rangle^2} \sim e^{-300} \left( \frac{H_e}{10^{10} GeV} \right)^2 \qquad (3.61)$$

We thus find that generically the pure fermion isocurvature is very small on scales relevant for the CMB.

## Next leading order result

We consider the diagrams (b)-(e) in Fig. 3.2, which contain the effects of the Yukawa interaction to the fermion production. We can perturbatively compute the diagrams using the "in-in" formalism (e.g. see Refs. [93, 198] and references therein).

Firstly, we estimate which diagram gives the largest contribution when $x$ and $y$ have large spatial separations. From the fact that equal-time correlator $\langle \sigma_x \sigma_y \rangle$ scales as $r^{2\nu-3}$ where $\nu^2 = 9/4 - m_\sigma^2/H^2$ from Eq. (3.105) and $\langle \psi_x \bar{\psi}_y \rangle$ scales as $r^{-3}$, we expect that diagrams that have fewer fermion lines stretched between $x$ and $y$ decreases slower as $r \to \infty$. Thus, we conclude diagram (b) gives the dominant contribution to the two-point function.

For diagram (b), we expand it using commutators

$$
\begin{aligned}
I_b(x,y) &= \langle (\bar{\psi}\psi)_{x,r}(\bar{\psi}\psi)_{y,r}\rangle_{c,diag(b)} \tag{3.62}\\
&= 4(i\lambda)^2 \int^x (dz)\int^y (dw)\langle \bar{\psi}\psi_{[x}\bar{\psi}\psi_{z]}\rangle \langle \bar{\psi}\psi_{[y}\bar{\psi}\psi_{w]}\rangle \langle \sigma_{\{z}\sigma_{w\}}\rangle \\
&\quad + 4(i\lambda)^2 \int^x (dz)\int^y (dw)\langle \bar{\psi}\psi_{\{x}\bar{\psi}\psi_{z\}}\rangle \langle \bar{\psi}\psi_{[y}\bar{\psi}\psi_{w]}\rangle \langle \sigma_{[w}\sigma_{z]}\rangle \Theta(w^0 - z^0) \\
&\quad + 4(i\lambda)^2 \int^x (dz)\int^y (dw)\,(\, x \leftrightarrow z \text{ and } w \leftrightarrow z\,) \tag{3.63}\\
&\approx (i\lambda)^2 \int^x (dz)\int^y (dw)\langle [\bar{\psi}\psi_x,\bar{\psi}\psi_z]\rangle \langle [\bar{\psi}\psi_y,\bar{\psi}\psi_w]\rangle \langle \sigma_{\{z}\sigma_{w\}}\rangle \tag{3.64}
\end{aligned}
$$

where $(dz) = \sqrt{-\det{(g_{\mu\nu})}}d^4z$, $[\cdots]$ means anti-symmetrization and $\{\cdots\}$ means symmetrization, and we have implicitly assumed the PV regulator. From the scalar and fermion mode functions in de Sitter spacetime, we know $\langle [\sigma_{x_1},\sigma_{x_2}]\rangle$ is suppressed by $a^{-2\nu}$ relative to $\langle \{\sigma_{x_1},\sigma_{x_2}\}\rangle$, whereas $\langle [\bar{\psi}\psi_{x_1},\bar{\psi}\psi_{x_2}]\rangle$ is suppressed by $a^{-1}$ relative to $\langle \{\bar{\psi}\psi_{x_1},\bar{\psi}\psi_{x_2}\}\rangle$. The last line is obtained by keeping only the dominant contribution.

Since the fermion particle production ends at $t_*$ and the previously produced particles have been diluted away, we expect the $z$ and $w$ integrals to peak around the time $t_*$. For late time and large spatial separations, the scalar correlator $\langle \sigma_{\{z}\sigma_{w\}}\rangle$ is slowly varying with respect to changes in $z$ and $w$. Thus we may approximately take $\langle \sigma_{\{z}\sigma_{w\}}\rangle = \langle \sigma_{\{z_0}\sigma_{w_0\}}\rangle$, where $z_0 = (t_*, \vec{x})$ and $w_0 = (t_*, \vec{y})$, and factor it outside of the $z, w$ integral:

$$
I_b(x,y) \approx (i\lambda)^2 \langle \sigma_{\{z_0}\sigma_{w_0\}}\rangle [\int^x (dz)\langle [\bar{\psi}\psi_x,\bar{\psi}\psi_z]\rangle][\int^y (dw)\langle [\bar{\psi}\psi_y,\bar{\psi}\psi_w]\rangle] \tag{3.65}
$$

The remaining fermion integral $\int^x (dz)\langle [\bar{\psi}\psi_x,\bar{\psi}\psi_z]\rangle$ is quadratically divergent. The counter-terms $\delta Z_4 \sigma + \delta Z_8 R\sigma$ in $(\bar{\psi}\psi)_r$ is in place to cancel such divergences. Furthermore, our choice of the renormalization conditions given in Section 3.3 ensures that a constant shift in $\sigma$ is equivalent to a shift of the fermion mass (see Eq. (3.38)). An explicit computation of the fermion loop integral using the adiabatic subtraction is

given in Appendix 3.12. Thus we have

$$\langle (\delta_S)_{r,x}(\delta_S)_{r,y} \rangle_{NLO} \approx \omega_\psi^2 \lambda^2 [\partial_m \ln n_\psi|_x][\partial_m \ln n_\psi|_y]\langle \sigma_{\{(\vec{x},t_*)}\sigma_{(\vec{y},t_*)\}} \rangle \qquad (3.66)$$

where $t_*$ is the time when fermion production ends (i.e. $m_\psi \sim H(t_*)$) and $\partial_m$ denotes the derivative with respect to $m_\psi$. Note that $\langle (\delta_S)_{r,x}(\delta_S)_{r,y} \rangle_{NLO}$ freezes for $t > t_*$ since $\partial_m n_\psi$ and $n_\psi$ behaves as $a^{-3}$ after the fermion production ends. We will discuss the numerical implications of this result below.

To summarize, we computed the isocurvature correlation function to the next leading order, as in Eq. (3.66). Intuitively, the light scalar's quantum fluctuation modulate the fermion's mass, which affect the fermion relic abundance. In the same line of thought, we may extrapolate this result to estimate higher order corrections

$$\langle (\delta_S)_{r,x}(\delta_S)_{r,y} \rangle_{full} \approx \omega_\psi^2 \frac{\langle n_\psi \left( m_\psi + \lambda\sigma(\vec{x},t_*) \right) n_\psi \left( m_\psi + \lambda\sigma(\vec{y},t_*) \right) \rangle_\sigma}{n_\psi^2} \qquad (3.67)$$

where we have treated $n_\psi$ to be a function of its mass and the expectation value is taken with respect of the $\sigma$ field.

## Isocurvature Power Spectrum

In the long wavelength limit, which corresponds to the low multipoles in the angular CMB anisotropy, the temperature fluctuations dominantly come from the Sach-Wolfe term [127], which is expressed as

$$\frac{\Delta T}{T} = -\frac{1}{5}\zeta - \frac{2}{5}\delta_S. \qquad (3.68)$$

Then the power spectrum of the temperature fluctuations

$$\Delta^2_{\frac{\Delta T}{T}}(k) \equiv \frac{k^3}{2\pi^2} \int d^3x \left\langle \frac{\Delta T}{T}(t,\vec{x}) \frac{\Delta T}{T}(t,\vec{0}) \right\rangle e^{-i\vec{k}\cdot\vec{x}} \tag{3.69}$$

$$= \frac{1}{25}\Delta^2_\zeta(k) + \frac{4}{25}\Delta^2_{\delta_S}(k), \tag{3.70}$$

$$\Delta^2_\zeta(k) \equiv \frac{k^3}{2\pi^2} \int d^3x \left\langle \zeta(t,\vec{x})\zeta(t,\vec{0}) \right\rangle e^{-i\vec{k}\cdot\vec{x}}, \tag{3.71}$$

$$\Delta^2_{\delta_S}(k) \equiv \frac{k^3}{2\pi^2} \int d^3x \left\langle \delta_S(t,\vec{x})\delta_S(t,\vec{0}) \right\rangle e^{-i\vec{k}\cdot\vec{x}}, \tag{3.72}$$

where the cross-correlation contribution $\langle \zeta \delta_S \rangle$ has been neglected because of the reason explained in Section 3.7. When the leading term approximation (3.66) is valid, Eq. (3.66) yields the isocurvature power spectrum

$$\Delta^2_{\delta_S}(t,k) = \omega^2_\psi(t)\lambda^2 \left( \frac{\partial_m n_\psi(m_\psi)}{n_\psi} \right)^2 \Delta^2_\sigma(t_*,k) + O(\lambda^4), \tag{3.73}$$

which includes the extra factor $\omega^2_\psi$ due to the thermal relics. Furthermore, when the mass of scalar field $\sigma$ is sufficiently light such that $\sigma$ does not start its coherent oscillation until the fermion particle production ends, i.e., $m_\sigma < H(t_*) < H_{inf}$, the power spectrum for $\sigma$ is

$$\Delta^2_\sigma(t_*,k) \approx \frac{H^2(t_k)}{4\pi^2} \tag{3.74}$$

where $t_k$ is the time when the scale $k$ exits horizon. Note that we have already shown that the correction of $m_\sigma$ due to the fermion loop is negligible in Section 3.4. Therefore, the isocurvature power spectrum becomes

$$\Delta^2_{\delta_S}(k) \approx \omega^2_\psi \lambda^2 \left( \frac{\partial_m n_\psi(m_\psi)}{n_\psi} \right)^2 \frac{H^2(t_k)}{4\pi^2}. \tag{3.75}$$

The currently known parametric bounds for this isocurvature power spectrum is presented in Section 3.6.

## 3.6 Result and Discussion

**Parameter bounds**

In this subsection, we present the allowed parameter region in the fermion isocurvature model from the observational constraints using the dark matter relic abundance and the CDM isocurvature power-spectrum. In this scenario, there are 5 independent parameters: $m_\psi$, $H_{inf}$, $\lambda$, $T_{RH}$ and $m_\sigma$, where $H_{inf}$ is the Hubble scale during inflation and $T_{RH}$ is the reheating temperature. We assume $H_{inf}$ and $T_{RH}$ are free parameters governed entirely by the inflaton and the reheating sector. As discussed in Section 3.2, as long as $m_\sigma \ll m_\psi$, the exact value of the scalar mass $m_\sigma$ is numerically unimportant in this model. Therefore, we are basically left with two parameter, namely $\lambda$ and $m_\psi$.[16]

For the light fermion, $m_\psi < H_{inf}$, the fermion particle number freezes when $H(t_*) \sim m_\psi$ as reviewed in Appendix 3.10. In particular, the Yukawa coupling works effectively as a mass shift in our scenario $m_{eff} = |m_\psi + \lambda\sigma(t_*)|$. The fermion relic abundance (3.128) becomes

$$\Omega_\psi h^2 \quad \sim \quad 3r\left(\frac{m_{eff}}{10^{11}\,\mathrm{GeV}}\right)^2\left(\frac{T_{RH}}{10^9\,\mathrm{GeV}}\right), \tag{3.76}$$

where the extra factor $r$ comes from the difference in the effective masses at $t_*$ and later time, at which the energy density of $\psi$ is not negligible, such as the MD era. For example, if $\sigma$ is treated as a Gaussian random variable with $\sqrt{\langle\sigma^2\rangle} \sim H_{inf}/2\pi$,

---

[16]Note that we implicitly assume that if $m_\psi$ and $T_{RH}$ are such that the dark matter relic abundance is not saturated by the $\psi$ energy density, the other CDM sector in Eq. (3.6) is adjusted to provide the rest of the dark matter. Note that when the $\psi$ dark matter abundance is small, no large tuning is needed to make this occur since the well known WIMP miracle can saturate the dark matter abundance.

we can approximate $r \approx m_\psi / \langle m_{eff} \rangle$ and write

$$\Omega_\psi h^2 \sim \begin{cases} \left(\frac{m_\psi}{10^{11}\,\text{GeV}}\right)^2 \left(\frac{T_{RH}}{10^9\,\text{GeV}}\right) & \text{if } m_\psi > \lambda H_{\text{inf}}/2\pi \\ \frac{2\pi m_\psi}{\lambda H_{inf}} \left(\frac{\lambda H_{inf}}{10^{11}\,\text{GeV}}\right)^2 \left(\frac{T_{RH}}{10^9\,\text{GeV}}\right) & \text{if } m_\psi < \lambda H_{\text{inf}}/2\pi \end{cases}, \qquad (3.77)$$

where $O(1)$ factors are neglected.

Furthermore, from the result (3.75) in Sec. 3.5, the fractional isocurvature amplitude [199] becomes

$$\alpha_S \equiv \frac{\Delta_{\delta_S}^2}{\Delta_\zeta^2 + \Delta_{\delta_S}^2} \sim \frac{\lambda^2}{2} \left(\frac{m_\psi}{10^4\,\text{GeV}}\right)^2 \left(\frac{H}{10^{13}\,\text{GeV}}\right)^2 \left(\frac{T_{RH}}{10^9\,\text{GeV}}\right)^2, \qquad (3.78)$$

where we have used

$$\frac{\partial_m n_\psi}{n_\psi} \sim \begin{cases} m_\psi^{-1} & \text{for } m_\psi > \lambda H_{inf}/2\pi \\ 2\pi \lambda^{-1} H_{inf}^{-1} & \text{for } m_\psi < \lambda H_{inf}/2\pi \end{cases}, \qquad (3.79)$$

because the number density $n_\psi$ at the time $t_*$ is determined by only one dimensionful scale $m_{eff} \sim H(t_*)$. The current observational bound [108, 109, 199–202] of the isocurvature for the uncorrelated case, i.e. $\langle \zeta \delta_S \rangle = 0$, is $\alpha_S < 0.016\,(95\%\,\text{CL})$ from the Planck+WP9 combined data, which yields the constraints on the parameters $\lambda$ and $m_\psi$. Combining the above consideration, we have the parameter plot shown in Fig. 3.3. We emphasize that the parameter region beyond the (left diagonal) bound from the $\sigma$ annihilation, Eq. 3.50, is not necessarily excluded. Due to the uncertainty of the $\sigma$ annihilation effect, we provide it as a conservative bound of this model.

The case that $m_\psi < \lambda H_{inf}/(2\pi)$ (which we will refer to as large mass correction regime) is potentially the most interesting case because the fermion number density $n_\psi$ depends on $|m_\psi + \lambda \sigma|$, not $m_\psi + \lambda \sigma$ as the sign of the fermion mass is irrelevant for particle production[17]. This may lead to interesting features such as large non-Gaussianities when the effective mass varies from negative to positive depending

---

[17]The sign of the fermion mass changes under a chiral transformation.
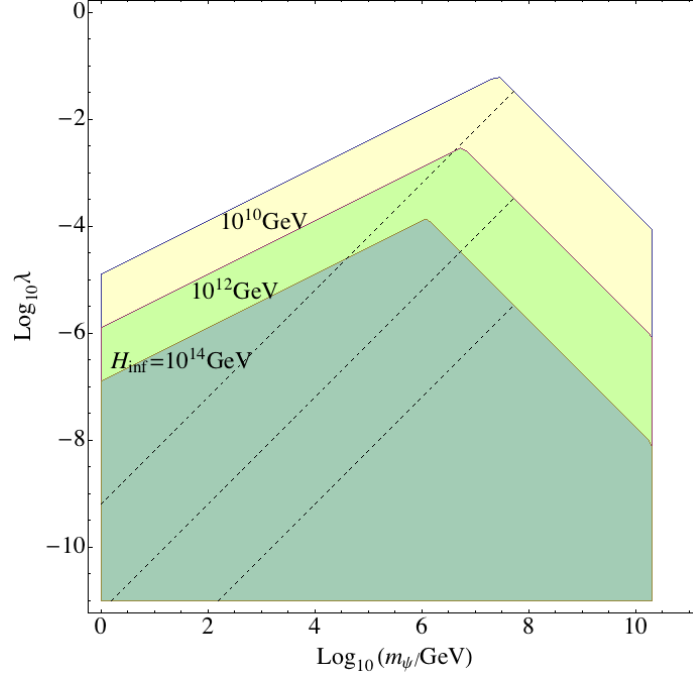
Figure 3.3: Bounds on the fermion mass and Yukawa coupling for various inflationary Hubble scales. The vertical bound corresponds to the total dark matter relic density constraint, the right diagonal and the left diagonal bounds correspond to the constraints from the CDM isocurvature and the $\sigma$ annihilation using Eq. (3.50), respectively. The splitting dashed lines in each region separates the small mass and large mass correction regime. In this plot, we set $T_{RH} = 10^9 GeV$.

on the local Hubble patches at $t_*$. However, this parametric region has couple of problems: 1) the perturbative calculation of $n_\psi$ may be unsuitable since we are not resuming the large mass corrections; 2) Eq. (3.46) may not be satisfied. Hence, for the rest of this section, we primarily focus on the case that $m_\psi > \lambda H_{inf}/(2\pi)$, which we will refer to as the small mass correction regime.

## Non-Gaussianities

In this subsection, we compute the bi-spectrum $B_S(\vec{p}_1, \vec{p}_2, \vec{p}_3)$ defined by

$$(2\pi)^3 \delta^{(3)}(\sum_i \vec{p}_i) B_S(\vec{p}_1, \vec{p}_2, \vec{p}_3) = \int d^3x_1 d^3x_2 d^3x_3 e^{-i\sum_i \vec{p}_i \cdot \vec{x}_i} \langle \delta_S(\vec{x}_1)\delta_S(\vec{x}_2)\delta_S(\vec{x}_3)\rangle .$$

$$(3.80)$$
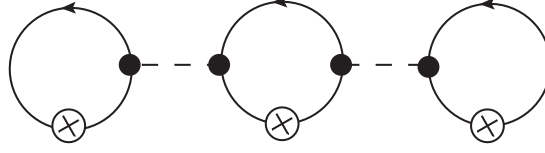
Figure 3.4: The leading order diagrams to 3-point function $\langle \delta_S \delta_S \delta_S \rangle$ is shown. The cross-dot vertices corresponds to $\bar{\psi}\psi/n_\psi$ insertions.

The fermion density fluctuation is intrinsically non-Gaussian since $n_\psi$ is the non-linear function of $\sigma$, which is treated as a Gaussian random variable. When the effective mass fluctuation due to $\lambda\sigma$ is small, we can Taylor-expand the number density with respect to $\lambda\sigma$,

$$n_\psi\left(m_\psi + \lambda\sigma\right) = n_\psi\left(m_\psi\right) + \lambda\left(\partial_{m_\psi} n_\psi(m_\psi)\right)\sigma + \frac{1}{2}\lambda^2\left(\partial_{m_\psi}^2 n_\psi(m_\psi)\right)\sigma^2 + O(\lambda^3). \quad (3.81)$$

Then the bispectrum is written as

$$B_S(\vec{p}_1,\vec{p}_2,\vec{p}_3) = \lambda^4 \omega_\psi^3 \frac{(\partial_m n_\psi)^2(\partial_m^2 n_\psi)}{n_\psi^3}\left[\Delta_\sigma^2(p_1)\Delta_\sigma^2(p_2) + 2\text{ perms}\right] + O(\lambda^6) \quad (3.82)$$

which is shown diagrammatically in Fig. 3.4. Now we compare this with the observational non-Gaussianities using the conventional non-Gaussian parameter $f_{NL}$ defined by

$$B_\zeta(\vec{p}_1,\vec{p}_2,\vec{p}_3) \equiv \frac{6}{5}f_{NL}\left[\Delta_\zeta^2(p_1)\Delta_\zeta^2(p_2) + 2\text{ perms}\right]. \quad (3.83)$$

Identifying $B_\zeta$ as the bispectrum of the temperature fluctuation using Eq.(3.68) and comparing it with $B_S$, we find in the squeezed triangle limit

$$f_{NL}^S = \frac{8\,B_S}{B_\zeta|_{f_{NL}=1}} = 8\frac{5}{6}\lambda^4\omega_\psi^3\frac{(\partial_m n_\psi)^2(\partial_m^2 n_\psi)}{n_\psi^3}\frac{\Delta_\sigma^2(p_1)\Delta_\sigma^2(p_2) + 2\text{ perms.}}{\Delta_\zeta^2(p_1)\Delta_\zeta^2(p_2) + 2\text{ perms.}}. \quad (3.84)$$

The factor 8 arises because the radiation transfer function for isocurvature is twice larger than that for adiabatic perturbation for the low multipoles of the CMB anisotropy as shown in Eq. (3.68). Although the isocurvature non-Gaussianities parameter $f_{NL}^S$

should not be compared directly with $f_{NL}$ defined by the curvature perturbation [203], this can be done with the extra $O(1)$ correction factor [18, 59, 90, 181, 204, 205]. The reason why $\partial_m^2 n_\psi$ appears instead of a first derivative is because of the squeezed triangle limit allows the short distance propagator to become important. Furthermore, subhorizon physics via the Yukawa interaction, in principle, gives rise to the non-Gaussianities of other types, e.g., the equilateral type. We postpone this study for future work.

In order to obtain the functional structure of $n_\psi(m, H; t)$, which relies on the background behavior, we specialize to the case of the inflaton coherent oscillation reheating scenarios, in which the total fermion number freezes during the reheating. During the early stage of the reheating when the inflaton field oscillates coherently, the equation of state of the inflaton is zero and the background behaves like the matter dominated (MD) era. After approximating the early stage of the reheating to the MD-like era (i.e. inflaton coherent oscillations period), we get (see Eq. (3.127))

$$n_\psi(t) \sim \frac{m_\psi^3}{3\pi^2} \left( \frac{a(t_m)}{a_t} \right)^3 \sim m_\psi H_e^2 \left( \frac{a_e}{a_t} \right)^3 \tag{3.85}$$

However, this leading order result gives $\partial_{m_\psi}^2 n_\psi = 0$ which renders $f_{NL}^S = 0$ via Eq.(4.31).

To find the non-zero result of $f_{NL}^S$, we need to study the mass dependence of $n_\psi$ in more detail, which in turn requires the knowledge of $|\beta_k(t; m)|^2$. To this point, we have approximated our spectrum by $|\beta_k(t; m)|^2 \sim 1/2\Theta(k_* - k)$, where $k_* = a(t_*)m$ and $t_*$ is the time when $m = H$. However, in general the spectrum should contain more than one characteristic scale, such as $k_e = a(t_e)H_e$ where $t_e$ marks the end of inflation. Thus, in general, the number density should contain a correction factor $f(\frac{m}{H_e})$ i.e.

$$n_\psi \sim m_\psi H_e^2 \left( \frac{a_e}{a_t} \right)^3 f(\frac{m_\psi}{H_e}) \tag{3.86}$$

and $f(0) = 1$. This higher order correction to $n_\psi$ would render $\partial_m^2 n_\psi \neq 0$ for the MD-like reheating scenario.

For simplicity, if we assume that $f(x) = 1 + a_1 x^{18}$, then in the limit where $\Delta_\sigma^2$, $\Delta_\zeta^2$, and $\Delta_{\delta_S}^2$ are scale invariant, we find

$$f_{NL}^S \quad \sim \quad a_1 \left( \frac{\alpha_S(\lambda, m_\psi, H_e, T_{RH})}{0.02} \right)^2 \left( \frac{\Omega_\psi h^2(m_\psi, T_{RH})}{10^{-7}} \right)^{-1} \left( \frac{m_\psi/H_e}{10^{-1}} \right). \quad (3.87)$$

Although we would naively guess $a_1 \sim O(1)$, the justification of the Taylor expansion for $f(x)$ and the estimation of the coefficient $a_1$ will be left for future work since the main thrust of this work is the computation of isocurvature perturbations and not the non-Gaussianities. The maximum $f_{NL}$ for the $m_\psi \gtrsim \lambda H_{inf}/(2\pi)$ case (consistent with small mass correction case) is achieved when this inequality is saturated and $\alpha_S$ is at its phenomenological maximum. We find this maximum to be at

$$f_{NL,\text{max}}^S \sim O(100) a_1 \frac{m_\psi}{H_{inf}/(2\pi)}. \quad (3.88)$$

Recall that our scenario assumes that $2\pi m_\psi/H_{inf} < 1$. Hence, although $f_{NL}^S$ cannot be made arbitrarily large, there may exist a parametric regime in which $f_{NL}^S$ is observable depending on $a_1$. Note that this extremum value corresponds to making the inhomogeneities $O(1)$ while staying consistent with phenomenology through the $\omega_\psi$ dilution factor: i.e. at this parametric point, the fermion abundance is $\Omega_\psi h^2 \approx 10^{-6}$ while most of the CDM is made up of assumed dark matter different from $\psi$.

---

[18] On very general grounds, $n_\psi$ cuts off exponentially at very large masses, $m_\psi \gtrsim H_e$, and we qualitative estimate the correction factor $f$ from this exponential cut-off, which gives an $O(1)$ value for $a_1$.

# 3.7 Natural Suppression of Gravitational Coupling to the Inflaton

As briefly discussed in 3.2, the gravity induced coupling of the fermion to the inflaton give a suppressed contribution to the isocurvature correlation function. We would like to consider this in more detail in this section. In addition, the argument below also shows that $\langle \bar{\psi}\psi\zeta \rangle$ cross-correlation is negligible, justifying the classification of this fermionic isocurvature perturbations as uncorrelated.

First, consider the $\zeta\psi\psi$ interaction given by Eq. (3.213) following the argument given in Ref. [94]. In this case, the most important coupling term is $a^2\zeta\delta_{ij}T_{\psi}^{ij} \in \mathcal{H}_{int}$ because the other interactions are derivatively suppressed, and decays as $O(1/a^2)$ or faster. Since $\zeta$ also freezes outside the horizon, using the similar argument given surrounding Eq. (3.65) we can factor the $\zeta$ correlation function out of the dominantly contributing integral, which corresponds to the diagram (b). Then we have

$$
\begin{aligned}
I_{\zeta\psi\psi}(x,y) &\approx (i)^2\langle\zeta_{\{z_0}\zeta_{w_0\}}\rangle[\int_{t_r}^{t} dt_z \int d^3z\, a^3(t_z)\langle[\bar{\psi}\psi_x, T_{\psi\,i}^i(z)]\rangle] \\
&\times [\int_{t_r}^{t} dt_w \int d^3w\, a^3(t_w)\langle[\bar{\psi}\psi_y, T_{\psi\,i}^i(w)]\rangle] + O\left(\frac{a^2(t_r)}{a^2(t)}\right) \quad (3.89)
\end{aligned}
$$

where $z_0 = (t_*, \vec{x})$, $w_0 = (t_*, \vec{y})$, $t = x^0 = y^0$, and $t_r$ denotes the time that the comoving distance $r = |\vec{x} - \vec{y}|$ crosses the horizon during inflation. In the integral, we have assumed the PV regulator. Note that $\lambda \int (dz) T_{\psi\,i}^i$ is a generator of the spatial dilatation, $x^i \rightarrow (1 + \lambda)x^i$ which is an element of diffeomorphism. Thus, we have

$$
\int_{-\infty}^{t} dt_z \int d^3z\, a^3(t_z)\langle[\bar{\psi}\psi_x, T_{\psi\,i}^i(z)]\rangle = 0 \quad (3.90)
$$

because $\bar{\psi}\psi$ is a diffeomorphism invariant scalar. Indeed, this is a Ward identity similar to that of Ref. [94]. Although the integral in Eq. (3.89) does not completely vanish (because of the time integral limit being $t_r$ and not $-\infty$), the mode function

of $\psi$ decays as $1/a^3$ (as shown Appendix 3.11) because of the classical conformal symmetry characterizing the massless fermionic sector[19], we have

$$\int_{t_r}^{t} dt_z \int d^3z \, a^3(t_z) \langle [\bar{\psi}\psi_x, T^i_{\psi\,i}(z)] \rangle \sim O\left(\frac{a^3(t_r)}{a^3(t)}\right). \qquad (3.91)$$

In a similar manner, we can have

$$\left\langle \zeta_x \left(\bar{\psi}\psi\right)_y \right\rangle \sim O\left(\frac{a^2(t_r)}{a^2(t)}\right). \qquad (3.92)$$

Therefore, we can conclude that large scale density perturbations of $\psi$ particles generated by $\zeta$ interaction and the curvature and isocurvature cross-correlation via the $\zeta\bar{\psi}\psi$ are negligible.

## 3.8   Summary and Conclusion

In this work, we have presented a fermionic isocurvature scenario which contains fermionic field fluctuation information during inflation. To our knowledge, this is the first work that describes isocurvature inhomogeneities of fermionic fields during inflation. Because massless free fermions have a tree-level conformal symmetry, such isocurvature models must couple to a conformal symmetry breaking sector. Because the $\zeta$ sector coupling to fermion $\psi$ is suppressed due to the dilatation symmetry, an additional scalar sector $\sigma$ is coupled to $\psi$ (with mass $m_\psi$) through a Yukawa coupling with strength $\lambda$. Composite operator renormalization in curved spacetime plays an important role in determining the isocurvature perturbations. We have computed the

---

[19]Thus, the result is different for a scalar case, which is minimally coupled to gravity. In particular, the cross-correlation for the light scalar case is computed in Ref. [94] and is

$$\langle \zeta(t,\vec{x}) \, \sigma^2(t,\vec{y}) \rangle \sim O\left( \left(\frac{a(t_r)}{a(t)}\right)^{3-2\nu}, \left(\frac{a(t_r)}{a(t)}\right)^2 \right),$$

where $\nu \equiv \sqrt{\frac{9}{4} - \frac{m_\sigma^2}{H^2}}$.

fermion isocurvature two point correlation function which has its dominant contribution in the long wavelength limit coming at one loop 1PI level. We have also estimated the local non-Gaussianity and found a value that is promising for observability for a particular corner of the parameter space.

As far as the existence proof inspired "minimal" model of this work is concerned, a large phenomenologically viable parameter region spanned by $\{\lambda, m_\psi\}$ exists for various inflationary models controlled by $\{H_{inf}, T_{RH}\}$. The large $\lambda$ parameter region is bounded either by current CMB constraints on isocurvature perturbations or the constraint of $\sigma$ not decaying to $\psi$. The large $m_\psi$ region is constrained by the relic abundance non-overclosure. The small $m_\psi$ region is constrained by requiring that $\sigma$ not decay to $\psi$ (for a fixed $\lambda$ and $H_{inf}$). The large non-Gaussianity parametric region is associated with largest $\lambda$ consistent with isocurvature bounds and the simplifying assumption $m_\psi \gtrsim \lambda H_{inf}/(2\pi)$. This intuitively corresponds to a large fermion inhomogeneity (i.e. $\delta\rho_\psi/\bar{\rho}_\psi \sim O(1)$) with a tiny $\bar{\rho}_\psi/(\bar{\rho}_\psi + \bar{\rho}_m)$ where $\bar{\rho}_m$ corresponds to an adiabatic cold dark matter component that helps saturate the phenomenologically measured cold dark matter abundance.

Our results regarding the gravitational fermion production give good dynamical intuition on many models with dynamical fermions existing during inflation. One shortcoming of the explicit model used in the current work is the tuning of the $\sigma$ sector imposed to keep it light and to prevent any $\sigma$ decay into $\psi$. In addition to model building issues, it would be interesting to consider in the future non-Gaussianities from such models more completely and carefully beyond the estimation presented in this work. It may also be interesting to see what UV model fermionic sector built independently of cosmological motivation can be constrained using the analysis presented in this work.

# 3.9 Appendix: Scalar and Spinor fields in Curved spacetime

First we list the relevant results about scalar field. Consider the following action

$$S = \int d^4x \sqrt{|g|} \left\{ -\frac{1}{2} g^{\alpha\beta} \partial_\alpha \phi \partial_\beta \phi - \frac{1}{2} m^2 \phi^2 - \frac{1}{2} \xi R \phi^2 \right\}, \tag{3.93}$$

This gives rises to equation of motion

$$\frac{1}{\sqrt{|g|}} \partial_\mu (g^{\mu\nu} \sqrt{|g|} \partial_\nu \phi) - (m^2 + \xi R)\phi = 0 \tag{3.94}$$

Scalar product between two solutions are defined as

$$(\phi_1, \phi_2) = -i \int_\Sigma [\phi_1 \partial_\mu \phi_2^* - \phi_2 \partial_\mu \phi_1^*] \sqrt{|g_\Sigma|} d\Sigma^\mu \tag{3.95}$$

where $\Sigma$ is a spacelike hypersurface.

For FRW metric, we can use mode decomposition

$$\phi(x) = \int d^3k (c_{\vec{k}} u_{\vec{k}}(x) + c_{\vec{k}}^\dagger u_{\vec{k}}^*(x)) \tag{3.96}$$

with the normalization condition

$$[c_{\vec{k}}, c_{\vec{p}}^\dagger] = \delta^3(\vec{k} - \vec{p}) \tag{3.97}$$

$$(u_{\vec{k}}, u_{\vec{p}}) = \delta^3(\vec{k} - \vec{p}) \tag{3.98}$$

The mode functions can be written explicitly as

$$u_{\vec{k}}(x) = \frac{e^{i\vec{k}\cdot\vec{x}}}{(2\pi)^{3/2}a(\eta)} f_k(\eta) \tag{3.99}$$

$$f_k \partial_\eta f_k^* - f_k^* \partial_\eta f_k = i \tag{3.100}$$

The time-part of the mode function obeys the differential equation

$$\frac{d^2}{d\eta^2} f_{k,\eta} + \{k^2 + a_\eta^2 [m^2 + (\xi - \frac{1}{6} R(\eta))]\} f_{k,\eta} = 0 \tag{3.101}$$

where $R(\eta) = 6a^{-1}\partial_\eta^2 a$, and $\eta$ is the conformal time. For de Sitter spacetime, the mode solution for a minimally coupled scalar ($\xi = 0$) is

$$f_k(\eta) = \frac{1}{\sqrt{2k}}\sqrt{\frac{\pi}{2}\left(\frac{k}{aH}\right)}e^{i\frac{\pi}{2}(\nu+\frac{1}{2})}H_\nu^{(1)}(\frac{k}{aH}) \tag{3.102}$$

where $\nu^2 = \frac{9}{4} - \frac{m^2}{H^2}$.

The following relations of first kind of Hankel functions are useful

$$H_\nu^{(1)}(z) \rightarrow -i\frac{\Gamma(\nu)}{\pi}\left(\frac{2}{z}\right)^\nu \quad (z \rightarrow 0) \tag{3.103}$$

$$H_\nu^{(1)}(z) \rightarrow \sqrt{\frac{2}{\pi z}}e^{-i\frac{\pi}{2}(\nu+\frac{1}{2})}e^{iz} \quad (z \rightarrow \infty) \tag{3.104}$$

From the mode expansion, we may construct the equal-time correlator in dS spacetime. In particular, we are interested in the large separation limit. For light scalar, when $\nu$ is real, we have

$$\langle\sigma_x\sigma_y\rangle \approx \frac{H^2}{8\pi}\frac{\Gamma(\frac{3}{2}-\nu)}{\Gamma(\frac{3}{2})\Gamma(1-\nu)\sin(\nu\pi)}(aHr)^{2\nu-3} \tag{3.105}$$

For heavy scalar, when $\nu = i\alpha$ and if $\alpha \sim \frac{m}{H} \gg 1$, then

$$\langle\sigma_x\sigma_y\rangle \approx \frac{H^{3/2}m^{1/2}}{\pi^{3/2}}e^{-\frac{m}{H}\pi}\sin[2\frac{m}{H}\ln(aHr) - \frac{1}{4}\pi](aHr)^{-3} \tag{3.106}$$

Next, we give the result for spinor field. Consider the free Dirac field $\psi$ action

$$S = \int(dx)\left(i\bar{\psi}\gamma^\mu\nabla_\mu\psi - m\bar{\psi}\psi\right). \tag{3.107}$$

where $(dx) = d^4x\sqrt{|g_x|}$ and $\gamma^\mu \equiv \gamma^a e_a^\mu$ with vierbein $e_a^\mu$. The covariant derivatives for $\psi$ is defined by

$$\nabla_\mu\psi = \partial_\mu\psi + \frac{1}{2}\omega_\mu^{ab}\Sigma_{ab}\psi \tag{3.108}$$

and the spin-connection is defined by

$$\omega_\mu^{ab} = e_\nu^a\nabla_\mu e^{b\nu} \tag{3.109}$$

and the Lorentz generator on the spinor field is given by

$$\Sigma_{ab} = -\frac{1}{4}[\gamma_a, \gamma_b], \tag{3.110}$$

where the $\gamma$ matrices satisfy the $\{\gamma_a, \gamma_b\} = -2\eta_{ab}$ with $\eta \equiv \operatorname{diag}(-1, 1, 1, 1)$. Note that the sign convention is chosen such that $[\Sigma^{12}, \Sigma^{23}] = \Sigma^{13}$.

Extremizing the action with respect to $\delta\bar{\psi}$ and $\delta\psi$ yields the equations of motion:

$$(i\gamma^\mu \nabla_\mu - m)\psi = 0, \qquad \nabla_\mu \bar{\psi}(-i\gamma^\mu) - \bar{\psi}m = 0. \tag{3.111}$$

The solution space can be endowed with a scalar product as

$$(\psi_1, \psi_2)_\Sigma = \int d\Sigma n_\mu \bar{\psi}_1 \gamma^\mu \psi_2 \tag{3.112}$$

in which $\Sigma$ is an arbitrary space-like hypersurface, $d\Sigma$ is the volume 3-form on this hypersurface computed with the induced metric, and $n_\mu$ is the future-pointing time-like unit vector normal to $\Sigma$. The current conservation condition

$$\nabla_\mu(\bar{\psi}_1 \gamma^\mu \psi_2) = 0 \tag{3.113}$$

implies the integral in the scalar product is independent of the choice of $\Sigma$.

If we adopt the Dirac basis for the $\gamma$ matrices, i.e.

$$\gamma^0 = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}, \quad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix} \tag{3.114}$$

the mode functions can be written as

$$U_{\vec{k},r}(x) = \frac{1}{a_x^{3/2}} \frac{e^{i\vec{k}\cdot\vec{x}}}{(2\pi)^{3/2}} \begin{pmatrix} u_{A,k,x^0} \\ r\, u_{B,k,x^0} \end{pmatrix} \otimes h_{\hat{k},r} \tag{3.115}$$

$$V_{\vec{k},r}(x) = -i\gamma^2 U_{\vec{k},r}^*(x) = \frac{1}{a_x^{3/2}} \frac{e^{-i\vec{k}\cdot\vec{x}}}{(2\pi)^{3/2}} \begin{pmatrix} r\, u_{B,k,x^0}^* \\ -u_{A,k,x^0}^* \end{pmatrix} \otimes (-i\sigma_2) h_{\hat{k},r}^* \tag{3.116}$$

where $h_{\hat{k},r}$ is eigenvector of $\hat{k} \cdot \vec{\sigma}$. The normalization conditions requires

$$h_{\hat{k},r}^{\dagger} h_{\hat{k},s} = \delta_{rs} \tag{3.117}$$

$$|u_{A,k,\eta}|^2 + |u_{B,k,\eta}|^2 = 1. \tag{3.118}$$

The time dependent parts of the mode functions obey the following equation

$$i\frac{d}{d\eta}\begin{pmatrix} u_A \\ u_B \end{pmatrix} = \begin{pmatrix} am & k \\ k & -am \end{pmatrix}\begin{pmatrix} u_A \\ u_B \end{pmatrix}. \tag{3.119}$$

In the special case of the de Sitter background with Bunch-Davies boundary condition, we have

$$\begin{pmatrix} u_A \\ u_B \end{pmatrix}_{k,\eta}^{in} = \begin{pmatrix} \sqrt{\frac{\pi}{4}(\frac{k}{aH_e})}e^{i\frac{\pi}{2}(1-i\frac{m}{H_e})}H_{\frac{1}{2}-i\frac{m}{H_e}}^{(1)}(\frac{k}{aH}) \\ \sqrt{\frac{\pi}{4}(\frac{k}{aH_e})}e^{i\frac{\pi}{2}(1+i\frac{m}{H})}H_{\frac{1}{2}+i\frac{m}{H_e}}^{(1)}(\frac{k}{aH}) \end{pmatrix} \tag{3.120}$$

$$\text{if } |kx^0| \ll 1 \longrightarrow \begin{pmatrix} \frac{1}{\sqrt{2\pi}}e^{\frac{\pi}{2}\frac{m}{H}}e^{-im(t-t_e)+i\frac{m}{H}\ln(2k/a_eH)}\Gamma(\frac{1}{2}-i\frac{m}{H}) \\ \frac{1}{\sqrt{2\pi}}e^{-\frac{\pi}{2}\frac{m}{H}}e^{+im(t-t_e)-i\frac{m}{H}\ln(2k/a_eH)}\Gamma(\frac{1}{2}+i\frac{m}{H}) \end{pmatrix} \tag{3.121}$$

Since the interaction picture operator $\psi(x)$ obeys the same classical equations, Eq. (3.111), we can expand the operator using $\{U_i, V_i\}$ as the basis:

$$\psi(x) = \sum_i a_i U_i(x) + b_i^{\dagger} V_i(x) \tag{3.122}$$

and the normalization conditions on $U_i, V_i$ gives the usual canonical anti-commutation relations of the creation and annihilation operators.

The first order WKB approximation is defined as

$$\begin{pmatrix} u_A \\ u_B \end{pmatrix}_{k,\eta}^{WKB} = \begin{pmatrix} \sqrt{\frac{\omega+am}{2\omega}} \\ \sqrt{\frac{\omega-am}{2\omega}} \end{pmatrix}e^{-i\int^{\eta}\omega d\eta'} \tag{3.123}$$

In the following, when we talk about fermion particle, we are implicitly referring to the WKB-mode.

Thus one can introduce the time-dependent Bogoliubov coefficients $\{\alpha_{k,\eta}, \beta_{k,\eta}\}$ between the in-modes and WKB-modes:

$$\begin{pmatrix} u_A \\ u_B \end{pmatrix}_{k,\eta}^{in} = \alpha_{k,\eta} \begin{pmatrix} u_A \\ u_B \end{pmatrix}_{k,\eta}^{WKB} + \beta_{k,\eta} \begin{pmatrix} u_B^* \\ -u_A^* \end{pmatrix}^{WKB}. \tag{3.124}$$

Clearly, $(\alpha, \beta) \to (1, 0)$ as $\eta \to -\infty$. We may also note that the Bogoliubov coefficients obey normalization condition as

$$|\alpha_{k,\eta}|^2 + |\beta_{k,\eta}|^2 = 1. \tag{3.125}$$

in agreement with fermion statistics.

Using Eq. (3.124), (3.123) and (3.119), we can derive the evolution equation for the Bogoliubov coefficients, as shown in Eq. (3.126).

## 3.10 Appendix: Review of fermion particle production

In this section, we give a brief review of the main result about fermion production during inflation [185]. The fermion number density can be obtained by solving this equations of Bogoliubov coefficients

$$\partial_\eta \begin{pmatrix} \alpha_{k,\eta} \\ \beta_{k,\eta} \end{pmatrix} = \frac{a^2 m \, k \, H}{2\omega^2} \begin{pmatrix} 0 & e^{2i \int^\eta \omega d\eta'} \\ -e^{-i \int^\eta \omega d\eta'} & 0 \end{pmatrix} \begin{pmatrix} \alpha_{k,\eta} \\ \beta_{k,\eta} \end{pmatrix} \tag{3.126}$$

We define the non-adiabaticity for a mode $k$ as $\epsilon_{k,\eta} = \frac{m \, k_p \, H}{\omega_p^3}$, where subscript $p$ stand for "physical", $\omega_p = \omega/a$ etc. As the system evolves from an initial vacuum condition of $(\alpha_{k,\eta}, \beta_{k,\eta}) = (1, 0)$, $\beta_{k,\eta}$ will only increase significantly when $\epsilon_{k,\eta} \sim O(1)$. This implies the following results,

1. In the heavy mass limit $(m_\psi \gg H_{inf})$, $\epsilon_{k,\eta}$ is always suppressed by $\frac{H}{m_\psi}$, we get $|\beta_{k,\eta}|^2 \sim \exp[-C\frac{m_\psi}{H(\eta_k)}] \ll 1$, where $C$ is some order one constant and $H(\eta_k)$ is the Hubble rate at the most non-adiabatic moment for mode $k$.

2. In the light mass limit $(m_\psi \ll H_{inf})$, $\epsilon_{k,\eta}$ is largest when $k_p \sim m_\psi$, we call this time $\eta_k$. If $m_\psi < H(\eta_k)$, we have $|\beta_k|^2 \sim \frac{1}{2}$, otherwise it is suppressed by $\exp[-C\frac{m_\psi}{H(\eta_k)}]$ as well.

Since the heavy fermion production is always exponentially suppressed by $m_\psi/H$ ratio, we focus on the light fermion case. The energy density at time $t$ is given by

$$\rho(t) \sim \frac{m_\psi^4}{3\pi^2} \left(\frac{a(t_*)}{a(t)}\right)^3,  \tag{3.127}$$

where $t_*$ is the time when $H(t) = m_\psi$. If $t_*$ occurs during reheating, one get the relic abundance today time as

$$\Omega_\psi h^2 \sim 3 \left(\frac{m_\psi}{10^{11}\text{GeV}}\right)^2 \left(\frac{T_{RH}}{10^9\text{GeV}}\right).  \tag{3.128}$$

## 3.11 Appendix: Asymptotic behavior of $\langle\psi_x\bar{\psi}_y\rangle$ at large $r$

In this section we derive the result about leading order contribution to $\langle n_{\psi,x}n_{\psi,y}\rangle$, i.e. Eq. (3.52). By Wick contraction, this reduces to computing the field correlator $\langle\psi_x\bar{\psi}_y\rangle$. The standard way to compute the correlator is to plug in the mode decomposition Eq. (3.122) and compute the mode functions $\{U_i, V_i\}$. The difficulties lie in how to obtain the mode functions on a curved spacetime. For inflationary background spacetime, one can use the de Sitter spacetime as an approximation and obtain exact analytic solutions. However, it is unclear how do these mode solutions evolve after

inflation ends. Such postinflationary solutions are relevant for our computation because the particle production freezes out after the end of inflation. Here we give an approach that answers this question.

First, we plug in the mode decomposition to the equal-time correlator:

$$
\langle \psi_x \bar{\psi}_y \rangle = \int d^3k \frac{1}{a_x^3} \frac{e^{i\vec{k}\cdot\vec{r}}}{(2\pi)^3}
$$
$$
\times \begin{pmatrix} |u_{A,k,x^0}|^2 \otimes I_2 & -u_{A,k,x^0} u_{B,k,x^0}^* \otimes (\hat{k}\cdot\vec{\sigma}) \\ u_{B,k,x^0} u_{A,k,x^0}^* \otimes (\hat{k}\cdot\vec{\sigma}) & -|u_{B,k,x^0}|^2 \otimes I_2 \end{pmatrix} \tag{3.129}
$$

where we have performed the spin-sum in the last step. Since

$$
\int d^3k \frac{e^{i\vec{k}\cdot\vec{r}}}{(2\pi)^3} |u_{A,k,x^0}|^2 = \int d^3k \frac{e^{i\vec{k}\cdot\vec{r}}}{(2\pi)^3} (1 - |u_{B,k,x^0}|^2) \tag{3.130}
$$
$$
= \delta^3(\vec{r}) - \int d^3k \frac{e^{i\vec{k}\cdot\vec{r}}}{(2\pi)^3} |u_{B,k,x^0}|^2 \tag{3.131}
$$

and $\vec{r} \neq 0$, we see the diagonal elements are the same. Then we perform the angular integral $d^2\hat{k}$. Recall that

$$
\int d^3k \, e^{i\vec{k}\cdot\vec{r}} f(k) = \int 4\pi k^2 dk \frac{\sin(kr)}{kr} f(k) \tag{3.132}
$$
$$
\int d^3k \, e^{i\vec{k}\cdot\vec{r}} \hat{k}_i f(k) = (-i\hat{r}_i\partial_r) \int 4\pi k^2 dk \frac{\sin(kr)}{kr} \frac{f(k)}{k} \tag{3.133}
$$

After the angular integral, we have

$$
\langle \psi_x \bar{\psi}_y \rangle = \int \frac{4\pi k^2 dk}{(2\pi)^3} \begin{pmatrix} A & B \\ B^* & C \end{pmatrix} \tag{3.134}
$$
$$
A = |u_{A,k,\eta}|^2 \cdot \frac{\sin(kr)}{kr} \tag{3.135}
$$
$$
B = (i\hat{r}\cdot\vec{\sigma})u_{A,k,\eta} u_{B,k,\eta}^* \cdot \partial_r \frac{\sin(kr)}{kr} \frac{1}{k} \tag{3.136}
$$
$$
C = -|u_{B,k,\eta}|^2 \cdot \frac{\sin(kr)}{kr} \tag{3.137}
$$

It is sufficient to study these two integrals for the diagonal and off-diagonal elements.

$$I_{11} = I_{22} = \int_0^\infty \frac{4\pi k^2 dk}{(2\pi)^3} |u_{A,k,\eta}|^2 \cdot \frac{\sin(kr)}{kr} \tag{3.138}$$

$$I_{12} = I_{21}^* = \partial_r \int_0^\infty \frac{4\pi k^2 dk}{(2\pi)^3} u_{A,k,\eta} u_{B,k,\eta}^* \frac{\sin(kr)}{kr} \frac{1}{k} \tag{3.139}$$

Now, we only need to find the mode function $u_A, u_B$, and perform the mode sum.

Let's consider the mode functions first. Since we are interested in evaluating the fermion field correlator at a time when the fermion production has ended, i.e. when $m \gg H(x^0)$ and in the limit $r \to \infty$, we can make the following approximations about the mode functions $\{u_{A,k,x^0}, u_{B,k,x^0}\}$. First, since the particle production has stopped, the non-adiabatic parameter is suppressed by $\frac{H(t)}{m}$, thus we can approximately replace the Bogoliubov coefficients by their late time asymptotic values, i.e.

$$\alpha_{k,x^0} \approx \alpha_k, \quad \beta_{k,x^0} \approx \beta_k. \tag{3.140}$$

Second, since we want to capture the particle production effect on the correlator and the produced particles are non-relativistic at the time of production, by the time $x^0$ which is sufficiently long after the production has ended, we may approximate the produced modes all have $k \ll a(x^0)m$. Thus, the WKB modes can be approximated by

$$\begin{pmatrix} u_A \\ u_B \end{pmatrix}_{k,\eta,IR}^{WKB} = \begin{pmatrix} \sqrt{\frac{\omega+am}{2\omega}} \\ \sqrt{\frac{\omega-am}{2\omega}} \end{pmatrix} e^{-i\int^\eta \omega d\eta'} \to \begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix} e^{-i\int^\eta \omega d\eta'}. \tag{3.141}$$

Combining these two approximations, we have

$$\begin{pmatrix} u_A \\ u_B \end{pmatrix}_{k,\eta,IR}^{in} \approx \begin{pmatrix} \alpha_k \frac{1}{\sqrt{2}} e^{-i\int^\eta \omega d\eta'} \\ -\beta_k \frac{1}{\sqrt{2}} e^{i\int^\eta \omega d\eta'} \end{pmatrix} \tag{3.142}$$

Thus we can easily evaluate $I_{11}, I_{12}$:

$$2\pi^2 I_{11,IR} = \frac{1}{r} \mathrm{Im} \int_0^\infty k dk \frac{1}{2}[1 - n(k)] \cdot e^{ikr} \tag{3.143}$$

We note that for the contribution from 1 vanishes

$$\frac{1}{r}\text{Im}\int_0^\infty kdk\,[1]\cdot e^{ikr} = \frac{1}{r}\text{Im}\int_0^\infty (is)ids\,[1]\cdot e^{-sr} = 0 \qquad (3.144)$$

For the contribution from $n(k)$ , we may assume it to be a real analytic function on $\mathbb{R}^+$ and can be analytically continuated to upper-right quadrant of the complex $k$ plane. The location of singularity of $n(k)$ determines contour of $k$. For example, we may consider the $n(k)$ for heavy fermion case ($m > H_{inf}$):

$$n(k)_{heavy} = \exp\left[-\frac{4(k/a_{nad})^2}{mH} - \frac{4m}{H}\right] \qquad (3.145)$$

where $a_{nad}$ is at the non-adiabatic time point. In this case, the non-adiabatic time is the transition from de Sitter era to the reheating era, i.e. $a_{nad} = a_e$. One can apply steepest descent to find that

$$2\pi^2 I_{11,heavy,IR}$$
$$\approx -\frac{1}{r}\exp[-\frac{4m}{H} - \frac{1}{16}mHr^2](a_e^2 mH)\text{Im}[-i\frac{1}{4}\sqrt{mH}a_e r\frac{1}{2}\sqrt{\pi}] \qquad (3.146)$$
$$= \frac{1}{8}\sqrt{\pi}a_e^3(mH)^{\frac{3}{2}}\exp[-\frac{4m}{H} - \frac{1}{16}a_e^2 mHr^2] \qquad (3.147)$$

For light fermion, we may approximate the number density spectrum as

$$n(k)_{light} = \frac{1}{1 + \exp(\frac{k^2}{(a_{nad}m)^2})} \qquad (3.148)$$

where the non-adiabatic point occurs when $H$ drops below $m$, i.e. $a_{nad} = a(\eta_*) = a_*$. This ansatz is only used to mimic the cut-off of the spectrum at $k \sim a_{nad}m$. The singularity lies at

$$\frac{k^2}{a_*^2 m^2} = (2n+1)\pi i, \qquad n = 0, 1, 2\cdots \qquad (3.149)$$

or $k_{*,n} = a_* m\sqrt{(2n+1)\pi}e^{\frac{\pi}{4}i}$. Again, one can perform the steepest descent around the $n = 0$ singularity $k_* = a_* m\sqrt{\pi}e^{\frac{\pi}{4}i}$. Let $\delta = (k - k_*)/a_* m$, we have

$$2\pi^2 I_{11,light,IR} = \pi a_*^3\frac{m^2}{a_* r}\exp[-\sqrt{\frac{\pi}{2}}a_* mr]\cos(\sqrt{\frac{\pi}{2}}a_* mr) \qquad (3.150)$$

For both the heavy and light fermion case, $I_{11} \propto \exp(-a_* M r)$, where $a_* M$ is the scale that $n(k)$ cuts off. We should also remind ourself that the UV vacuum contributions also exist, which scales as

$$I_{11,UV} \propto \exp[-a_\eta m r] \tag{3.151}$$

due to the singularity at $k = a_\eta m$ in the mode functions $u_A^{WKB}, u_B^{WKB}$. Thus we have shown that the diagonal element of Eq. (3.134) is always exponentially suppressed.

Next, we turn to look at the off diagonal element $I_{12}$. Unlike the $I_{11}$ case, whose integrand $|u_A|^2$ has constant asymptotic value in the IR region, the $I_{12}$'s IR contribution

$$u_{A,k,\eta} u_{B,k,\eta}^* = \alpha_k \beta_k^* e^{-2i \int^\eta \omega d\eta'} \tag{3.152}$$

contains $e^{-2imt}$ time dependence. Physically, if we decompose the in-state into WKB vacuum and excitation state

$$|\text{in,vac}\rangle =\sim |\text{WKB,vac}\rangle + \sim |\text{WKB,2-particles}\rangle + \sim |\text{WKB,4-particles}\rangle \tag{3.153}$$

then this term comes from the interference term

$$\langle \text{WKB, vac}|\psi_x \bar{\psi}_y|\text{WKB, 2-particles}\rangle \in \langle \text{in, vac}|\psi_x \bar{\psi}_y|\text{in, vac}\rangle. \tag{3.154}$$

If we care about $r$ large enough, for example corresponding to the CMB observation scale at recombination, we may assume the relevant $k$ scale exit horizon and become non-relativistic during inflation. Thus we may safely use the dS mode function to evaluate $I_{12,IR,CMB}$.

Recall that during dS era, we have Eq. (3.120), where we choose the end of inflation time $t_e$ as the reference point. Thus

$$u_{A,k,\eta} u_{B,k,\eta}^* = \frac{1}{2\pi} e^{-2im(t-t_e)+2i\frac{m}{H} \ln(2k/a_e H)} \Gamma^2 \left(\frac{1}{2} - i\frac{m}{H}\right) \tag{3.155}$$

Performing the integral using steepest descent, we find the leading contribution comes from $k \sim 0$ singularity in $u_{A,k,\eta} u_{B,k,\eta}^*$. We note that the $k$ dependent phase factor

$e^{2i\frac{m}{H}\ln(2k/H)}$ cannot be absorbed by a redefinition of the mode functions $u_{A,k,\eta}, u_{B,k,\eta}$, since this phase factor depends on the relative phase of $u_{A,k,\eta}, u_{B,k,\eta}$ which is fixed by the Bunch-Davies initial condition.

Plugging in the Eq. (3.139), we have

$$
\begin{aligned}
&2\pi^2 I_{12,IR} \\
&= -e^{-2im(t-t(r))+i\phi(\frac{m}{H})}r^{-3}\sqrt{\frac{2\pi\frac{m}{H}}{\sinh(2\pi\frac{m}{H})}\left(1+\left(\frac{m}{H}\right)^2\right)}
\end{aligned}
\tag{3.156}
$$

where $\phi(\frac{m}{H}) = \mathrm{Arg}(\Gamma(2+ix)\Gamma(\frac{1}{2}-ix))$ and $t(r)$ is the time when $a(t_r)Hr = 4$. We may consider the light mass limit

$$
2\pi^2 I_{12,IR,light} \approx -e^{-2im(t-t(r))}r^{-3}
\tag{3.157}
$$

and the heavy mass limit

$$
2\pi^2 I_{12,IR,heavy} \approx -(4\pi)^{\frac{1}{2}}\left(\frac{m}{H}\right)^{\frac{3}{2}}\exp(-\pi\frac{m}{H})e^{-2im(t-t(r))}r^{-3}
\tag{3.158}
$$

We may also consider the effect of having an IR cut-off $k_{IR}$, which is the scale that exit horizon at the beginning of inflation. Such an IR cut-off will introduce a $\exp(-k_{IR}r)$ type of exponential suppression factor. However, for observable universe with comoving radius $R_{obs}$, as long as $k_{IR}R_{obs} \ll 1$, we may ignore this suppression factor.

After evaluating the matrix element for the fermion correlators, we find that

1. For the light fermion case, i.e. $m \ll H_{inf}$, in the limit $r \to \infty$

$$
\langle \psi_x \bar{\psi}_y \rangle \approx \frac{1}{a_x^3}\frac{1}{2\pi^2}\begin{pmatrix} A & B \\ B^* & A \end{pmatrix}
\tag{3.159}
$$

where

$$
A = \frac{1}{2}\pi a_*^3 \frac{m^2}{a_* r}\exp[-\sqrt{\frac{\pi}{2}}a_* mr]\cos(\sqrt{\frac{\pi}{2}}a_* mr)
\tag{3.160}
$$

$$
B = -i\hat{r}\cdot\vec{\sigma}e^{-2im(t-t_r)}r^{-3}
\tag{3.161}
$$

where $a_*$ in evaluated at $\eta_*$.

2. For the heavy fermion case, i.e. $m \gg H_{inf}$, in the limit $r \to \infty$, we find in Eq. (3.159)

$$A = \frac{1}{16}\sqrt{\pi}a_e^3(mH_e)^{\frac{3}{2}}\exp[-\frac{4m}{H_e} - \frac{1}{16}a_e^2 mH_e r^2] \tag{3.162}$$

$$B = -i\hat{r}\cdot\vec{\sigma}(4\pi)^{\frac{1}{2}}\left(\frac{m}{H_e}\right)^{\frac{3}{2}}\exp(-\pi\frac{m}{H_e})e^{-2im(t-t(r))}r^{-3} \tag{3.163}$$

and $a_e$ is evaluated at the end of inflation.

Finally, we plug in the field correlator to $\langle n_{\psi,x}n_{\psi,y}\rangle$, and drop the term that are exponentially suppressed when $r \to \infty$, to get Eq. (3.52).

## 3.12    Appendix: Relative suppression of Commutators

In this subsection, we want compare the dependence on the scale factor $a(t)$ between $\langle in|[O_x, O_y]|in\rangle$ and $\langle in|\{O_x, O_y\}|in\rangle$, where $O_x$ is a bosonic hermitian operator and $x, y$ are spacetime points located near the end of inflation. For simplicity, we take $H$ as a constant. In particular, we are interested in the cases where $O = \sigma, \bar{\psi}\psi, \zeta$. We want to show that the commutator of $O$ suffers from additional suppression factor compared to the anti-commutator.

In general, the diagonal matrix elements of products of hermitian operator obeys

$$(\langle in|O_x O_y|in\rangle)^* = \langle in|O_y O_x|in\rangle \tag{3.164}$$

therefore

$$\langle in|[O_x, O_y]|in\rangle = 2i\text{Im}\langle in|O_x O_y|in\rangle \tag{3.165}$$

$$\langle in|\{O_x, O_y\}|in\rangle = 2\text{Re}\langle in|O_x O_y|in\rangle \tag{3.166}$$

We can just study $\langle in|O_x O_y|in\rangle$. We may use the mode expansion of the field operator to evaluate such an expression, and focus on modes that are outside of horizon at both times $\eta_x, \eta_y$.

We shall first take $O = \sigma$, and we assume that the scalar is light, i.e. $m_\sigma < \frac{3}{2}H$, such that $\nu$ is real:

$$
\begin{aligned}
\langle in|\sigma_x \sigma_y|in\rangle &= \int 4\pi k^2 dk \frac{[\int d^2\hat{k}\, e^{i\vec{k}\cdot(\vec{x}-\vec{y})}]}{(2\pi)^3 a_x^{3/2} a_y^{3/2}} \\
&\times \frac{1}{H}\frac{\pi}{4}[J_x J_y + Y_x Y_y + i(Y_x J_y - J_x Y_y)]
\end{aligned}
\tag{3.167}
$$

where $J_x = J_\nu(\frac{k}{a_x H}), Y_x = Y_\nu(\frac{k}{a_x H})$ are the first and second kinds of Bessel functions with real values. The $d^2\hat{k}$ is the angular integral with normalization $\int d^2\hat{k} = 1$, and $\int d^2\hat{k}\, e^{i\vec{k}\cdot(\vec{x}-\vec{y})} = \sin(kr)/kr$ is real. If we focus on the $k$ modes that are outside of horizon, i.e. $k/aH \ll 1$, we may use the small argument expansion of the Bessel function, i.e. when $(0 < z < \sqrt{1+\nu})$

$$
J_\nu(z) \approx \frac{1}{\Gamma(\alpha+1)}\left(\frac{z}{2}\right)^\nu
\tag{3.168}
$$

$$
Y_\nu(z) \approx -\frac{\Gamma(\alpha)}{\pi}\left(\frac{2}{z}\right)^\nu.
\tag{3.169}
$$

Then, under the common scaling of $a_x \to \lambda a_x, a_y \to \lambda a_y$, with $\lambda$ increasing, we see the various term in the correlator scales as

$$
a_x^{-3/2} a_y^{-3/2} J_x J_y \propto \lambda^{-2\nu-3}
\tag{3.170}
$$

$$
a_x^{-3/2} a_y^{-3/2} Y_x Y_y \propto \lambda^{2\nu-3}
\tag{3.171}
$$

$$
a_x^{-3/2} a_y^{-3/2}(Y_x J_y - J_x Y_y) \propto \lambda^{-3}
\tag{3.172}
$$

Thus, we see under this common scaling, the IR contribution to the two point func-

tions are

$$\langle in|\{\sigma_x, \sigma_y\}|in\rangle_{IR} = 2\int_{IR} 4\pi k^2 dk \frac{[\int d^2\hat{k} e^{i\vec{k}\cdot(\vec{x}-\vec{y})}]}{(2\pi)^3 a_x^{3/2} a_y^{3/2}} \frac{1}{H}\frac{\pi}{4}(J_x J_y + Y_x Y_y) \quad (3.173)$$

$$\propto \lambda^{2\nu-3} \quad (3.174)$$

$$\langle in|[\sigma_x, \sigma_y]|in\rangle_{IR} = 2i\int_{IR} 4\pi k^2 dk \frac{[\int d^2\hat{k} e^{i\vec{k}\cdot(\vec{x}-\vec{y})}]}{(2\pi)^3 a_x^{3/2} a_y^{3/2}} \frac{1}{H}\frac{\pi}{4}(Y_x J_y - J_x Y_y) \quad (3.175)$$

$$\propto \lambda^{-3} \quad (3.176)$$

Thus, we have shown under the scaling $a \to \lambda a$, the commutator of $\sigma$ is suppressed by $\lambda^{-2\nu}$ factor relative to its anti-commutator. For small mass scalar, $\lambda^{-2\nu} \approx \lambda^{-3+\frac{2m^2}{3H^2}}$.

For the case of $O = \zeta$, we have similar statements as the scalar case with $\nu = \frac{3}{2}$, i.e. $\langle[\zeta_x, \zeta_y]\rangle_{IR}$ is suppressed by $\lambda^{-3}$ relative to $\langle\{\zeta_x, \zeta_y\}\rangle_{IR}$ under the scaling of $a \to \lambda a$.

Next, we consider the case of $O = \bar{\psi}\psi$. Using the mode decomposition Eq.(3.122) and mode functions Eq. (3.115,3.116), we have

$$\langle\bar{\psi}\psi_x\bar{\psi}\psi_y\rangle = \sum_{i,j} \frac{1}{a_x^3 a_y^3} \frac{e^{i(\vec{k}_i+\vec{k}_j)\cdot(\vec{x}-\vec{y})}}{(2\pi)^6}[h_i^T(i\sigma_2)h_j][h_j^\dagger(-i\sigma_2)h_i^*]F_{ij,x}F_{ij,y}^* \quad (3.177)$$

where

$$F_{ij,x} = r_i u_{B,i,x} u_{A,j,x} + (i \leftrightarrow j) \quad (3.178)$$

$$F_{ij,x}F_{ij,y}^* = 2[r_i u_{B,i,x} u_{A,j,x} + (i \leftrightarrow j)](r_i u_{B,i,y}^* u_{A,j,y}^*) \quad (3.179)$$

$$= 2[u_{B,i,x} u_{A,j,x} u_{B,i,y}^* u_{A,j,y}^* + r_i r_j u_{B,i,x} u_{A,j,x} u_{B,j,y}^* u_{A,i,y}^*]. \quad (3.180)$$

We note that in Eq. (3.177), the factor $e^{i(\vec{k}_i+\vec{k}_j)\cdot(\vec{x}-\vec{y})}$ after angular average is real, and the factor $[h_i^T(i\sigma_2)h_j][h_j^\dagger(-i\sigma_2)h_i^*] = |[h_i^T(i\sigma_2)h_j]|^2$ is also real, thus the imaginary and real part of $F_{ij,x}F_{ij,y}^*$ correspond to the commutator and anti-commutator respectively.

Next, we consider the two terms in Eq. (3.180) one by one, using explicit expression

of Eq. (3.121) to get

$$
\begin{aligned}
u_{B,i,x}u_{A,j,x}u^*_{B,i,y}u^*_{A,j,y} &= \sqrt{\frac{\pi}{4}\frac{k_i}{a_x H}}\sqrt{\frac{\pi}{4}\frac{k_j}{a_x H}}\sqrt{\frac{\pi}{4}\frac{k_i}{a_y H}}\sqrt{\frac{\pi}{4}\frac{k_j}{a_y H}} \\
&\times (J_{+,i,x}+iY_{+,i,x})(J_{-,j,x}+iY_{-,j,x}) \\
&\times (J_{-,i,y}-iY_{-,i,y})(J_{+,j,y}-iY_{+,j,y}) \qquad (3.181)
\end{aligned}
$$

where

$$
J_{\pm,i,x} = J_{\frac{1}{2}\pm i\frac{m}{H}}(\frac{k_i}{a_x H}), \quad Y_{\pm,i,x} = Y_{\frac{1}{2}\pm i\frac{m}{H}}(\frac{k_i}{a_x H}). \qquad (3.182)
$$

Using the small $z$ expansion of Bessel function again, where Re $(\nu) = \frac{1}{2}$ in all the cases, we can extract its scaling behavior under $a \to \lambda a$,

$$
\begin{aligned}
&(J_{+,i,x}+iY_{+,i,x})(J_{-,j,x}+iY_{-,j,x})(J_{-,i,y}-iY_{-,i,y})(J_{+,j,y}-iY_{+,j,y}) \\
&= Y_{+,i,x}Y_{-,j,x}Y_{-,i,y}Y_{+,j,y}\cdots\cdots \propto \lambda^2, \text{real} \\
&\quad -iJ_{+,i,x}Y_{-,j,x}Y_{-,i,y}Y_{+,j,y} - iY_{+,i,x}J_{-,j,x}Y_{-,i,y}Y_{+,j,y}\cdots\cdots \propto \lambda^1, \text{imaginary} \\
&\quad +iY_{+,i,x}Y_{-,j,x}J_{-,i,y}Y_{+,j,y} + iY_{+,i,x}Y_{-,j,x}Y_{-,i,y}J_{+,j,y}\cdots\cdots \propto \lambda^1, \text{imaginary} \\
&\quad +\text{terms subdominant in }\lambda\text{expansion.} \qquad (3.183)
\end{aligned}
$$

Thus the imaginary part is suppressed by $\lambda^{-1}$ relative to the real part. We can do similar analysis to the second part $r_i r_j u_{B,i,x}u_{A,j,x}u^*_{B,j,y}u^*_{A,i,y}$ in Eq. (3.180) and found the same behavior. Thus, for $\bar{\psi}\psi$ operator, we have the following scaling law

$$
\langle \{\bar{\psi}\psi_x, \bar{\psi}\psi_y\} \rangle_{IR} \propto \lambda^{-6} \qquad (3.184)
$$

$$
\langle [\bar{\psi}\psi_x, \bar{\psi}\psi_y] \rangle_{IR} \propto \lambda^{-7}. \qquad (3.185)
$$

Thus, we see the commutator for $\bar{\psi}\psi$ gives additional suppression of $a^{-1}$ factor compared with the anti-commutator, whereas the commutator for $\sigma$ and $\zeta$ gives additional suppression of $a^{-3}$ factor.

# 3.13    Appendix: Explicit check of the mass insertion formula

In this section, we show that the particle production part of the following equation holds using the adiabatic subtraction.

$$-i \int^y (dw) \langle [\bar{\psi}\psi_x, \bar{\psi}\psi_z] \rangle = \partial_m \langle \bar{\psi}\psi_x \rangle = \partial_m n_\Psi(x) \tag{3.186}$$

Expressing both side of Eq.(3.186) using the mode sum, we see the left hand side is

$$-i \int^y (dw) \langle [\bar{\psi}\psi_x, \bar{\psi}\psi_w] \rangle = \frac{16}{a_x^3} \int^{y^0} dw^0 \, a_w \int \frac{d^3k}{(2\pi)^3} \mathrm{Im}[(u_{A,k}u_{B,k})_x (u_{A,k}u_{B,k})_w^*] \tag{3.187}$$

and the right hand side is

$$\partial_m \langle \bar{\psi}\psi_x \rangle = \frac{2}{a_x^3} \int \frac{d^3k}{(2\pi)^3} \partial_m(|u_B|^2 - |u_A|^2) \tag{3.188}$$

Thus, we only need to check for each given $k$, the following equation is right

$$\partial_m(|u_B|^2 - |u_A|^2) = 8 \int^{y^0} dw^0 \, a_w \mathrm{Im}[(u_{A,k}u_{B,k})_x (u_{A,k}u_{B,k})_w^*] \tag{3.189}$$

From the left hand side, we have

$$\partial_m(|u_B|^2 - |u_A|^2) = -2\mathrm{Re} \left[ \begin{pmatrix} u_A^* & u_B^* \end{pmatrix} \sigma_3 \frac{\partial}{\partial m} \begin{pmatrix} u_A \\ u_B \end{pmatrix}_{k,x} \right] \tag{3.190}$$

and upon expressing mode function at time $x^0$ in term of evolution operator acting on the initial value, we have

$$\frac{\partial}{\partial m} \begin{pmatrix} u_A \\ u_B \end{pmatrix}_{k,x} = -i \int_{\eta_i}^{x^0} dz^0 \, U(x^0 \leftarrow z^0) \frac{\partial}{\partial m} \begin{pmatrix} am & k \\ k & -am \end{pmatrix}$$

$$\times U(z^0 \leftarrow \eta_i) \begin{pmatrix} u_A \\ u_B \end{pmatrix}_{k,i} \tag{3.191}$$

Combining these two expression, we can obtain the desired result after some algebra.

However, the remaining $d^3k$ integrals in Eq. (3.187) and Eq. (3.188) are UV divergent. To make them finite, we express both side in terms of Bogoliubov coefficients and dropped the pure vacuum contribution to get

$$
\begin{aligned}
-i \int^{x^0} (dw) \langle [\bar{\psi}\psi_x, \bar{\psi}\psi_w] \rangle &\approx 16 \int \frac{d^3k}{(2\pi a_x)^3} (\frac{am}{\omega_k})_x \int^x d\eta_w \, a_w (\frac{am}{\omega})_w \\
&\qquad \times \mathrm{Im}[(\alpha\beta)_x (\alpha\beta)^*_w] &(3.192) \\
\partial_m \langle \bar{\psi}\psi_x \rangle &\approx \frac{2}{a_x^3} \int \frac{d^3k}{(2\pi)^3} \partial_m [2|\beta_{k,x}|^2 \frac{a_x m}{\omega_{k,x}}] &(3.193) \\
&\approx \frac{4}{a_x^3} \int \frac{d^3k}{(2\pi)^3} (\frac{a_x m}{\omega_{k,x}}) \partial_m |\beta_{k,x}|^2 &(3.194)
\end{aligned}
$$

Now, we only need to check

$$
\partial_m |\beta_{k,x}|^2 = 4 \int^x d\eta_w \, a_w (\frac{am}{\omega})_w \mathrm{Im}[(\alpha\beta)_x (\alpha\beta)^*_w] \tag{3.195}
$$

Suppose, $x^0$ is late enough such that $\beta_{k,x}$ is constant and equals to its value at asymptotic future $\beta_k$, then we get

$$
\partial_m |\beta_k|^2 = 4 \int^{x^0}_{\eta_i} dz^0 a_z \frac{am}{\omega} \mathrm{Im}(\alpha_k \beta_k)_x (\alpha\beta)^*_z \tag{3.196}
$$

Thus, Eq. (3.186) is compatible with the Bogoliubov projection.

## 3.14  Appendix: Gravitational Interaction

Here we derive the gravitational interaction. Consider the action

$$
\begin{aligned}
S &= S_{EH} + S_\phi + S_\sigma + S_\psi &(3.197) \\
&= \int (dx) \left\{ \frac{1}{2} M_p^2 R + [-\frac{1}{2} g^{\mu\nu} \partial_\mu \phi \partial_\nu \phi - V(\phi)] + [-\frac{1}{2} g^{\mu\nu} \partial_\mu \sigma \partial_\nu \sigma - \frac{1}{2} m_\sigma^2 \sigma^2] \right. \\
&\qquad \left. + \bar{\psi}(i\gamma^\mu \nabla_\mu - m_\psi)\psi - \lambda\sigma\bar{\psi}\psi \right\}, &(3.198)
\end{aligned}
$$

where $M_p^2 = \frac{1}{8\pi G} = 1$. The metric is given in ADM formalism[20] [206] by

$$g_{\mu\nu} = \begin{pmatrix} -N^2 + h_{ij}N^iN^j & h_{ij}N^j \\ h_{ij}N^j & h_{ij} \end{pmatrix}, \quad g^{\mu\nu} = \begin{pmatrix} -N^{-2} & N^iN^{-2} \\ N^iN^{-2} & h^{ij} - N^iN^jN^{-2} \end{pmatrix},$$
(3.199)

where $h_{ij}$ is the metric tensor on the constant time hypersurface, and $h^{ij}$ is the inverse metric. We use Latin indices $i, j \cdots$ for objects on the 3-dimensional constant time hypersurface, and we use $h_{ij}$ and $h^{ij}$ to raise and lower the indices. Then we use the Hamiltonian and the momentum constraints to determine the lapse function $N$ and the shift vector $N^i$:

$$0 = \frac{1}{N}[R^{(3)} - \frac{1}{N^2}(E_{ij}E^{ij} - E^2)] - 2NT^{00}$$
(3.200)

$$0 = \frac{2}{N}\nabla_i^{(3)}[\frac{1}{N}(E^{ij} - Eh^{ij})] + 2N^jT^{00} + 2T^{0j},$$
(3.201)

where $T^{\mu\nu}$ is the total matter stress tensor, $R^{(3)}$ is the Ricci scalar calculated with the three-metric $h_{ij}$, and

$$E_{ij} = \frac{1}{2}(\dot{h}_{ij} - \nabla_i^{(3)}N_j - \nabla_j^{(3)}N_i).$$
(3.202)

$$E = E_{ij}h^{ij}.$$
(3.203)

In order to consider the perturbation around the background configuration

$$\phi^{(0)} = \bar{\phi}(t), \quad \sigma^{(0)} = 0, \quad g_{\mu\nu}^{(0)} = \begin{pmatrix} -1 & 0 \\ 0 & a^2(t)\delta_{ij} \end{pmatrix}$$
(3.204)

where the background fields satisfy the background equations of motion

$$3H^2 = \frac{1}{2}\dot{\bar{\phi}}^2 + V(\bar{\phi})$$
(3.205)

$$\dot{H} = -\frac{1}{2}\dot{\bar{\phi}}^2$$
(3.206)

$$\ddot{\bar{\phi}} + 3H\dot{\bar{\phi}} + V'(\bar{\phi}) = 0,$$
(3.207)

---

[20]We use $(-+++)$ sign convention for the metric, and physical time $t$.

we choose the comoving gauge, defined by [21]

$$\delta\phi = 0, \quad \gamma_{ii} = 0, \quad \partial_i \gamma_{ij} = 0 \tag{3.208}$$

where

$$h_{ij} = a^2(t)[e^\Gamma]_{ij}, \quad \Gamma_{ij} = 2\zeta\delta_{ij} + \gamma_{ij}. \tag{3.209}$$

Then we solve the constraint equations (3.200) and (3.201) perturbatively using $\zeta$ and $\gamma$, and putting their solutions for $N$ and $N^i$ back into the action, we can get the perturbed action:

$$S^{(C)} = S^{(C)}_{\zeta\zeta} + S^{(C)}_{\sigma\sigma} + S^{(C)}_{\psi\psi} + S^{(C)}_{\gamma\gamma} + S^{(C)}_{\zeta\zeta\zeta} + S^{(C)}_{\zeta\sigma\sigma} + S^{(C)}_{\zeta\psi\psi} + S^{(C)}_{\zeta\sigma\sigma} \cdots . \tag{3.210}$$

For the interaction terms $S^{(C)}_{\zeta\sigma\sigma}$ and $S^{(C)}_{\zeta\psi\psi}$, we need the solutions of $N$ and $N^i$ up to linear order in $\zeta$

$$N^{(1,C)} = 1 + \frac{\dot\zeta}{H}, \quad N_i^{(1,C)} = \partial_i[-\frac{\zeta}{H} + \epsilon\frac{a^2}{\nabla^2}\dot\zeta], \tag{3.211}$$

where $\epsilon \equiv \dot{H}/H^2$. Hence, the metric perturbations becomes

$$\delta g_{\mu\nu}^{(C)} = \begin{pmatrix} -2\frac{\dot\zeta}{H} & (-\frac{\zeta}{H} + \epsilon\frac{a^2}{\nabla^2}\dot\zeta)_{,i} \\ (-\frac{\zeta}{H} + \epsilon\frac{a^2}{\nabla^2}\dot\zeta)_{,i} & a^2\left(\delta_{ij}2\zeta + \gamma_{ij}\right) \end{pmatrix}, \tag{3.212}$$

and we have the $\zeta$-matter cubic interaction action

$$S^{(C)}_{\zeta\sigma\sigma} + S^{(C)}_{\zeta\psi\psi} = \frac{1}{2} \int d^4x \sqrt{-g} \left(T_\sigma^{\mu\nu} + T_\psi^{\mu\nu}\right) \delta g_{\mu\nu}^{(C)}, \tag{3.213}$$

where $T_\sigma^{\mu\nu}$ and $T_\psi^{\mu\nu}$ is the stress energy tensors for $\sigma$ and $\psi$, respectively, which are written as

$$T_\sigma^{\mu\nu} = g^{\mu\alpha}g^{\nu\beta}\partial_\alpha\sigma\partial_\beta\sigma + g^{\mu\nu}\mathcal{L}_\sigma, \tag{3.214}$$

$$T_\psi^{\mu\nu} = -\frac{i}{2}[\bar\psi\gamma^{(\mu}\nabla^{\nu)}\psi - \nabla^{(\mu}(\bar\psi)\gamma^{\nu)}\psi] + g^{\mu\nu}\mathrm{Re}\left(\mathcal{L}_\psi\right). \tag{3.215}$$

Particularly, up to the cubic interaction, $\mathcal{L}_{int} = -\mathcal{H}_{int}$. Thus $S^{(C)}_{\zeta\sigma\sigma} + S^{(C)}_{\zeta\psi\psi} = -\int dt\, H_{\zeta\sigma\sigma}(t) + H_{\zeta\psi\psi}(t)$.

---

[21] In this section, Latin indices $i, j$ are raised and lowered by $\delta_{ij}$, and repeated indices are contracted.

# Chapter 4

# Isocurvature Cross-correlation and Ward Identity

Sources of isocurvature perturbations and large non-Gaussianities include field degrees of freedom whose vacuum expectation values are smaller than the expansion rate of inflation. The inhomogeneities in the energy density of such fields are quadratic in the fields to leading order in the inhomogeneity expansion. Although it is often assumed that such isocurvature perturbations and inflaton-driven curvature perturbations are uncorrelated, this is not obvious from a direct computational point of view due to the form of the minimal gravitational interactions. We thus compute the irreducible gravitational contributions to the quadratic isocurvature-curvature cross-correlation. We find a small but non-decaying cross-correlation, which in principle serves as a consistency prediction of this large class of isocurvature perturbations. We apply our cross-correlation result to two dark matter isocurvature perturbation scenarios: QCD axions and WIMPZILLAs. On the technical side, we utilize a gravitational Ward identity in a novel manner to demonstrate the gauge invariance of the computation. Furthermore, the detailed computation is interpreted in terms of a soft-

$\zeta$ theorem and a gravitational Ward identity. Finally, we also identify explicitly all the counterterms that are necessary for renormalizing the isocurvature perturbation composite operator in inflationary cosmological backgrounds.

This work was performed in collaboration with Daniel Chung and Peng Zhou. It was published in Physical Review D in March of 2013 [94].

## 4.1   Introduction

As physics beyond the Standard Model is expected to contain many fields in addition to the inflaton, there are many candidates for isocurvature perturbations in the context of inflationary cosmology, including those of the dark matter. Indeed, the current data is consistent with the existence of an $O(2\%)$ isocurvature component [91, 97, 112, 200, 201, 207–209]. Furthermore, it is well known that quadratic isocurvature perturbations (i.e. the vacuum expectation value of the field is much smaller than the Hubble expansion rate) are one of the very few ways to generate measurably large local non-Gaussianities [18,40,45,47,51,59,132,152–174,181,210] in the context of the slow-roll inflationary paradigm. The only nontrivial requirement that the isocurvature field degree of freedom must possess is that it be light enough to be excited by the inflationary quasi-de Sitter (dS) background and that it not be conformally invariant. In the literature [67, 150, 151], quadratic isocurvature perturbations are often assumed to have negligible cross-correlations with the curvature perturbations (which corresponds to the inflaton field degree of freedom dressed by gravity). However, the gravitational interactions lead to a minimum cross-correlation, which in principle can be observationally important. We present a computation of this minimal gravitational cross-correlation in this work.

As explained below, the form of the gravitational interaction between the cur-

vature and isocurvature perturbations naively suggests that there can be cross correlators which do not vanish in the long wavelength limit. If this was true, the cross correlation can dominate over the isocurvature two-point function in the observables since the latter vanishes in the long wavelength limit for a massive field. By an explicit rigorous computation, we show that the cross correlator vanishes in the long wavelength in such a way that the cross correlation induced by gravity never dominates over the isocurvature two-point function, given that the curvature inhomogeneity perturbation is characterized by a strength of order $10^{-5}$. We explain this qualitatively as well using a combination of a soft-$\zeta$ theorem [19, 95, 211–226] and a Ward identity associated with a spatial dilatation diffeomorphism. We also check the gauge invariance of our computation using a Ward identity.

Among the possible isocurvature candidates, thermal dark matter is usually produced copiously by the inflaton decay products, which typically leads to a large suppression of isocurvature effects. On the other hand, nonthermal dark matter that is not produced by the inflaton decay can easily generate large isocurvature effects that survive until today. Hence, as an illustration, we apply our computation of the cross correlation to two different nonthermal dark matter models: QCD axions and WIMPZILLAs. In both cases, we find a cross-correlation characterized by the parameter $|\beta| \sim O(10^{-5})$ (the parameter definition is given in Eq. (4.29)) which is below the current observable sensitivity of $10^{-2}$. In principle, $\beta$ can be measured and is a generic consistency prediction of this class of nonthermal dark matter quadratic isocurvature models. Note that even though the nonthermal dark matter fields can be identified with the isocurvature degrees of freedom, this scenario is consistent with the WIMP dark matter scenario since the isocurvature perturbations can be as small as an order $10^{-5}$ fraction of the total dark matter and still leave an isocurvature imprint on the CMB spectrum.

The order of presentation is as follows. In Section 4.2, we present our assumptions about the inflationary cosmology, review gauge invariant variables in the perturbation theory, and summarize the observational constraints on the isocurvature scenario relevant to our work. One of the most important aspects of this section is our review of features of the $\beta$ variable that we compute. In Section 4.3, we first explain two naive estimates, one leading to the wrong observationally large result, and the other leading to the correct suppressed result. In explaining the correct estimate (which requires assumptions that cannot be known without the justification of a full computation), we present the interpretation in terms of a soft-$\zeta$ theorem and the Ward identity. The rigorous explicit computation at one loop is then presented, demonstrating how the correct naive estimate result is achieved. We also present in this section how gauge invariance is achieved for these quadratic isocuvature computations using the gravitational Ward identity. Next, we apply these results to the axion and the WIMPZILLA scenarios in Section 4.4. This section contains a detailed explanation for choosing nonthermal dark matter to illustrate the computations of our work instead of thermal dark matter. Finally, we summarize our results in Section 4.5. In appendices, we collect technical details and also supplementary computational results: the radiation transfer functions is derived in Appendix 4.6, the brief review of the gravitational Ward Identity is given in Appendix 4.7, ADM formalism is reviewed in Appendix 4.8, the details about the Pauli-Villars regulator is explained in Appendix 4.9, and the two point function computation in the uniform curvature gauge is presented in Appendix 4.10.

## 4.2 A Class of Curvature and Isocurvature Perturbations

Inflation through quantum correlator dynamics generates "classical" initial conditions for superhorizon cosmological fluid perturbations [8, 227–229]. The resulting initial conditions for the classical equations governing classical fluid variables (which are set during radiation domination before the CMB last scattering time) are categorized into two types: adiabatic and isocurvature [3, 230–232]. An adiabatic initial condition is intuitively characterized by all species composing the fluid having the same initial number overdensities. In the context of inflation, if there is a *single dynamical degree of freedom* $\phi$ during inflation such that after a few efolds of inflation, the quantum vacuum boundary can be approximated as Bunch-Davies initial conditions (for a discussion of number of efold requirement see e.g. [85]), and if all the degrees of freedom during radiation domination come from the inflaton decay, then this adiabatic condition is the resulting approximate classical boundary condition during radiation domination era of the universe. An isocurvature initial condition intuitively corresponds to setting nonzero the initial difference of the number overdensities of at least one pair of fluid element species while setting to zero the total energy density inhomogeneity on long wavelength scales. Because these two types of initial conditions are linearly independent, a generic initial condition to the linearized perturbation equations can be written as a linear combination of them.

In this work, we are concerned with the following physical system which is generic for isocurvature scenarios. One real scalar slow-roll inflaton degree of freedom $\phi$ dominates the energy density during inflation. During this time period, there exists also another light degree of freedom $\sigma$ which has no coupling to $\phi$ stronger than gravity. We assume that this system carries an approximately conserved discrete

charge (such as $\mathbb{Z}_2$ broken at most by a model dependent non-renormalizable operator) such that the one particle states are stable and can act as dark matter. Note that since we do not require all of the dark matter to come from $\sigma$, this system is consistent with the existence of the weakly interacting massive particle (WIMP) dark matter. If WIMP dark matter exists, the parameter $\omega_\sigma \equiv \Omega_\sigma/\Omega_{CDM} < 1$ will play a role, and this scenario can yield interesting isocurvature signatures for $\omega_\sigma$ as small as $10^{-5}$ [18]. The action of this system can thus be written as

$$
\begin{aligned}
S[\phi, \sigma, \{\psi\}] \;=\; & \int (dx) \left\{ \frac{1}{2} M_p^2 R + [-\frac{1}{2} g^{\mu\nu} \partial_\mu \phi \partial_\nu \phi - V(\phi)] + [-\frac{1}{2} g^{\mu\nu} \partial_\mu \sigma \partial_\nu \sigma - U(\sigma)] \right\} \\
& + S_{rh}[\phi, \{\psi\}]
\end{aligned}
\tag{4.1}
$$

where $R$ is the Ricci scalar, $M_p^2 = \frac{1}{8\pi G}$, $(dx) = d^4 x \sqrt{|\det(g_{\mu\nu})|}$, and $S_{rh}$ corresponds to the action of the reheating degrees of freedom $\{\psi\}$. We assume that $\{\psi\}$ is heavy during inflation such that it can be integrated out or if $\{\psi\}$ are light, they are conformal such that they are not excited during inflation. After inflation ends, we assume $\{\psi\}$ fields are light, leading to a successful reheating scenario. The only special initial condition dependent assumption that we make in this isocurvature scenario is that $\langle\sigma\rangle \ll H/(2\pi)$ during inflation even when $\partial^2 U(\sigma)/\partial\sigma^2 \ll H$.[1]   Because $\langle\sigma\rangle = 0$ during inflation, $\sigma$ by itself does not spontaneously break time translation invariance and therefore does not mix with $\delta\phi$ in forming the gauged time translation Nambu-Goldstone boson $\zeta$. Hence, we can treat the scalar fluid variable $\zeta(\delta g_{\mu\nu}, \delta\phi)$ as the curvature degree of freedom and $\delta_S(\sigma, \zeta)$ as the isocurvature degree of freedom. (As we will show in detail below, the isocurvature degree of freedom $\delta_S$ will be quadratic in $\sigma$ and will involve $\zeta$ as a difference).

---

[1]Note that even with a Gaussian distributed values of $\langle\sigma\rangle$ on an inflationary patch with a Gaussian width $H/(2\pi)$, there is about a 2/3 probability that such initial condition configurations can be found. Also, an unbroken discrete symmetry such as $\mathbb{Z}_2 : \sigma \to -\sigma$ can stabilize the VEV. In the context of supergravity, generic terms in the effective potential however can appear leading to $\langle\sigma\rangle \neq 0$ during inflation. In the end, whether or not $\langle\sigma\rangle = 0$ is model dependent, but it is not fine tuned.

Thus, the basic physics picture of the classical fluid that we are concerned with in this work is the following. To predict CMB temperature fluctuation $\langle \Delta T \Delta T \rangle$, we must compute the cross correlation $\langle \delta_S \zeta \rangle$ since at the linearized level, Einstein-Boltzmann equations give the relationship $\Delta T / T \sim c_1 \zeta + c_2 \delta_S$ for computable order unity (for long wavelengths) coefficients $c_i$. Up until this work, there has never been an explicit computation of the $\langle \delta_S \zeta \rangle / \sqrt{\langle \zeta \zeta \rangle \langle \delta_S \delta_S \rangle}$ coming from irreducible gravitational interactions.[2] What will emerge is a clean universal result that applies to a wide range of isocurvature models including those of the QCD axions (in a particular initial condition regime) and WIMPZILLAs. We find that $\langle \delta_S \zeta \rangle$ contribution is generically subdominant to $\langle \delta_S \delta_S \rangle$ in the case of pure gravitational interactions.

In the following, we establish our conventions in describing this isocurvature degree of freedom carrying the non-adiabatic initial condition information. In the process, we review the gauge invariant construction of these cosmological perturbations and the current CMB observational constraint, which represents the strongest constraint on the isocurvature initial condition derived from inflation.

## Gauge Invariant Construction

The cosmological inhomogeneity perturbation variables are generally spacetime coordinate gauge-dependent because of the coordinate dependent definition of fictitious background metric slices. From the perspective of matching classical equation initial conditions to inflationary quantum correlator computations, identifying gauge invariant combinations is helpful [233–235]. On the other hand, the gauge freedom involved in computing gauge invariant quantities facilitates the quantum computation. Hence, understanding the gauge dependences of the correlation computations is helpful. In

---

[2]As we will later explain, we do not compute $\langle \delta_S \zeta \rangle$ analytically fully beyond the time of the end of inflation. However, the importance of the isocurvature cross correlation can be generically predicted by $\langle \delta_S \zeta \rangle / \sqrt{\langle \zeta \zeta \rangle \langle \delta_S \delta_S \rangle}$ which is insensitive to the post-inflationary evolution for superhorizon modes.

this subsection, we review the gauge invariant variable construction and establish our notation. For a more general discussion, see for example [21, 89, 233, 234, 236–244].

In $(t, \vec{x})$ coordinates, we parameterize the metric as $g_{\mu\nu} = \bar{g}_{\mu\nu} + \delta g_{\mu\nu}^{(S)}$ where the scalar metric perturbation is

$$\delta g_{\mu\nu}^{(S)} = \begin{pmatrix} -E & aF_{,i} \\ aF_{,i} & a^2[A\delta_{ij} + B_{,ij}] \end{pmatrix}, \tag{4.2}$$

the background metric is $\bar{g}_{\mu\nu} \equiv \mathrm{diag}\{-1, a^2(t), a^2(t), a^2(t)\}$, and derivatives are denoted as usual as $X_{,i} \equiv \partial X/\partial x^i$. Under the diffeomorphism $x \to x + \epsilon$ where

$$\epsilon^{\mu} = (\epsilon^0, a^{-2}\partial_i(\epsilon^S)), \tag{4.3}$$

the scalar metric perturbation components transform as

$$\Delta A = -2H\epsilon^0, \quad \Delta B = -\frac{2}{a^2}\epsilon^S, \tag{4.4}$$

$$\Delta E = -2\dot{\epsilon}^0, \quad \Delta F = \frac{1}{a}(\epsilon^0 - \dot{\epsilon}^S + 2H\epsilon^S) \tag{4.5}$$

which is obtained from $\delta g_{\mu\nu}^{(S)} \to \delta g_{\mu\nu}^{(S)} + \Delta(\delta g_{\mu\nu}^{(S)})$ with $\Delta(\delta g_{\mu\nu}^{(S)}) = -\mathcal{L}_{\epsilon^{\mu}\partial_{\mu}}\bar{g}_{\mu\nu}$.

Similarly, we parameterize the perfect fluid stress tensor for a fluid element $a$ as

$$T_{\mu\nu}^{(a)} = \bar{T}_{\mu\nu}^{(a)} + \delta T_{\mu\nu}^{(a)} \tag{4.6}$$

where $\bar{T}_{\mu\nu}^{(a)} \equiv \mathrm{diag}\{\bar{\rho}_a, \bar{P}_a, \bar{P}_a, \bar{P}_a\}$ contains the average energy density and pressure seen by a comoving observer, $\delta T_{ij}^{(a)} = \bar{P}_a \delta g_{ij}^{(S)} + a^2\delta_{ij}\delta P_a$, $\delta T_{i0}^{(a)} = \bar{P}_{(a)}\delta g_{i0}^{(S)} - (\bar{\rho}_a + \bar{P}_a)\delta U_i^{(a)}$ (where $\delta U_i^{(a)}$ is the velocity perturbation), and $\delta T_{00}^{(a)} = -\bar{\rho}_a\delta g_{00}^{(S)} + \delta\rho_a$. Under the diffeomorphism of Eq. (4.3), the energy density perturbation transforms as

$$\Delta\delta\rho_a = -\epsilon^0\dot{\bar{\rho}}_a. \tag{4.7}$$

In practice, gauge-invariant variables are constructed by combining metric perturbations and other perturbations, such as densities. A popular choice is

$$\zeta_a \equiv \frac{A}{2} - H\frac{\delta\rho_a}{\dot{\bar{\rho}}_a}.$$

(4.8)

For example, the first-order gauge-invariant perturbation associated with the inflaton $\phi$ is usually defined as

$$\zeta_\phi \;\; \equiv \;\; \frac{A}{2} - H\frac{\delta\rho_\phi}{\dot{\bar{\rho}}_\phi}$$

(4.9)

(see for example Ref. [235] and references therein). Now, one can form a quantity that is conserved through reheating by defining

$$\zeta_{\text{tot}} \equiv \sum_i r_i \zeta_i$$

(4.10)

where

$$r_i \equiv \frac{\bar{\rho}_i + \bar{P}_i}{\sum_n \bar{\rho}_n + \bar{P}_n}.$$

(4.11)

Because there must be reheating dynamical degrees of freedom, $\zeta_{\text{tot}}$ must involve at least 2 degrees of freedom by the end of inflation of any single field slow-roll model. In single field slow-roll scenarios, what is done in practice is to argue that the reheating degrees of freedom are integrated out during inflation and then integrated back in at the end of inflation due to the different location of the inflaton VEV at the end of inflation. Alternatively, another often used assumption is that the main reheating degree of freedom are conformal such that no isocurvature fluctuations are appreciably excited during inflation. This means that in single field models, we have

$$\zeta_{\text{tot}} \approx \zeta_\phi$$

(4.12)

up to ambiguities in how one hides the reheating degrees of freedom.

One reason why the combination of Eq. (4.10) is convenient is because the superhorizon mode of this is approximately conserved through reheating if this mode object can be shown to obtain an initial conditions of what is sometimes referred to as the adiabatic solution [211, 235] and there are no non-adiabatic processes that mix superhorizon modes of isocurvature degrees of freedom with $\zeta_{\text{tot}}$. Such classical adiabatic solution initial conditions are generated by the Bunch-Davies quantum fluctuations for $\zeta_\phi$, and we will restrict the couplings of the isocurvature degrees of freedom (discussed below) such as to avoid non-adiabatic mixing. This means that Eq. (4.12) ensures that $\zeta_{\text{tot}}$ is approximately conserved if $\bar{\rho}_\phi + \bar{P}_\phi$ dominates over others. More explicitly, as discussed in the introduction to this section, suppose there exists only one isocurvature field degree of freedom which we call $\sigma$ during the inflationary period.[3] The total curvature perturbation can be written as

$$\zeta_{\text{tot}} = \zeta_\phi + r_\sigma(\zeta_\sigma - \zeta_\phi) \tag{4.13}$$

with the sum over $n$ runs over $\phi$ and $\sigma$ (assuming that $\psi$ has been integrated out during inflation). However, one can estimate that the coefficient of $\zeta_\sigma$ during inflation is

$$r_\sigma \lesssim \frac{1}{(2\pi)^2}\Delta_\zeta^2 \sim 10^{-11} \tag{4.14}$$

which makes the approximation of $\zeta_{\text{tot}} \approx \zeta_\phi$ accurate, just as in the single field case of Eq. (4.12). Thus just as in the single field scenarios without $\sigma$, $\zeta_{\text{tot}}$ acquires an approximately adiabatic boundary condition from the Bunch-Davies vacuum field fluctuations.

To complete the examination of how $\zeta_{\text{tot}}$ is used in the scenario of concern in this work, let's look at the time period surrounding the reheating transition when

---

[3]The species $\sigma$ will later be identified dark matter candidates such as the axions and WIMPZILLAs.

the universe reaches radiation domination. Near the time of the completion of the reheating, the variable $\zeta_{\text{tot}}$ is approximately

$$\zeta_{\text{tot}} \approx r_\phi \zeta_\phi + \sum_i r_{\psi_i} \zeta_{\psi_i} \tag{4.15}$$

such that after the inflaton decays, we have $r_\phi = 0$ and

$$\zeta_{\text{tot}} \approx \sum_i r_{\psi_i} \zeta_{\psi_i}. \tag{4.16}$$

[4] (The approximation used in Eq. (4.15) neglects the $r_\sigma$ contribution because of Eq. (4.14).) It is also a standard assumption that

$$\zeta_{\psi_i} = \zeta_{\text{tot}}, \tag{4.17}$$

which is rigorously true if one relativistic species dominate the fluid (e.g. $r_{\psi_1} \approx 1$) or if the decay process does not redistribute the spatial inhomogeneities of $\psi_i$ in a distinct configuration from that of $\phi$. [5] This justifies the usual statement in the literature that $\zeta_{\text{tot}}$ defined in Eq. (4.10) is primarily useful for arguing how a combination of quantities involving the inflaton and the reheating decay products remain unchanged through the reheating phase transition. Here, we have merely described how this argument is not changed by the presence of $\sigma$ because of the smallness of $r_\sigma$ in Eq. (4.14) during the primordial periods of interest.

In summary, as long as boundary conditions for the classical fluid equation are evaluated at a time when $r_\sigma$ is small (compared to the accuracy desired), we can neglect the $r_\sigma$ contribution from $\zeta_{\text{tot}}$ both through reheating and until the time that boundary conditions for the classical fluid equations are imposed. Hence, if $\zeta_{\text{tot}}$ remains constant on long wavelengths (due to the initial conditions set by the

---

[4]In the case that $\psi_i$ is integrated back in at the end of inflation, we have made the assumption that this does not change $\zeta_{\text{tot}}$

[5]However this need not be true for more general reheating scenarios.

Bunch-Davies vacuum), Eqs. (4.13) and (4.14) imply that the effective curvature perturbation during this early primordial epoch is given by Eq. (4.12). Hence, in the discussion below, we will drop the $\phi$ subscript and write

$$\zeta \equiv \zeta_\phi \approx \zeta_{\text{tot}}. \tag{4.18}$$

During this radiation dominated early primordial time $t_p$, the relationship between super horizon $A(t_p, \vec{k})$ and the value of $\zeta(t_e, \vec{k})$ evaluated at the end of inflation time $t_e$ is

$$\frac{A(t_p, \vec{k})}{2} \approx \frac{2}{3}\zeta(t_e, \vec{k}) \tag{4.19}$$

in the Newtonian gauge ($B = F = 0$) and the presence of $\zeta_\sigma$ gives a small error controlled by $r_\sigma$.

At the same radiation dominated era[6] when initial condition is set by $\zeta_{\text{tot}} \approx \zeta$, the inhomogeneity of the small mixture of dark matter component $\sigma$ can be related to the isocurvature perturbation $\zeta_\sigma$. Conventionally, this information is parameterized by the gauge-invariant isocurvature perturbation [3, 89, 232]

$$\delta_S(t, \vec{k}) \equiv 3\left(\zeta_\sigma(t, \vec{k}) - \zeta_{\text{tot}}(t, \vec{k})\right). \tag{4.20}$$

The physical interpretation of this quantity can be see by noting that when $\sigma$ particles are dominantly non-relativistic and the universe is radiation dominated, this expression becomes

$$\delta_S(t, \vec{k}) = \frac{\delta\rho_\sigma(t, \vec{k})}{\bar{\rho}_\sigma} - \frac{3}{4}\frac{\delta\rho_\gamma(t, \vec{k})}{\bar{\rho}_\gamma} \tag{4.21}$$

where $\rho_\gamma$ represents the photon energy densities. This clearly represents the difference in number densities of $\sigma$ and $\gamma$.[7] Assuming that the radiation inhomogeneity is

---

[6]During this time period, there is possibly a population of thermal dark matter components such as thermal WIMPs.

[7]It is interesting to note that since number densities can diverge while gravitational physics does not care about number densities (in favor of energy densities), this choice of variables is unfortunate in situations when there are IR divergences. In this work, we stick to this convention which is prevalent in literature.

characterized by $\zeta$ as explained in Eqs. (4.17) and (4.18) during radiation domination, we have

$$\delta_S(t, \vec{k}) \approx 3(\zeta_\sigma(t, \vec{k}) - \zeta(t_e, \vec{k})) \tag{4.22}$$

Similarly to the case of $\zeta_{\text{tot}}$, long wavelength limit of $\zeta_\sigma$ generated from Bunch-Davies initial conditions simplify (partly because of causality) in the absence of non-adiabatic processes mixing of $\zeta_\sigma$ with other superhorizon degrees of freedom. The $\zeta_\sigma$ mode for a comoving wave vector $\vec{k}$ becomes constant once $|\vec{k}/a| \ll H$ and $m_\sigma \ll H$ because the mode functions involved in $\zeta_\sigma$ are governed by the Hubble friction once these conditions are satisfied.

Although the key correlator computation result of this work involving $\beta$ evaluated at the end of inflation is independent of the transfer function evolving the isocurvature degrees of freedom after the end of inflation, because its immediate phenomenological application to CMB requires a transfer function describing this post-inflationary evolution, we will restrict our illustration in Section 4.4 to the situation when the chemical reaction rates that mix $\sigma$ and the radiation components are negligible. We will discuss in more detail the cross section constraint for this condition in Appendix 4.4.

## Observational Constraints on Isocurvature Perturbation

The current observational data shows that the CMB power spectrum is consistent with the adiabatic initial conditions. However, it does not rule out mixed boundary condition contributions from CDM isocurvature perturbations. Schematically, the temperature fluctuations depend linearly on $\zeta$ and $\delta_S$ initial conditions as

$$\frac{\Delta T}{T} = c_1 \zeta + c_2 \delta_S \tag{4.23}$$

where $c_i \sim O(1)$. Hence, the CMB temperature correlation data constrains

$$\frac{k^3}{2\pi^2} \int \frac{d^3p}{(2\pi)^3} \langle \frac{\Delta T(\vec{p})}{T} \frac{\Delta T^*(\vec{k})}{T} \rangle = \Delta_\zeta^2(k) \left[ |c_1|^2 + |c_2|^2 \frac{\alpha}{1-\alpha} - 2\Re \left( c_1^* c_2 \beta \sqrt{\frac{\alpha}{1-\alpha}} \right) \right]$$

(4.24)

where [97]

$$\int \frac{d^3p}{(2\pi)^3} \langle \zeta(\vec{p})\zeta^*(\vec{k}) \rangle = \Delta_\zeta^2(k)\frac{2\pi^2}{k^3}$$

(4.25)

$$\int \frac{d^3p}{(2\pi)^3} \langle \delta_S(\vec{p})\delta_S^*(\vec{k}) \rangle = \Delta_{\delta_S}^2(k)\frac{2\pi^2}{k^3}$$

(4.26)

$$\int \frac{d^3p}{(2\pi)^3} \langle \delta_S(\vec{p})\zeta^*(\vec{k}) \rangle = \Delta_{\zeta\delta_S}^2(k)\frac{2\pi^2}{k^3}$$

(4.27)

$$\alpha \equiv \frac{\Delta_{\delta_S}^2(k)}{\Delta_\zeta^2(k) + \Delta_{\delta_S}^2(k)},$$

(4.28)

$$\beta \equiv -\frac{\Delta_{\zeta\delta_S}^2(k)}{\sqrt{\Delta_\zeta^2(k)\Delta_{\delta_S}^2(k)}},$$

(4.29)

[8]which are customarily evaluated in the primordial epoch when $k$ corresponds to a far superhorizon scale such that the $\Delta_X^2(k)$ objects are constant in time. Typically the data constraints are parameterized by evaluating $\alpha$ and $\beta$ at a pivot scale $k = k_0$ [96, 200]. An important utility of this parameterization is the following fact: a necessary and sufficient condition for the cross correlation to be a significant part of the isocurvature contribution is to have $|\beta| \gtrsim |c_2/c_1|\sqrt{\alpha}$ for $\alpha < 1$. For example, in order to have approximately the same level of the angular power spectra from both pure isocurvature correlation and and cross-correlation at the intermediate scale $l \sim 200$, i.e. $C_l^{pure\,iso} \sim C_l^{cross\,cor}$, the fractional cross-correlation should satisfy $|\beta| \gtrsim 4 \times 10^{-2}$. Another utility of the $\beta$ variable comes from the fact that when there are non-trivial transfer functions governing $\Delta_{\zeta\delta_S}^2$ and $\Delta_{\delta_S}^2$ after the end of inflation, the

---

[8]Our sign conventions are such that *negative* values for $\beta$ correspond to a positive contribution of the cross-correlation term to the Sachs-Wolfe component of the total temperature spectrum. See, e.g., [96, 200].

transfer function factors can cancel in the expression for $\beta$. We will use this feature later to compute $\beta$ based on just the (quasi)-dS mode function behavior.[9]

As far as the experimental numbers are concerned, the isocurvature contribution to the CMB temperature perturbation is expected to be roughly less than 10% compared to the curvature contribution. More precisely, the Planck+WP limits [112, 201, 209] are

$$\alpha|_{\beta=0} < 0.016 \text{ (95\% CL) and } \alpha|_{\beta=-1} < 0.0011 \text{ (95\% CL)}, \qquad (4.30)$$

where the isocurvature power spectrum is assumed to be scale-invariant, i.e. $n_{iso} = 1$. The significant difference in the upper-bound of $\alpha$ between uncorrelated and totally (anti-)correlated cases can be explained by the ratio $\beta/\sqrt{\alpha}$ already discussed above. The difficulty in improving the current isocurvature bound with data on short wavelengths can be seen in Fig. 4.1, where one sees a fall-off of the isocurvature spectrum on short scales ($l \gtrsim 100$). This fall-off is generic and can be attributed to the transfer function effect encoded by $c_1(k)/c_2(k)$ in Eq. (4.24) for $k \gtrsim k_{eq}$ (where $k_{eq}/a_0 \sim 10^{-2}$ Mpc$^{-1}$ is the wave vector associated with matter radiation equality). To understand why $c_1(k)/c_2(k)$ generically becomes large for $k \gtrsim k_{eq}$, note that isocurvature modes with $k \gtrsim k_{eq}$ enter the horizon during radiation domination. Because the isocurvature effect on the temperature spectrum is gravitational, the value of $c_1(k)/c_2(k)$ is proportional to the ratio $\rho_R(t(k))/\rho_\sigma(t(k))$ of the radiation energy density to the energy density in the isocurvature degree of freedom at the time $t(k)$ when mode $k \gtrsim k_{eq}$ enters the horizon. Since shorter wavelengths enter the horizon earlier, $\rho_R(t(k))/\rho_\sigma(t(k))$ is larger for shorter wavelengths, making $c_1(k)/c_2(k)$ larger. For those readers not familiar with this physics, some of the details of the

---

[9]We will use the exact dS approximation for the massive $\sigma$ and use the quasi-dS approximation for only the massless scenario. The corrections coming from the the deviations away from the exact dS background in principle can be absorbed into the transfer function multiplying the superhorizon mode function which cancel out in $\beta$ due to a common appearance in the numerator and the denominator.
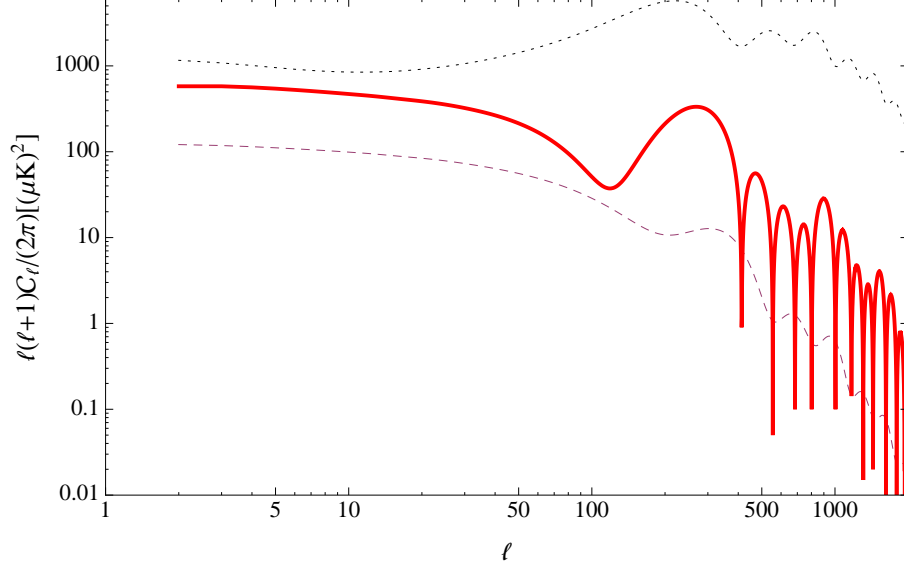
Figure 4.1: Angular power spectra $C_l$ from pure adiabatic(dotted), pure isocurvature(dashed), and cross-correlation(solid) contributions. The plotted pure adiabatic perturbation has the spectral index $n_s = 0.96$. For isocurvature perturbations, the spectral index $n_{iso}$ is 1 and the isocurvature fraction $\alpha = 0.067$ defined at $k_0 = 0.002 \, Mpc^{-1}$, and the fractional cross-correlation $\beta$ is 1.

transfer function are reviewed in Appendix 4.6.

Because of the large differences in the constraints between $\beta = 0$ and $\beta = -1$, estimating the cross-correlation is crucial to restrict parameters and give observable predictions of isocurvature models. In particular, the axion scenario with a negligible homogeneous vacuum misalignment angle (and similarly the WIMPZILLA scenario with a negligible homogenous background field value) predicts detectable non-Gaussianity [18,59,90]

$$f_{NL} \sim 30 \left( \frac{\alpha}{0.067} \right)^{3/2} \qquad (4.31)$$

provided the assumption the cross-correlation is zero, i.e. $\beta = 0$. However, as we will explain, this assumption is not obvious for massive field quadratic isocurvature scenarios, and the reexamination of this assumption is one of the goals of this work.

## 4.3   Computation of Correlators

In order to provide the initial condition of the classical fluid equations, it is standard to compute the quantum equal time correlators with the inflationary background approximated as a Bunch-Davies vacuum. In this section, we compute the correlators using the "in-in" formalism (e.g. see Weinberg [93]). More specifically, in the context of canonical quantization, we perturbatively compute the expectation value of an operator $\hat{Q}(t)$

$$
\begin{aligned}
\left\langle \hat{Q}(t) \right\rangle &= \sum_n (-i)^n \int_{-\infty}^t dt_1 \int_{-\infty}^{t_1} dt_2 \cdots \int_{-\infty}^{t_{n-1}} dt_n \\
&\quad \times \left\langle \left[ \left[ \left[ \hat{Q}^I(t), \hat{H}^I(t_n) \right], \hat{H}^I(t_{n-1}) \right], \cdots \hat{H}^I(t_1) \right] \right\rangle,
\end{aligned} \tag{4.32}
$$

where the superscript $I$ stands for the interaction picture and $\hat{Q}(t)$ represents a product of canonically quantized operators.

In the scenario explained in Sec. 4.2, we consider the gravitational coupling whose interaction Hamiltonian is derived from the ADM formalism with a given choice of gauge. For the computation of the cross-correlation to leading order in gravitational coupling, we need at least up to the cubic coupling $H^I_{\zeta\sigma\sigma}$, where $\sigma$ is a spectator field during inflation. The interaction Hamiltonian is diffeomorphism gauge-dependent. For two commonly used gauges, the comoving gauge($\delta\phi = 0$) and the uniform curvature gauge($A = 0$), we have

$$
H^I_{\zeta\sigma\sigma}(t) = -\frac{1}{2} \int d^3x\, a^3(t)\, T^{\mu\nu}_\sigma(t, \vec{x}) \delta g_{\mu\nu}(t, \vec{x}), \tag{4.33}
$$

$$
\delta g^{(C)}_{\mu\nu} = \begin{pmatrix} -2\frac{\dot{\zeta}}{H} & \left(-\frac{\zeta}{H} + \epsilon \frac{a^2}{\nabla^2} \dot{\zeta}\right)_{,i} \\ \left(-\frac{\zeta}{H} + \epsilon \frac{a^2}{\nabla^2} \dot{\zeta}\right)_{,i} & a^2 \delta_{ij} 2\zeta \end{pmatrix}, \tag{4.34}
$$

$$
\delta g^{(U)}_{\mu\nu} = \begin{pmatrix} 2\epsilon\zeta & \epsilon \frac{a^2}{\nabla^2} \dot{\zeta}_{,i} \\ \epsilon \frac{a^2}{\nabla^2} \dot{\zeta}_{,i} & 0 \end{pmatrix}, \tag{4.35}
$$

where $T_\sigma^{\mu\nu}$ is the stress energy tensor of the field $\sigma$, and $\delta g_{\mu\nu}$ is the metric perturbation and the superscript $(C)$ and $(U)$ denote the comoving gauge and uniform curvature gauge, respectively. A detailed derivation of the interaction Hamiltonian using the ADM formalism is presented in Section 4.8.

The isocurvature perturbation $\delta_S$ should be also written in terms of quantum operators associated with the energy density $\rho_\sigma$ of the particle $\sigma$. Since the energy density $\rho_\sigma$ is written in bilinear form of $\sigma$ and since the energy density of CDM are often those of non-relativistic particles at the time of matching to classical equations, we may approximate the energy density $\rho_\sigma \approx m_\sigma^2 \sigma^2$. We then promote field $\sigma$ to a quantum operator:

$$\delta_\sigma \equiv \frac{\delta\rho_\sigma}{\rho_\sigma} \approx \frac{\sigma^2 - \bar{\sigma}^2}{\bar{\sigma}^2} \rightarrow \hat{\delta}_\sigma = \frac{\hat{\sigma}^2 - \langle\hat{\sigma}^2\rangle}{\langle\hat{\sigma}^2\rangle}. \tag{4.36}$$

The field $\hat{\sigma}$ can be decomposed into the classical homogeneous background and the quantized perturbation, i.e. $\hat{\sigma} = \bar{\sigma} + \delta\hat{\sigma}$. Unlike the inflaton $\phi$ whose classical background is non-zero, because we consider the field $\hat{\sigma}$ without classical background, the leading density perturbation starts with the quadratic in the operator $\delta\hat{\sigma}^2$. As with any quantum composite operator, we renormalize it with counter terms invariant under the underlying gauge symmetry (here, it is diffeomorphism):

$$\left(\hat{\sigma}^2\right)_r = \left(\delta\hat{\sigma} + \sum_i \hat{\chi}_i\right)^2 + \delta Z_0 + \delta Z_1 R, \tag{4.37}$$

where the subscript $r$ denotes that the operator is a renormalized composite operator, $R$ is the Ricci scalar, and $\hat{\chi}_i$ are Pauli-Villars fields, which is described in Section 4.9. We apply this to gauge-invariant isocurvature variable $\delta_S$ defined in Section 4.2. Then we have

$$\hat{\delta_S}^{(C)} = -\frac{3H}{\partial_t\langle(\hat{\sigma}^2)_r\rangle}\left[(\hat{\sigma}^2)_r - \langle(\hat{\sigma}^2)_r\rangle\right], \tag{4.38}$$

$$\hat{\delta_S}^{(U)} = -\frac{3H}{\partial_t\langle(\hat{\sigma}^2)_r\rangle}\left[(\hat{\sigma}^2)_r - \langle(\hat{\sigma}^2)_r\rangle\right] - 3\hat{\zeta}. \tag{4.39}$$

We will not write the hat explicitly from now on.

In the next subsection, we present how a non-diffeomorphism-invariant estimation of the cross-correlation leads to an observationally attractive but grossly incorrect result. In subsections after that, we identify the problems with the wrong estimate and calculate the cross-correlation properly.

## Plausible but Wrong Estimation of the Cross-Correlation

In this subsection, we present a plausible estimation of the cross correlation that leads to a large value that is observationally interesting. Unfortunately, we will see in later subsections that the estimate presented in this subsection can be many orders of magnitude off due to the explicit breaking of diffeomorphism invariance in the treatment of the UV physics. Nonetheless, what is presented in this subsection is interesting both as a lesson in field theory and as a motivation for the careful correct computation that follows later.

The isocurvature cross-correlation in the comoving gauge is written as

$$\left\langle \delta_S^{(C)} \zeta \right\rangle \approx \frac{\langle (\sigma^2)_r \, \zeta \rangle}{\langle (\sigma^2)_r \rangle}, \tag{4.40}$$

where we have used $\partial_t \langle (\sigma^2)_r \rangle + 3H \langle (\sigma^2)_r \rangle \approx 0$ for the isocurvature field number density. For an order of magnitude estimation, we consider a non-derivatively coupled part of the gravitational interaction, $2\zeta a^2 \delta_{ij} T_\sigma^{ij} \in H_{\zeta\sigma\sigma}^I$. Then the two-point function,

shown diagrammatically in Fig. 4.2, is written in the Fourier space as

$$
\widetilde{\langle (\sigma^2)_r \, \zeta \rangle}^C_p
$$

$$
\sim \int d^3 x \, e^{-i\vec{p}\cdot\vec{x}} \int^t d^4 z \, a^3(t_z) \left\langle \left[ \sigma^2(t,\vec{x})\zeta(t,\vec{0}), \frac{i}{2}\left( 2\zeta a^2 \delta_{ij}T_\sigma^{ij} \right)_z \right] \right\rangle \quad (4.41)
$$

$$
\sim -4 \int \frac{d^3 k_1}{(2\pi)^3} d^3 k_2 \delta^3(\vec{k}_1 + \vec{k}_2 - \vec{p}) \int_{-\infty}^t dt_z \, a_z^3
$$

$$
\times \mathrm{Im}\left[ \zeta_p(t)\zeta_p^*(t_z) u_{k_1}(t) u_{k_2}(t) \right.
$$

$$
\times \left. \left\{ \frac{1}{2}\frac{\vec{k}_1 \cdot \vec{k}_2}{a^2} + 3\left( \frac{1}{2}\partial_t^{(1)}\partial_t^{(2)} - \frac{1}{2}m_\sigma^2 \right) \right\} u_{k_1}^*(t_z) u_{k_2}^*(t_z) \right] \quad (4.42)
$$

where

$$
\widetilde{\langle AB \rangle}_p \equiv \int d^3 x \, e^{-i\vec{p}\cdot\vec{x}} \left\langle A(t,\vec{x})B(t,0) \right\rangle, \quad (4.43)
$$

$\zeta_p$ and $u_k$ are mode functions for $\zeta$ and $\sigma$, respectively, and $\partial_t^{(i)}$ means the time derivative with respect to $u_{k_i}^*(t_z)$.[10] This integral is UV divergent, and thus we introduce the horizon scale UV cut-off

$$
\Lambda_{UV} \sim aH_{inf}. \quad (4.44)
$$

Moreover, we neglect the contribution from the time range $t < t_p$, where $t_p$ is the time when the scale $p$ exits the horizon since $\zeta_p$ is oscillatory before the horizon exit. Using the super-horizon approximation for mode functions during inflation

$$
\zeta_k(t) \;=\; \frac{1}{\sqrt{4\epsilon}M_p}\frac{H}{k^{\frac{3}{2}}}e^{i\frac{k}{aH}}(1 - i\frac{k}{aH}), \quad (4.45)
$$

$$
u_k(t) \;\approx\; a^{-\frac{3}{2}}H^{-\frac{1}{2}}\left\{ \frac{2^{\nu-1}\Gamma(\nu)}{\pi^{\frac{1}{2}}}\left( \frac{k}{aH} \right)^{-\nu} + i\frac{\pi^{\frac{1}{2}}}{2^{\nu+1}\Gamma(1+\nu)}\left( \frac{k}{aH} \right)^{\nu} \right\}, \quad (4.46)
$$

---

[10] It is also helpful to remember that in terms of Fourier space operators/fields, the tilde notation is equivalent to

$$
\widetilde{\langle AB \rangle}_p = \int \frac{d^3 p_2}{(2\pi)^3}\left\langle A(t,\vec{p})B(t,\vec{p}_2) \right\rangle
$$

where

$$
A(t,\vec{p}) \equiv \int d^3 x \, e^{-i\vec{p}\cdot\vec{x}} A(t,\vec{x})
$$

for generic operators/fields $A$ and $B$.

where $\nu \equiv \sqrt{9/4 - m^2/H^2}$, the cross-correlation at the end of inflation time $t_e$ is approximately

$$\widetilde{\langle(\sigma^2)_r \zeta\rangle}_p^C \sim \frac{-1}{8\pi^2}\left|\zeta_p^o\right|^2 \frac{H^4}{m_\sigma^2}\left[1 - \left(\frac{p}{a_e H}\right)^{\frac{2m_\sigma^2}{3H^2}}\right] \tag{4.47}$$

where we used the relations $m_\sigma^2 \ll H^2$ and $\left|\zeta_p^o\right|^2 p^3 = H^2/4M_p^2\epsilon$ is the mode function behavior in the long wavelength limit. To understand the magnitude of this expression, note that for physical CMB scale comoving momenta, we have

$$\frac{p}{a_e} = e^{-N(p)}H \tag{4.48}$$

for $N(p) \sim O(50)$. As long as

$$1 \gg m_\sigma^2/H^2 \gtrsim 1/N(p), \tag{4.49}$$

we can estimate

$$\widetilde{\langle(\sigma^2)_r \zeta\rangle}_p^C \sim \frac{-1}{8\pi^2}\left|\zeta_p^o\right|^2 \frac{H^4}{m_\sigma^2} \tag{4.50}$$

which is an expression that is valid when the $p$ is far outside of the horizon and a constant $H$ is a good approximation. Note that this does not vanish in the limit $p \to 0$. We will soon see that this non-vanishing behavior is incorrect and is a signal of explicit breaking diffeomorphism invariance coming from Eq. (4.44). Note that if Eq. (4.49) is not satisfied because $m_\sigma = 0$, we have

$$\widetilde{\langle(\sigma^2)_r \zeta\rangle}_p^C \sim \frac{H^2}{12\pi^2}\left|\zeta_p^o\right|^2 \ln\frac{p}{a_e H} \tag{4.51}$$

$$\sim -N(p)\frac{H^2}{12\pi^2}\left|\zeta_p^o\right|^2 \tag{4.52}$$

which again does not vanish and is negative.

As explained around Eq. (4.24), the importance of the cross-correlation in the isocurvature bound depends on whether $\beta$ is of order $10^{-2}$ or larger and not by whether the cross correlation by itself is of the order of curvature perturbations. To

compute $\beta$ defined in Eq. (4.29), we need an estimate of $(\sigma^2)_r$ correlator which we can take from [18]:

$$\langle \widetilde{(\sigma^2)_r \, (\sigma^2)_r} \rangle_p^C \sim \frac{1}{2\pi^2} \frac{H^4}{p^3} f(m_\sigma/H, p/a_e H) \tag{4.53}$$

where $f$ is a function which can have an exponentially small value owing to the functional behavior

$$f \sim \frac{H^2}{m_\sigma^2} \left(\frac{p}{a_e H}\right)^{\frac{4}{3}\frac{m_\sigma^2}{H^2}}. \tag{4.54}$$

Combining Eqs. (4.29), (4.47), and (4.53), we find

$$\beta^{\text{wrong}} \sim \sqrt{\Delta_\zeta^2} \frac{H}{4m_\sigma} \left(\frac{p}{a_e H}\right)^{-\frac{2}{3}\frac{m_\sigma^2}{H^2}} \tag{4.55}$$

$$\sim \frac{H}{4m_\sigma} e^{\frac{2}{3}\frac{m_\sigma^2}{H^2}N - 12} \tag{4.56}$$

which after recalling that $N \sim O(50)$ and Eq. (4.49) gives some hope that a proper computation would give a large value for $\beta$ with $m_\sigma/H$ satisfying Eq. (4.49).[11] For example, if $|\beta| = O(1)$, then any appreciable isocurvature perturbation would be ruled out with the current data, affecting predictions of [18,59,90].

Recall from Eq. (4.24) that the role of the cross correlation can become important if $\beta$ can become sizable while keeping $\alpha$ also sizable. One may worry that the enhancement factor in $\beta$ of Eq. (4.53) which is approximately proportional to $\alpha$ may make $\alpha$ negligible in the parameter regime in which $\beta$ is enhanced. However, note that $\alpha$ is controlled not just by Eq. (4.53) but by

$$\widetilde{\langle \delta_S \delta_S \rangle}_p = \frac{\langle \widetilde{(\sigma^2)_r \, (\sigma^2)_r} \rangle_p^C}{[\langle (\sigma^2)_r \rangle]^2} \tag{4.57}$$

which has a one point function squared in the denominator proportional to the energy density squared of $\sigma$. One can straight forwardly check from Ref. [18] that

---

[11]It is important to keep in mind that we are making an assumption here about the isocurvature evolution when identifying the primordial computations of Eqs. (4.47) and (4.53) with the CMB observables of Eq. (4.29) where $c_i$ are computed according to the simple transfer treatment of Appendix 4.6. We will discuss this assumption more in detail in subsection 4.4.

the denominator of Eq. (4.57) can be tuned such that $\alpha$ can remain constant while $\langle \widetilde{(\sigma^2)_r} \, (\sigma^2)_r \rangle_p^C$ is sufficiently small as to enhance $\beta$ as described in Eq. (4.56).

Given this generic possibility of ruling out a large class of isocurvature perturbation models, we consider below the leading gravitational interaction contribution to $\beta$ carefully. We find that unlike the naive estimate given in Eq. (4.47), there is a suppression in the limit $p/(aH) \to 0$ for the mass in the range of Eq. (4.49). The suppression in the numerator of $\beta$ precisely cancels the denominator suppression factor coming from $f$ in Eq. (4.54) such that no enhancement is obtained, contrary to the naive expectation of Eq. (4.56). This suppression of the numerator in the proper computation not seen in the naive estimate can be attributed to a Ward identity associated with the diffeomorphism group element of constant scaling of the spatial coordinates. Furthermore, a careful computation that we give below will show that the sign of the cross-correlation will be opposite to the naive estimate, owing to the fact that the cross correlation here is tied to particle production instead of volume dilution.

The detailed computation will address also explicitly how same answer to the gauge invariant correlator results in two different gauges of comoving gauge and uniform curvature gauge (one can verify this is not obvious from the naive estimate presented in this subsection). Another technical care that is taken in the computations below is to explicitly specify how diffeomorphism invariant counter terms are introduced to renormalize the composite operators intrinsic to $\delta_S$. Since the correct answer relies on a gravitational Ward identity, identifying proper diffeomorphism invariant regulator and counter terms is important for a trustworthy computation. On the other hand, note that the finite parts of the counter terms that remain after the divergences are canceled will not affect the results to the leading $\hbar$ expansion that we

are concerned with.[12]

## Plausible and Correct Estimation Using a Soft-$\zeta$ Theorem

Before we describe the actual computation, we give in this subsection a method akin to the soft-$\zeta$ theorem used by [19, 212–217, 219–225] to estimate the correct answer without a detailed computation. We will also point out what ad-hoc assumptions are needed to make this estimate using this theorem. A rigorous computation will be given in subsection (4.3).

In the soft-$\zeta$ theorem application to the correlators in inflation, one factorizes $N$-point function including at least one soft external $\zeta$ into $(N-1)$-point function times the two point function $\langle \zeta\zeta \rangle$. The well-known example is the three-point function $\langle \zeta\zeta\zeta \rangle$ in the squeezed limit in quasi-dS space:

$$\int \frac{d^3q}{(2\pi)^3} \langle \zeta_{\vec{q}}\zeta_{\vec{k}}\zeta_{\vec{p}} \rangle \xrightarrow{p\to 0} -\left|\zeta_p^o\right|^2 \frac{1}{k^3}\frac{\partial}{\partial \ln k}\left[k^3 \widetilde{\langle \zeta\zeta \rangle_k}\right] \sim -(n_s-1)\left|\zeta_p^o\right|^2\left|\zeta_k^o\right|^2 \quad (4.58)$$

where the superscript on the $\zeta$ mode functions denote long wavelength parts. To use this, note that if we neglect renormalization of the composite operators, we can write

$$\int \frac{d^3q}{(2\pi)^3}\langle \zeta_{\vec{p}}\sigma^2(\vec{q}) \rangle = \int \frac{d^3k_2}{(2\pi)^3}\int \frac{d^3k_1}{(2\pi)^3}\langle \zeta_{\vec{p}}\sigma(\vec{k}_1)\sigma(\vec{k}_2) \rangle. \quad (4.59)$$

Using Eq. (4.58) and replacing two $\zeta$ fields with $\sigma$ fields, we can estimate

$$\int \frac{d^3q}{(2\pi)^3}\langle \zeta_{\vec{p}}\sigma^2(\vec{q}) \rangle \xrightarrow{p\to 0} -\left|\zeta_p^o\right|^2 \int_p \frac{d^3k_2}{(2\pi)^3}\frac{1}{k_2^3}\frac{\partial}{\partial \ln k_2}\left[k_2^3 \widetilde{\langle \sigma\sigma \rangle_{k_2}}\right] \quad (4.60)$$

where the comoving IR cutoff $p$ is required to treat $\zeta_p^o$ as a constant background field. This effective lower cutoff $p$ cannot be justified without explicit computation, but this is physically plausible because $\langle \sigma\sigma \rangle$ does not have any IR divergence as long as $m_\sigma^2 > 0$. One can rewrite the integral in Eq. (4.60) as

$$\int \frac{d^3q}{(2\pi)^3}\left\langle \zeta_{\vec{p}}\sigma^2(\vec{q}) \right\rangle \xrightarrow{p\to 0} \left|\zeta_p^o\right|^2 \frac{\partial}{\partial \ln a}\left\langle \sigma^2(t,\vec{x}) \right\rangle_p \quad (4.61)$$

---

[12]Note that particle production is non-perturbative in $\hbar$.

where the $\sigma^2$ on the right hand side corresponds to spacetime field (and not its Fourier transform), the $p$ subscript on the bracket corresponds to the IR cutoff in the mode function integral, and we assume that there is no contribution from the UV cutoff. It is easy to prove that if $p \to 0$ is well defined and a UV cutoff is not required, then the right hand side of Eq. (4.61) vanishes in the limit $p \to 0$. This is in contrast with Eq. (4.47).

The vanishing of this function in the $p \to 0$ limit for $m_\sigma^2 > 0$ is intuitively understood from the fact that in that limit, $\zeta_p^o$ acts as a spatial diffeomorphism

$$\vec{x} \to \vec{x}(1 + \zeta_p^0) \tag{4.62}$$

(which in turn effectively rescales the scale factor $a$ by a constant factor if we neglect spatial derivatives on long wavelengths) which cannot change $\langle \sigma^2(t, \vec{x}) \rangle = \langle \sigma^2(t, 0) \rangle$. More explicitly, one can show that the explicit computation can be rewritten as

$$\int \frac{d^3 q}{(2\pi)^3} \left\langle \zeta_{\vec{p}} \sigma^2(\vec{q}) \right\rangle \xrightarrow{p \to 0} |\zeta_p|^2 \int_p \frac{d^3 k}{(2\pi)^3} \int d^3 x i \langle [\hat{Q}(t), \hat{\sigma}(t, \vec{x}) \hat{\sigma}(t, 0)] \rangle e^{i\vec{k} \cdot \vec{x}} \tag{4.63}$$

where

$$\hat{Q}(t) \equiv \int^t d^4 z a^2(t_z) \delta_{ij} T_\sigma^{ij}(z) \tag{4.64}$$

is the generator of the diffeomorphism associated with Eq. (4.62). Note that the right hand side formally vanishes when the IR cutoff is removed (i.e. $p = 0$) because in that limit, we find the commutator

$$\langle [\hat{Q}(t), \hat{\sigma}^2(t, 0)] \rangle = 0. \tag{4.65}$$

This can be interpreted also as a Ward identity. On the flip side, as long as $p \neq 0$, $\langle \sigma^2(t, \vec{x}) \rangle_p$ is not invariant under the diffeomorphism Eq. (4.62). The crucial point from this perspective is that diffeomorphism invariance is extremely important to see that the cross correlation vanishes for $p \to 0$ for a massive scalar field. It is this that one failed to preserve in Eq. (4.44).

As we will show in detail, Eq. (4.61) is consistent with the explicit computation. Note that a couple of assumptions that we already mentioned in deriving Eq. (4.61) can only be justified by an explicit computation: namely, the effective lower cutoff $p$ in Eq. (4.60) and UV cutoff details associated with renormalizing the composite operator $\sigma^2$. Such complications do not arise in isocurvature scenarios without composite operators. Hence, one of the main technical merits of this work is to provide a explicit justification of Eq. (4.61). Note that because the diffeomorphism gauge invariance plays a crucial role in obtaining the correct $p$ dependence in Eq. (4.61) as explained around Eq. (4.65), we choose a UV regulator that preserves diffeomorphism invariance in the computation below.

## Gauge Invariance of Correlators

Before we begin our explicit computation, we will check the setup of our computation by demonstrating that the manifestly gauge invariant quantities $\langle \delta_S \zeta \rangle$ and $\langle \delta_S \delta_S \rangle$ yield the same values in comoving and in the uniform curvature gauges. To accomplish this, we use a gravitational Ward identity.

We first note that the $\zeta$ dependent metric perturbations $\delta g^{(C)}$ and $\delta g^{(U)}$ differs by a gauge transformation, i.e.

$$\Delta g_{\mu\nu} = \delta g_{\mu\nu}^{(U)} - \delta g_{\mu\nu}^{(C)} = \begin{pmatrix} 2\frac{d}{dt}(\frac{\zeta}{H}) & (-\frac{\zeta}{H})_{,i} \\ (-\frac{\zeta}{H})_{,i} & -a^2 \delta_{ij} 2\zeta \end{pmatrix} = -[\mathcal{L}_X \bar{g}]_{\mu\nu}, \qquad (4.66)$$

where

$$X^0 = -\frac{\zeta}{H}, \quad X^i = 0. \qquad (4.67)$$

Their interaction actions differ by

$$\Delta S_{\sigma\sigma\zeta} = S_{\sigma\sigma\zeta}^{(U)} - S_{\sigma\sigma\zeta}^{(C)} = -\int^{t_f} dt d^3x a_x^3 \, T^{\mu\nu}(\bar{g}, \sigma) \nabla_\mu X_\nu \qquad (4.68)$$

Their interaction Hamiltonians differ by

$$\Delta H_{\zeta\sigma\sigma}(t) = H_{\zeta\sigma\sigma}^{(U)}(t) - H_{\zeta\sigma\sigma}^{(C)}(t) = \int d^3x \, a^3(t) T^{\mu\nu}(\bar{g}, \sigma; t, \vec{x}) \nabla_\mu X_\nu(t, \vec{x}) \quad (4.69)$$

Then we compare $\langle \sigma_x^2 \zeta_y \rangle$ in the two gauges:

$$\langle \sigma^2(t_f, \vec{x}) \zeta(t_f, \vec{y}) \rangle^U - \langle \sigma^2(t_f, \vec{x}) \zeta(t_f, \vec{y}) \rangle^C \quad (4.70)$$

$$= -i \int^{t_f} dt \, \langle \left[ \sigma_x^2 \zeta_y, \Delta H_{\zeta\sigma\sigma}(t) \right] \rangle \quad (4.71)$$

$$= -i \int^{t_f} dt d^3z \, \langle \left[ \sigma_x^2 \zeta_y, \nabla_\mu \left( a^3(t) T^{\mu\nu}(\bar{g}, \sigma; t, \vec{x}) X_\nu(t, \vec{x}) \right) \right] \rangle \quad (4.72)$$

where we have integrated by parts and used the quantum version of $\nabla_\mu T_\sigma^{\mu\nu} = 0$: i.e. in-in formalism gravitational Ward identities

$$i\nabla_\mu \langle in | T_z^{\mu\nu+} \sigma_x^+ \sigma_y^+ | in \rangle_g = \frac{1}{\sqrt{g_x}} \delta^4(x - z) g_x^{\alpha\nu} \frac{\partial}{\partial x^\alpha} \langle in | \sigma_x^+ \sigma_y^+ | in \rangle_g$$

$$+ \frac{1}{\sqrt{g_y}} \delta^4(y - z) g_y^{\alpha\nu} \frac{\partial}{\partial y^\alpha} \langle in | \sigma_x^+ \sigma_y^+ | in \rangle_g \quad (4.73)$$

$$i\nabla_\mu \langle in | T_z^{\mu\nu-} \sigma_x^+ \sigma_y^+ | in \rangle_g = 0 \quad (4.74)$$

whose the notation is explained in Section 4.7. Note that the remaining term in Eq. (4.72) is a total derivative. Hence, we are left with the boundary contribution

$$\langle \sigma^2(t_f, \vec{x}) \zeta(t_f, \vec{y}) \rangle^U - \langle \sigma^2(t_f, \vec{x}) \zeta(t_f, \vec{y}) \rangle^C \quad (4.75)$$

$$= -i \int d^3z \, a^3(t_f) \frac{1}{H} \langle \left[ \sigma_x^2, T_{\sigma,z}^{00} \right] \rangle \langle \zeta_z \zeta_y \rangle \quad (4.76)$$

$$= -\frac{\partial_t \langle \sigma_x^2 \rangle}{H} \langle \zeta_x \zeta_y \rangle. \quad (4.77)$$

To make these composite operator correlators well defined while maintaining diffeomorphism invariance (see the discussion surrounding Eq. (4.65)), we need a proper covariant regulator, such as the Pauli-Villars (PV) regulator. It is straightforward to use the PV regulator here because the above identity holds for PV fields as well. See Appendix 4.9 for a more detailed discussion of the prescription of the PV regulator.

Using Eq. (4.77), it is now trivial to show that $\langle \delta_S \zeta \rangle^U = \langle \delta_S \zeta \rangle^C$ and $\langle \delta_S \delta_S \rangle^U = \langle \delta_S \delta_S \rangle^C$. Because $\delta_S \ni \sigma_x^2 / \langle \sigma_x^2 \rangle$, the denominator of this expression also transforms:

$$\Delta \delta_S \ni -\frac{\Delta \langle \sigma_x^2 \rangle}{\langle \sigma_x^2 \rangle} \frac{\sigma_x^2}{\langle \sigma_x^2 \rangle} = \frac{\zeta_x}{H} \frac{\partial_t \langle \sigma_x^2 \rangle}{\langle \sigma_x^2 \rangle} \frac{\sigma_x^2}{\langle \sigma_x^2 \rangle} \tag{4.78}$$

which leads to a cancellation of Eq. (4.77) consistently to leading $\hbar \to 0$ approximation. Hence, we have a nontrivial consistency check of the computation that we are setting up.

## Two-point Functions

In this subsection, we present a rigorous computation of $\beta$ defined in (4.29). To this end, we need to calculate the two-point function $\langle (\sigma^2)_r \zeta \rangle$ and $\langle (\sigma^2)_r (\sigma^2)_r \rangle$ where the renormalized composite operator [3, 5, 6, 245–254] is

$$(\sigma^2)_r \equiv (\sigma + \sum_n \chi_n)^2 + \delta Z_0(\Lambda, m_\sigma) + \delta Z_1(\Lambda, m_\sigma) R \tag{4.79}$$

which is discussed in greater detail in Sec. (4.9). Here we are going to use the comoving gauge for the computation because of its advantages that we state below.[13] As shown in Eqs. (4.33) and (4.34), the gravitational interactions in the comoving gauge are derivatively (i.e. $p^2/a^2$) suppressed except the $(ij)$-components. In other words, the contributions from $T_\sigma^{00} \delta g_{00}^{(C)}$ and $T_\sigma^{0i} \delta g_{0i}^{(C)}$ interactions are $O(p^2/a^2)$, where $\vec{p}$ is an external 3-momentum. Furthermore, all counter term contributions are also derivatively suppressed in the comoving gauge: $\delta Z_0 \langle \zeta \rangle = 0$ and $\delta Z_1 \widetilde{\langle R\zeta \rangle}_p^C = O(p^2/a^2)$. Therefore, we don't need the counter terms to compute the non-derivatively suppressed contributions, but we still need a regulator for UV divergences in the computation. The

---

[13]This computation has been done also in the uniform curvature gauge, which is presented in Appendix 4.10. Particularly, in the massless limit, we explicitly calculate up to the next leading term including all gravitational couplings. This shows that the next leading terms are indeed suppressed by the factor $p^2/a^2$.
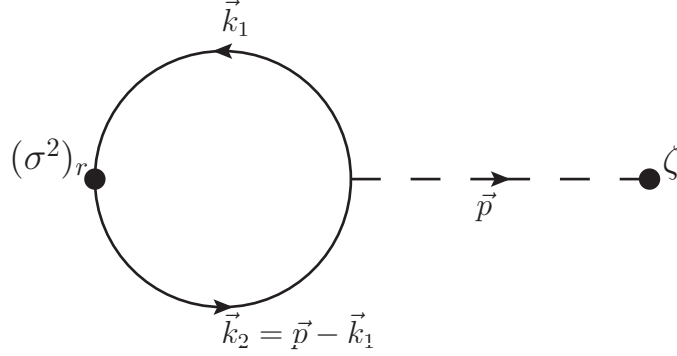
Figure 4.2: Two-point function at one loop order.

regulator dependences and the UV divergences will automatically disappear together in our final result.

Now we compute the two-point function shown in Fig. 4.2, which is written in the Fourier space as

$$
\widetilde{\langle (\sigma^2)_r \zeta \rangle}_p^C \;=\; \int d^3x\, e^{-i\vec{p}\cdot\vec{x}} \left\langle \left( \sigma^2(t,\vec{x}) \right)_r \zeta(t,\vec{0}) \right\rangle^C \tag{4.80}
$$

$$
=\; \int d^3x\, e^{-i\vec{p}\cdot\vec{x}} \int^t d^4z\, a^3(t_z)
$$

$$
\times \sum_{N=0}^{s} \left\langle \left[ \sigma_N^2(t,\vec{x})\zeta(t,\vec{0}),\, \frac{i}{2}\left( 2\zeta a^2 \delta_{ij} T_\sigma^{ij} \right)_z \right] \right\rangle + O\left( \frac{p^2}{a^2} \right), \tag{4.81}
$$

where we have introduced the Pauli-Villars (PV) regulator (see Appendix 4.9 for more details) and

$$
a^2 \delta_{ij} T_\sigma^{ij} \;=\; -3\mathcal{L}_\sigma + \sum_{N=0}^{s} C_N \left( \frac{\nabla}{a}\sigma_N \right)^2, \tag{4.82}
$$

where $\sigma_0$ and $\sigma_n$ are the physical field $\sigma$ and the PV field $\chi_n$ (here, $n \in \{1, 2, ..., s\}$), respectively, and $s$ is the number of introduced PV fields.

Interestingly, this integral can be computed in any FRW space-time. We first compute the second term contribution in Eq. (4.82) defined as

$$
I_N^{(2)}(p) \equiv \int d^3x\, e^{-i\vec{p}\cdot\vec{x}} \int^t d^4z \sqrt{-g_z} \left\langle \left[ \sigma_N^2(t,\vec{x})\zeta(t,\vec{0}),\, i\zeta_z C_N \left( \frac{\nabla}{a}\sigma_N \right)^2 \right]_z \right\rangle \tag{4.83}
$$

Expanding in mode functions, this becomes

$$
\begin{aligned}
I_N^{(2)}(p) &= -4C_N^{-1} \int \frac{d^3k_1}{(2\pi)^3} d^3k_2 \delta^3(\vec{k}_1 + \vec{k}_2 - \vec{p}) \int_{-\infty}^t dt_z a_z^3 \left(-\frac{\vec{k}_1 \cdot \vec{k}_2}{a_z^2}\right) \\
&\quad \times \mathrm{Im}\left[\zeta_p(t)\zeta_p^*(t_z) u_{N,k_1}(t) u_{N,k_1}^*(t_z) u_{N,k_2}(t) u_{N,k_2}^*(t_z)\right],
\end{aligned} \tag{4.84}
$$

where $u_N$ are the mode functions for fields $\sigma_N$. Because $\zeta$ oscillates before and freezes after the horizon exit, we neglect the contribution before the horizon exit. Furthermore, we can neglect the $O(p^2/a^2)$ term and factor $\zeta_p$ out of the time integral. We thus find

$$
\begin{aligned}
I^{(2)}(p) &\approx 4\left|\zeta_p^o(t)\right|^2 \int \frac{d^3k_1}{(2\pi)^3} d^3k_2 \delta^3(\vec{k}_1 + \vec{k}_2 - \vec{p}) \\
&\quad \times \int_{t_p}^t dt_z a_z^3 \left(\frac{\vec{k}_1 \cdot \vec{k}_2}{a_z^2}\right) \mathrm{Im}\left[u_{k_1}(t) u_{k_1}^*(t_z) u_{k_2}(t) u_{k_2}^*(t_z)\right] + O\left(\frac{p^2}{a^2}\right)
\end{aligned} \tag{4.85}
$$

where $t_p$ is the time at which scale $p$ exits the horizon. Note that we drop subscript $N$ and field normalization $C_N$ for convenience, but we will put it back later in the final result. Moreover, we neglect the low momentum phase space, i.e. $min\{k_1, k_2\} < p$, because of $|u_k|^2 \lesssim O(k^{-3})$ and the spatial gradient factor $\vec{k}_1 \cdot \vec{k}_2/a^2$.

$$
\begin{aligned}
\int_{k_1<p} \frac{d^3k_1}{(2\pi)^3} d^3k_2 \delta^3(\vec{k}_1 + \vec{k}_2 - \vec{p}) \int_{t_p}^t dt_z a_z^3 \left(\frac{\vec{k}_1 \cdot \vec{k}_2}{a_z^2}\right) \\
\times \mathrm{Im}\left[u_{k_1}(t) u_{k_1}^*(t_z) u_{k_2}(t) u_{k_2}^*(t_z)\right] \lesssim O\left(\frac{p^2}{a^2}\right).
\end{aligned} \tag{4.86}
$$

Then the main contribution of the integral comes from the phase space $k_1, k_2 > p$, and thus $p$ behaves as an IR cut-off (see the importance of this IR cutoff in the discussion surrounding Eq. (4.61)).

Since $k_1, k_2 > p$, we Taylor-expand the integrand with respect to $p$ and take the leading term. Then we have

$$
I^{(2)}(p) \approx 4\left|\zeta_p^o(t)\right|^2 \int_p \frac{d^3k_1}{(2\pi)^3} \int_{t_p}^t dt_z a_z^3 \left(-\frac{k_1^2}{a_z^2}\right) \mathrm{Im}\left[u_{k_1}^2(t) u_{k_1}^{*2}(t_z)\right] + O\left(\frac{p^2}{a^2}\right). \tag{4.87}
$$

Now we are going to compute the time integral. Recall that the differential equation for mode function $u_k$ is

$$\ddot{u}_k + 3Hu_k + \left(\frac{k^2}{a^2} + m^2\right)u_k = 0. \tag{4.88}$$

Applying $\frac{\partial}{\partial \ln k}$ to the equation, we obtain

$$\ddot{y}_k + 3Hy_k + \left(\frac{k^2}{a^2} + m^2\right)y_k = -2\frac{k^2}{a^2}u_k, \tag{4.89}$$

where $y_k \equiv \frac{\partial}{\partial \ln k}u_k$. Note that the homogeneous solutions for $y_k$ are $u_k$ and $u_k^*$. Thus, we use the Green function method to find a solution

$$y_k(t) = \int^t dt' \frac{a^3(t')}{i}\left(u_k^*(t)u_k(t') - u_k(t)u_k^*(t')\right)\left(-2\frac{k^2}{a^2}\right)u_k(t'). \tag{4.90}$$

From this, we find

$$\frac{d}{d\ln k}|u_k(t)|^2 = 2\mathrm{Re}\left[u_k^*(t)y_k(t)\right] \tag{4.91}$$

$$= 4\int_{-\infty}^t dt_z a_z^3 \frac{k^2}{a_z^2}\mathrm{Im}\left[u_k^2(t)u_k^{*2}(t_z)\right] \tag{4.92}$$

$$= \left[\int_{t_p}^t dt_z + \int_{-\infty}^{t_p} dt_z\right]4a_z^3\frac{k^2}{a_z^2}\mathrm{Im}\left[u_k^2(t)u_k^{*2}(t_z)\right]. \tag{4.93}$$

The second term is oscillatory with respect to $k$ so that we can safely neglect it after the momentum integral. Inserting this back to the integral (4.87), we obtain

$$I_N^{(2)}(p) \approx -C_N^{-1}\left|\zeta_p^o(t)\right|^2\int_p \frac{d^3k}{(2\pi)^3}\frac{d}{d\ln k}|u_{N,k}(t)|^2 + O\left(\frac{p^2}{a^2}\right) \tag{4.94}$$

$$= -C_N^{-1}\left|\zeta_p^o(t)\right|^2\left[-\frac{k^3}{2\pi^2}|u_{N,k}(t)|^2\Big|_p^{\Lambda_{UV}} + 3\left\langle(\sigma_N^2)_p\right\rangle\right] + O\left(\frac{p^2}{a^2}\right) \tag{4.95}$$

where we have put the subscript $N$ and the field normalization $C_N$ back, and

$$\left\langle(\sigma_N^2)_p\right\rangle \equiv \int_p \frac{d^3k}{(2\pi)^3}|u_{N,k}(t)|^2, \tag{4.96}$$

where the subscript $p$ stands for the comoving IR cut-off of momentum. One can then compute the contribution of the first term in Eq. (4.82) in a similar manner:

$$I_N^{(1)} \equiv \int d^3x e^{-i\vec{p}\cdot\vec{x}} \int^t d^4z \sqrt{-g_z} \left\langle \left[ \sigma_N^2(t,\vec{x})\zeta(t,\vec{0}), i\,(-3)\,\mathcal{L}_\sigma(z)\zeta(z) \right] \right\rangle \quad (4.97)$$

$$= 3C_N^{-1} \left| \zeta_p^o \right|^2 \left\langle \left( \sigma_N^2 \right)_p \right\rangle + O\left( \frac{p^2}{a^2} \right). \quad (4.98)$$

Hence, we obtain

$$\widetilde{\langle (\sigma^2)_r \zeta \rangle_p^C} = \sum_{N=0}^s I_N^{(1)} + I_N^{(2)} + O\left( \frac{p^2}{a^2} \right) \quad (4.99)$$

$$= \left| \zeta_p^o \right|^2 \frac{p^3}{2\pi^2} \left| u_p(t) \right|^2 + O\left( \frac{p^2}{a^2} \right) \quad (4.100)$$

where $u_p$ is the mode function for physical field $\sigma$.

Comparing the computation of Eq. (4.100) with the estimate in Sec. 4.3, we see two crucial differences:

1. There is a cancellation of the $3C_N^{-1} \left| \zeta_p^o \right|^2 \left\langle \left( \sigma_N^2 \right)_p \right\rangle$ term that is sensitive to mode summation that extends to sub horizon modes.

2. The $\Lambda_{UV}$ dependent term in Eq. (4.95) in the present computation disappears after accounting for the PV regulator fields. In contrast, the estimate in Sec. 4.3 leaves behind a $\Lambda_{UV} = aH_{inf}$ dependent contribution due to the ad hoc nature of the UV cutoff which does not preserve diffeomorphism.

Finally, putting the results (4.95) and (4.98) together, the two-point function becomes

$$\widetilde{\langle (\sigma^2)_r \zeta \rangle_p^C} \bigg|_{t_e} = \left| \zeta_p^o \right|^2 \times \begin{cases} \frac{\Gamma^2(\nu)H^2}{\pi^3} \left( \frac{p}{2a(t_e)H} \right)^{3-2\nu} & \text{massive scalar in dS} \\ \frac{H_p^2}{4\pi^2} & \text{massless during quasi-dS} \end{cases} \quad (4.101)$$

up to the $O\left( \frac{p^2}{a^2} \right)$ correction where $H_p$ denote the Hubble scale at which scale $p$ exits the horizon, $\nu = \sqrt{9/4 - m^2/H^2}$, and $t_e$ reminds us that we are evaluating this at the

end of inflation. We have applied (quasi)-dS mode function in evaluating (4.101).[14] One can easily check that Eq. (4.47) is consistent with Eq. (4.60).

As explained near Eq. (4.65), the vanishing of the cross-correlation in the limit $p \to 0$ is expected from the diffeomorphism Ward identity. For a nonvanishing $p$, one might expect the cross-correlation should be $O(p^2/a^2)$ by Taylor-expanding the cross-correlation at $p = 0$. However, Eq. (4.101) interestingly shows that the leading term of the cross-correlation is not analytic at $p = 0$ and thus not $p^2/a^2$-suppresed. Indeed, for any small $p/a(t_e)$, we can diminish the suppression by making $3 - 2\nu \to 0^+$ through the limit $m/H \to 0$.

To finish the computation of $\beta$, we also consider the two-point correlator $\langle (\sigma^2)_r (\sigma^2)_r \rangle$ showing up in the denominator. Again, the comoving gauge is convenient for this computation. Although the correlator is UV divergent, because the counter terms associated with the divergence are derivatively suppressed, we do not need to include the counter terms in computing the IR contributions and the non-derivative contribution of the correlator is insensitive to renormalization. Furthermore, the IR contribution using the super-horizon approximation is not UV divergent. That means the UV contribution and the IR contribution are cleanly separated. Thus, we can estimate $\langle \widetilde{(\sigma^2)_r (\sigma^2)_r} \rangle$ using only the super-horizon approximation unlike in the computation of $\langle \widetilde{(\sigma^2)_r} \zeta \rangle$. We find

$$\langle \widetilde{(\sigma^2)_r (\sigma^2)_r} \rangle_p^C = 2 \int_{\Lambda_{IR}} \frac{d^3 k_1}{(2\pi)^3} d^3 k_2 \delta^3(\vec{k}_1 + \vec{k}_2 - \vec{p}) |u_{k_1}(t)|^2 |u_{k_2}(t)|^2 + O\left(\frac{p^2}{a^2}\right) \quad (4.102)$$

---

[14]After inflation ends at time $t_e$, the cross correlation is expressed as

$$\langle \widetilde{(\sigma^2)_r} \zeta \rangle_p^C = f_T \langle \widetilde{(\sigma^2)_r} \zeta \rangle_p^C \Big|_{t_e}$$

where $f_T$ accounts for the change in the mode-function behavior after the end of inflation. As alluded to in the discussion near Eq. (4.24), the factor $f_T$ cancels out of the expression in $\beta$ due to its appearance in the denominator $\sqrt{\Delta_\zeta^2 \Delta_{\delta_S}^2}$. The factor $f_T$ can also account for the corrections in the superhorizon mode function behavior during inflation due to deviations away from the exact dS background.

where $\Lambda_{IR}$ is a comoving IR cutoff. Evaluating this with dS super horizon modes and assuming $m < 3H/2$, we find the value at the end of inflation to be

$$\langle (\widetilde{(\sigma^2)_r} \, \widetilde{(\sigma^2)_r})_p^C \rangle \Big|_{t_e}$$

$$\approx 2 \int_{\Lambda_{IR}} \frac{d^3 k_1}{(2\pi)^3} d^3 k_2 \delta^3(\vec{k}_1 + \vec{k}_2 - \vec{p})$$

$$\times \frac{2^{-4+4\nu} |\Gamma(\nu)|^4}{\pi^2} \frac{1}{a^6(t_e)H^2} \left( \frac{k_1}{a(t_e)H} \right)^{-2\nu} \left( \frac{k_2}{a(t_e)H} \right)^{-2\nu} \quad (4.103)$$

$$\approx \frac{1}{2\pi^2} \frac{H^4}{p^3} \frac{1}{3-2\nu} \left( \frac{p}{a(t_e)H} \right)^{6-4\nu} \left[ 1 - \left( \frac{\Lambda_{IR}}{p} \right)^{3-2\nu} \right]. \quad (4.104)$$

In Eq. (4.102), we have introduced a comoving IR cutoff $\Lambda_{IR}$ which corresponds to the statement that inflationary era had a beginning in the finite past. Explicitly, we cannot use the Bunch-Davies vacuum boundary condition for modes that left the horizon before the beginning of inflation. This means that

$$\frac{\Lambda_{IR}}{p} \sim e^{-(N_{\text{tot}} - N(p))} \quad (4.105)$$

where $N_{\text{tot}}$ is the total number of efolds of inflation, $N(p)$ is the number of efolds before the end of inflation at which the mode $p$ left the horizon: i.e. $p/a(N) = H$. This cutoff is related to the box cutoff introduced in [87, 90, 255]. Numerically, $\Lambda_{IR} \ll p$ is irrelevant when

$$\frac{m_\sigma^2}{H^2} \gg \frac{1}{N_{\text{tot}} - N(p)}. \quad (4.106)$$

For situations in which this condition is violated, IR effects are important, and our computation is only qualitatively suggestive since $\Lambda_{IR}$ has to be resolved using more detailed description of the beginning of inflation. In particular, since we do not physically expect $N_{\text{tot}} = \infty$, $m_\sigma = 0$ situation is not accurately captured by our computation. Of course, the IR sensitivity here is not important as far as the importance of the cross correlation is concerned since the qualitative behavior of having

$p/\Lambda_{IR} \to \infty$ is to make the correlation even larger making the $\beta$ parameter even smaller. Finally, note that Eq. (4.106) can easily be more stringent than Eq. (4.49).

Hence, we conclude

$$\beta \approx \begin{cases} -\frac{m_\sigma}{H}\sqrt{\frac{\Delta_\zeta^2}{6}} & \text{massive scalar in dS} \\ -\frac{\sqrt{\Delta_\zeta^2}}{2}\left(\ln\frac{p}{\Lambda_{IR}}\right)^{-1/2} & \text{massless during quasi-dS} \end{cases} \tag{4.107}$$

where for the massive scalar case is assume to satisfy Eq. (4.49). Although this in principle is a generic prediction of isocurvature scenario, the magnitude of around $10^{-5}$ is difficult to probe experimentally since the current sensitivity is at the level of $10^{-2}$.

## 4.4   Application

The $\beta$ computation presented in Eq. (4.107) is not sensitive to $\dot{\rho}_\sigma$ that is involved in the definition of the isocurvature perturbation $\delta_S$. Instead, it is a property of quadratic nature of the scalar composite operator during inflation. Since Eq. (4.107) does depend on the masses, in this section, we motivate couple of the mass parameters from well-motivated nonthermal dark matter models: WIMPZILLAs [65, 66, 68, 86, 100, 186, 256] and axions [257–259]. Although these two particles have different physical origins, they share some common properties as a cosmological component. Firstly, since they are massive (at the CMB time at least) and weakly interacting, they both are good CDM candidates. Also, they can be gravitationally produced during or after inflation, and this gives rise to isocuvature from their density perturbations. Furthermore, when their background field values are negligibly small, the isocurvature perturbation from these particles is approximated by quadratic form $\sigma^2$. In that case, they would present detectable non-Gaussianties [18, 59, 90] and their cross correlation is characterized by Eq. (4.107).

## Weakness of $\sigma$ Interactions with $\psi$

To connect our computation of $\beta$ to observables, a post inflationary isocurvature scenario is necessary. For the illustrative situations of axions and WIMPZILLAs, it is sufficient to assume that $\sigma$ has an extremely weak interaction with the reheating degrees of freedom $\psi$ and the inflaton $\phi$ such that the transfer function of $\sigma$ is trivial after inflation: with sufficiently small interactions, $\alpha$ and $\beta$ of Eqs. (4.28) and (4.29) computed during inflation can be directly matched without any further transfer function computations to isocurvature initial condition for CMB codes such as CMBFAST. In this section, we quantify the requisite weakness of the interactions and qualitatively discuss the situation when the weakness assumption is invalid. For example, we will show below that ordinary WIMPs are too strongly interacting with the reheating degrees of freedom for this assumption to be valid while axions and WIMPZILLAs are sufficiently weakly interacting. We also qualitatively describe what extra work needs to be done to apply this work for observations in situations in which the dark matter particles are not extremely weakly interacting.[15]

At the linearized classical equation of motion level, we have the gauge invariant perturbations $\{\zeta_j\}$ being governed by a linear time evolution operator

$$O[\{\zeta_j\}] = 0 \tag{4.108}$$

where the initial condition for the isocurvature species $j = \sigma$[16] is given by

$$\zeta_\sigma(t_i) = f(t_i) \tag{4.109}$$

---

[15]Because of the cross correlation result in this work is small, the discussion here is a bit academic if this discussion applied only to the cross correlation result. However, the discussion here applies to the isocurvature 2-point function found in the literature [18, 40, 67, 142, 149–151, 255] which has a realistic chance of being observable in near future experiments.

[16]In our scenario, the isocurvature species stand for the degrees of freedom constrast with the radiation degrees of freedom.

which in turn is set by the inflationary physics. For example, the initial time $t_i$ can be set to be the time of end of inflation. The final $\zeta_\sigma(t_f)$ will contain contribution which does not vanish in the limit $f \to 0$. Hence, one can write

$$\zeta_\sigma(t_f) = G^\sigma_{t_f}[f(t_i), 0] + G^\sigma_{t_f}[0, \zeta_{j \neq \sigma}(t_i)] \tag{4.110}$$

where $G^\sigma_{t_f}[D]$ is the $\sigma$ component of the Green's function derived from the linear operator $O$ which takes the initial data $D$ and maps it to the final value of $\zeta_\sigma(t_f)$. Note that we have implicitly assumed the boundary condition such that $G^\sigma_{t_f}[0, 0] = 0$ which means that $G^\sigma_{t_f}[f(t_i), 0]$ vanishes as $f(t_i) \to 0$.

Now, we will consider two situations in which bound the picture of super weakly interacting scenarios. In the first scenario, the thermal plasma generated by the inflaton decay will interact with $\sigma$ sufficiently strongly to make $\delta_S$ mix strongly with $\zeta$. In the second scenario, the inflaton decay to $\sigma$ directly will realign $\sigma$ fluctuations during radiation domination to those of $\zeta$, even though $\sigma$ and reheating products are not interacting appreciably.

First, consider the effects of radiation dominated thermal plasma on $\sigma$. The mixing rate governing $G^\sigma_{t_f}[0, \zeta_{j \neq \sigma}(t_i)]$ is the production rate of $\sigma$ particles from the thermal plasma. Typically a single channel involving particle $y$ dominates the production of the $\sigma$ particle from the plasma. (If there are more channels, the discussion below can easily be generalized.) We thus expect a qualitative behavior of

$$G^\sigma_{t_f}[0, \zeta_{j \neq \sigma}(t_i)] \sim \left(1 + \tanh\left[\frac{\Gamma(yy \to \sigma\sigma, t_{\max})}{H(t_{\max})}\right]\right) \zeta_y \tag{4.111}$$

where $\Gamma(yy \to \sigma\sigma, t_{\max})$ is the reaction rate for this process at the time that the production rate is maximum (in $\Gamma(yy \to \sigma\sigma, t)$ is maximum at $t = t_{\max}$ where $t_{\max} \in [t_i, t_f]$), $H$ is the expansion rate, and $\zeta_y = O(\zeta_{\text{tot}})$.

Hence, one sees that the information about the isocurvature perturbations depend

not only on

$$\Gamma(yy \to \sigma\sigma, t_{\max})/H(t_{\max})$$

but on $t_f$ since $t_{\max}$ is restricted to be in the range $t_{\max} \in [t_i, t_f]$. For example, the usual CMB code is run starting with an initial condition at $T \ll T_{\mathrm{BBN}}$. This means that $t_f \gg t_{\mathrm{BBN}}$ is required to use the inflationary correlator computations in the CMB code. A general computation of $G^\sigma_{t_f}$ needed for the prediction of isocurvature perturbation effect on CMB temperature is beyond the scope of current work. To be able to trust the trivial transfer function of

$$G^\sigma_{t_f}[f(t_i), 0] \approx f(t_i) \gg G^\sigma_{t_f}[0, \zeta_{j\neq\sigma}(t_i)], \tag{4.112}$$

for superhorizon modes (where $t_i$ is say at the end of inflation[17]), we can require

$$\frac{\Gamma(yy \to \sigma\sigma, t_{\max})}{H(t_{\max})} \ll \frac{\zeta_\sigma(t_i)}{\zeta_{\mathrm{tot}}} \tag{4.113}$$

where $t_{\max}$ can be at any time between inflation and the time at which boundary conditions are set for the CMB code. This sets a bound on the cross section $\langle \sigma v \rangle$ for $yy \to \sigma\sigma$ to be

$$\langle \sigma v \rangle \ll \frac{\zeta_\sigma(t_i)}{\zeta_{\mathrm{tot}}(t_i)} \frac{g_*^{3/4}}{g_y} \left(\frac{T_{\mathrm{RH}}}{10^6 \text{ GeV}}\right)^{-1} 4.2 \times 10^{-25} \text{ GeV}^{-2} \tag{4.114}$$

where the bound becomes more stringent for higher reheating temperatures.

This number should be compared to typical thermal WIMP DM candidate annihilation cross section of $10^{-9}$ GeV$^{-2}$ and a high energy s-channel scattering at $T_{RH}$ mediated through a vector boson with a dimensionless coupling $g = \sqrt{4\pi\alpha_g}$:

$$\langle \sigma v_{yy \to A^\mu \to \text{light states}} \rangle \sim \frac{\alpha_g^2}{T_{RH}^2} \tag{4.115}$$

$$= \left(\frac{\alpha}{10^{-1}}\right)^2 \left(\frac{T_{RH}}{10^6 \text{ GeV}}\right)^{-2} 10^{-14} \text{ GeV}^{-2}. \tag{4.116}$$

[17]Note that as discussed in footnote 14, $\alpha$ can also receive corrections from the departures from the ideal dS mode function evolution as well as from the time when $m/H$ becomes larger than unity. As discussed there, the quantity $\beta$ is not as sensitive to these corrections.

Hence, one sees that WIMP dark matter cannot play the role of the isocurvature perturbations. That is why if we are to identify our computation of $\alpha$ and $\beta$ directly to physical observables, we have to choose the isocurvature degree of freedom to be nonthermal.[18]

Even though the current work applies most immediately without changes to nonthermal dark matter scenarios having extremely weak interactions, Eq. (4.114) is still much bigger than gravity mediated s-channel interactions

$$\langle \sigma v_{yy \to g_{\mu\nu} \to \sigma\sigma} \rangle \quad \sim \quad \frac{1}{16\pi^2} \frac{T_{RH}^2}{M_p^4} \tag{4.117}$$

$$\sim \quad \left( \frac{T_{RH}}{10^6 \text{ GeV}} \right)^2 10^{-64} \text{ GeV}^{-2}. \tag{4.118}$$

For example, axion cross sections for gluon coannihilation behave as [261]

$$\langle \sigma v_{ag \to X} \rangle \quad \sim \quad \frac{\alpha_s^2}{8\pi^2} \frac{1}{f_a^2} \tag{4.119}$$

$$\sim \quad \left( \frac{f_a}{10^{12} \text{ GeV}} \right)^{-2} 10^{-28} \text{ GeV}^{-2} \tag{4.120}$$

where $f_a$ is the PQ breaking VEV. Hence, there is a large class of weakly interacting models for which this work directly applies without modification. For models for which Eq. (4.114) is not satisfied, one needs to compute the transfer function associated with the mixing. Nonetheless, this work will still be useful in setting the initial conditions for such computations.

Let's see qualitatively what happens when Eq. (4.114) is not satisfied. In that case, we expect mixing between isocurvature and curvature perturbations

$$\zeta_\sigma(t_f) = G_{t_f}^\sigma[f(t_i), 0] + G_{t_f}^\sigma[0, \zeta_{j \neq \sigma}(t_i)] \sim O(\zeta_\sigma) + O(\zeta_{\text{tot}}). \tag{4.121}$$

Since the curvature perturbations will analogously be

$$\zeta(t_f) = G_{t_f}^R[f(t_i), 0] + G_{t_f}^R[0, \zeta_{j \neq \sigma}(t_i)], \tag{4.122}$$

---

[18]Similar arguments can also be made from unitarity [260].

we would then have

$$\delta_S = 3 \left( \left\{ G^\sigma_{t_f}[f(t_i), 0] - G^R_{t_f}[f(t_i), 0] \right\} + \left\{ G^\sigma_{t_f}[0, \zeta_{j\neq\sigma}(t_i)] - G^R_{t_f}[0, \zeta_{j\neq\sigma}(t_i)] \right\} \right).$$
(4.123)

Up to the accuracy that all species are equipartitioned, this quantity may vanish since there is cancellation in each of the terms in the bracket. It is beyond the scope of the current work to compute more precisely this cancellation we are focusing on scenarios which satisfy Eq. (4.114).

Suppose there is a direct decay of the inflaton to $\sigma$, and suppose there is no other appreciable interaction between $\sigma$ and other decay products of the inflaton. In that case, it is better to set the initial time $t_i$ to be at the time of inflaton decay completion such that $G^\sigma_{t_f}[f(t_i), 0]$ is still trivial. In that case, we have

$$\zeta_\sigma \equiv -\frac{A}{2} + \frac{\delta\rho_\sigma^{(\text{grav})} + \delta\rho_\sigma^{(\text{decay})}}{3(\bar{\rho}_\sigma^{(\text{grav})} + \bar{\rho}_\sigma^{(\text{decay})} + \bar{P}_\sigma^{(\text{grav})} + \bar{P}_\sigma^{(\text{decay})})}$$
(4.124)

$$= r_\sigma^{(\text{grav})}\zeta_\sigma^{(\text{grav})} + r_\sigma^{(\text{decay})}\zeta_\sigma^{(\text{decay})}$$
(4.125)

where $r_i$ has been defined in Eq. (4.11). Hence, we have

$$\delta_S = 3(\zeta_\sigma - \zeta_R)$$
(4.126)

$$= 3(r_\sigma^{(\text{grav})}\zeta_\sigma^{(\text{grav})} + r_\sigma^{(\text{decay})}\zeta_\sigma^{(\text{decay})} - \zeta_R).$$
(4.127)

If $\zeta_\sigma^{(\text{decay})} = \zeta_R$ is assumed, then

$$\delta_S = 3 \left[ 1 - r_\sigma^{(\text{decay})} \right] \left( \zeta_\sigma^{(\text{grav})} - \zeta_R \right).$$
(4.128)

This equation says that if most of the inflaton energy density goes to $\sigma$, then the isocurvature is negligible.

In the next two subsections, we now consider couple of mass motivations for nonthermal dark matter isocurvature candidates.

# WIMPZILLA

The WIMPZILLA was originally proposed to avoid the restriction from the assumption that the dark matter is a thermal relic. Thus, the WIMPZILLA is supposed to either be very heavy and/or very weakly interacting. In particular, we consider the possibility that the WIMPZILLA is gravitationally produced during the phase transition out of the quasi-de-Sitter phase of inflation. In that case, the model is controlled by two parameters: the ratio of mass to the Hubble scale of inflation $m_X/H_{\text{inf}}$, and the reheating temperature $T_{RH}$, where $X$ denotes a massive scalar field. Since the energy density is approximated as $\rho_X \sim m_X^2 X^2$ the relic density of $X$ is estimated as

$$\Omega_X h^2 \sim 10^{-1} \left( \frac{H_e}{10^{12} GeV} \right)^2 \left( \frac{T_{RH}}{10^6 GeV} \right), \tag{4.129}$$

where we have assumed that $m_X \sim H_e$, because a priori we know that we can find proper isocurvature and relic density in this mass range. (For a more detailed discussion of the relic abundance, see for example [18].) The isocurvature power spectrum depends on the details of the evolution of the background during inflation because the mode function of massive particle decays as $a^{-3+2\nu}$ (see a related discussion in footnote 17). However, we can generally obtain $\alpha \sim 0.067$ if $m_X \lesssim H_{inf}$, where $H_{inf}$ is the Hubble expansion rate when the CMB scale crosses the horizon [18]. The WIMPZILLA isocurvature has also the quadratic form like the axion. It thus generates the observable non-Gaussianities estimated as Eq. (4.31). Eq. (4.107) translates to the fractional cross-correlation of

$$\beta_{WIMPZILLA} \approx -0.4 \frac{m_X}{H_{inf}} \sqrt{\Delta_\zeta^2} \tag{4.130}$$

which justifies the constraint used in [18]. Since the naive estimate of Eq. (4.56) gives a gross overestimate $\beta$, one of the merits of this work is to put such worries to rest through the proper computation.

## Axion

In this subsection we assess the relevance of Eq. (4.107) to the axion scenario. Firstly, we review the axion scenario. In 1997, Peccei and Quinn proposed the global $U(1)_{PQ}$ symmetry in order to solve to the strong CP problem in the QCD. The axion is the Nambu-Goldstone boson associated with the symmetry after it is broken spontaneously. Many mechanisms have been proposed to produce axions in the early universe. We focus only on the "vacuum misalignment" mechanism here following Refs. [143–149, 151, 197]. In early universe, the axions are effectively massless and gain their mass when the QCD anomaly term (which explicitly breaks PQ symmetry) becomes physical after the chiral symmetry breaking QCD phase transition. After the universe cools down and the Hubble friction drops below the axion mass, the axions begin to coherently oscillate and they contributes to the CDM component of the universe because of their long lifetime.

Let us denote the PQ symmetry breaking scale by $f_a$. Because $n_a \propto \theta^2$ where $\theta$ is the axion angle, the relic axion density is estimated as

$$\Omega_a h^2 \sim \begin{cases} 2 \times 10^4 \left( \frac{f_a/N}{10^{16} \text{ GeV}} \right)^{7/6} \langle \theta^2 \rangle & \text{for } T_{osc} \gtrsim \Lambda_{QCD} \\ 5 \times 10^3 \left( \frac{f_a/N}{10^{16} \text{ GeV}} \right)^{3/2} \langle \theta^2 \rangle & \text{for } T_{osc} \lesssim \Lambda_{QCD}, \end{cases} \tag{4.131}$$

where we have neglected $O(1)$ factors due to diffusion, anharmonic correction, and temperature-dependent mass correction, and $T_{osc}$ is the temperature at which the axion starts to oscillate. The axion isocurvature in comoving gauge is written as

$$\delta_s^{(C)} = \omega_a \frac{\theta^2 - \langle \theta^2 \rangle}{\langle \theta^2 \rangle} = \omega_a \frac{2\theta_i \delta\theta + \delta\theta^2 - \langle \delta\theta^2 \rangle}{\langle \theta^2 \rangle}, \tag{4.132}$$

where $\omega_a \equiv \Omega_a / \Omega_{CDM}$, $\theta_i$ is the average of initial QCD vacuum angle $\theta$ over the observable universe, and $\delta\theta$ is inhomogeneity of $\theta$, i.e. $\theta(t, \vec{x}) = \theta_i(t) + \delta\theta(t, \vec{x})$. Then

the isocurvature power spectrum becomes

$$\widetilde{\langle \delta_s \delta_s \rangle} \sim \omega_a^2 \begin{cases} 3.5 \times 10^{10} \left( \frac{f_a/N}{10^{16} \text{ GeV}} \right)^{7/3} \widetilde{F} & \text{for } f_a/N \gtrsim 6 \times 10^{17} GeV \\ 2 \times 10^9 \left( \frac{f_a/N}{10^{16} \text{ GeV}} \right)^3 \widetilde{F} & \text{for } f_a/N \lesssim 6 \times 10^{17} GeV, \end{cases} \tag{4.133}$$

where

$$\widetilde{F} = 4\theta_i^2 \widetilde{\langle \delta\theta\delta\theta \rangle} + \widetilde{\langle \delta\theta^2\delta\theta^2 \rangle} + \theta_i \left[ \langle \delta\theta\delta\theta^2 \rangle + \langle \delta\theta^2\delta\theta \rangle \right]. \tag{4.134}$$

Since our primary interest is in the cross correlation with $\theta_i \approx 0$, we set it to zero.

Therefore, the adiabaticity parameter $\alpha$ defined in Eq. (4.28) is estimated as

$$\alpha \sim \omega_a^2 \begin{cases} 1.3 \times 10^{19} \left( \frac{f_a/N}{10^{16} GeV} \right)^{7/3} \Delta_\theta^2 & \text{for } f_a/N \gtrsim 6 \times 10^{17} GeV \\ 8.1 \times 10^{17} \left( \frac{f_a/N}{10^{16} GeV} \right)^3 \Delta_\theta^2 & \text{for } f_a/N \lesssim 6 \times 10^{17} GeV, \end{cases} \tag{4.135}$$

$$\begin{aligned} \Delta_\theta^2(p) &= \frac{p^3}{2\pi^2} \widetilde{\langle \delta\theta^2\delta\theta^2 \rangle} \\ &= \left( \frac{f_a}{N} \right)^{-4} \left( \frac{H_p^2}{2\pi^2} \right)^2 \ln \frac{p}{\Lambda_{IR}}, \end{aligned} \tag{4.136}$$

where $H_p$ is the Hubble scale at the horizon exit of mode $p$, and $\Lambda_{IR}$ is an IR cut-off. Here we have used Eq. (4.104) with the assumption that the axion is effectively massless during inflation. In the case that $\theta_i \ll \delta\theta$, the isocurvature has the quadratic form of gaussian variable $\delta\theta$, and it naturally becomes non-Gaussian perturbation. The isocurvature non-Gaussianity is estimated as Eq. (4.31).

These parameter constraints and predictions (4.131), (4.135) and (4.31) already have been investigated in the literature [59, 90, 149, 151] with the assumption that the axion isocurvature and the curvature is uncorrelated. Our result from Eq. (4.107) is

$$\beta_{axion} = -\frac{\sqrt{\Delta_\zeta^2}}{2} \left( \ln \frac{p}{\Lambda_{IR}} \right)^{-1/2} \lesssim 2.5 \times 10^{-5} \tag{4.137}$$

which is consistent with the assumptions made in the literature.

## 4.5 Summary

In this work, we have presented the first explicit computation of the gravitational interaction contribution to the cross-correlation between the curvature and quadratic isocurvature perturbations (which include dark matter isocurvature candidates such as the axion and the WIMPZILLA). Since the necessary and sufficient condition for the cross-correlation to dominate over the isocurvature perturbations in observational constraints is $|\beta| \gtrsim 4 \times 10^{-2}$, we have explicitly computed $\beta$, which incidentally is not sensitive to the background number density of the isocurvature degrees of freedom and post-inflationary mode function changes on superhorizon scales. Although a naive estimate of $\beta$ based on a diffeomorphism violating UV cutoff leads to the possibility $\beta \sim O(1)$ due to a large ratio that can appear between the numerator and the denominator of the expression for $\beta$, our explicitly diffeomorphism invariant computation leads to $|\beta| \lesssim \Delta_\zeta / 2 \approx 2.5 \times 10^{-5}$ because the numerator has a suppression as a consequence of a diffeomorphism Ward identity. Unfortunately, this is far below the current observational sensitivity of $|\beta| \gtrsim 10^{-2}$.

The smallness of the cross-correlation is explained by the fact that the super-horizon mode of the curvature perturbation $\zeta$ can be smoothly connected to the gauge mode, which is the spatial dilatation, in the zero external momentum limit. Hence, Eq. (4.100) vanishes when $p = 0$ and $m \neq 0$. In other words, this can be seen as a suppression due to a diffeomorphism Ward identity (i.e. uniform spatial rescaling invariance). A nontrivial structure revealed through our explicit computation is the suppression's non-analytic structure with respect to $p$: the cross correlation cannot be Taylor-expanded at $p = 0$, and this contribution is not $p^2/a^2$-suppressed.

Our rigorous result which incorporates UV renormalization of the composite operator in the curved background is also shown to be consistent with an estimate based

on a soft-$\zeta$ theorem, which allows one to factorize $\langle\zeta\zeta\rangle$ from $\langle\sigma^2\zeta\rangle$ as explained in Eq. (4.61). However, Eq. (4.61) requires two assumptions that can only be justified by an honest computation such as what is presented in subsection 4.3:

1. There is an effective IR cutoff of $p$ in evaluating $\langle\sigma^2\rangle$ due to the external momentum $p$ inserted into the composite operator.

2. The only UV renormalization property of $\langle\sigma^2\rangle$ that is relevant to leading $\hbar$ approximation is the preservation of diffeomorphism invariance.

Note that the proper diffeomorphism invariant UV treatment also allowed us to demonstrate that the cross-correlation is indeed gauge-invariant with one-loop correction through the gravitational coupling. This gauge invariance is checked explicitly by computing our cross correlation in both the comoving gauge and the uniform curvature gauge.

Physically, the curvature perturbation $\zeta$ can affect the particle density $\rho_\sigma$ and generate correlations only at its horizon crossing, because the perturbation $\zeta$ begins to freeze after its horizon exit, and it can be effectively treated as a gauge mode after the horizon exit. Positive cross correlation corresponds to the situation in which the $1 + \zeta$ enhancement in the expansion enhances the particle production (assuming that this enhances inhomogeneity) while the negative cross correlation corresponds to the situation in which the $1 + \zeta$ enhancement in the expansion dilutes the particle inhomogeneity. The latter dilution effect leads to $\beta > 0$, while the particle production enhancement effect corresponds to the quadratic scenario that we were interested in this work. This explains the sign $\beta < 0$ of our result.

Given the robustness of the smallness of $\beta$, the gravitational interaction contribution to the cross correlation should be negligible in most nonthermal dark matter isocurvature scenarios. In addition to giving a concrete computation supporting this,

our work serves as an interesting lesson in computing correlators of composite operators in curved spacetime in the context of inflationary cosmology.

## 4.6   Appendix: Transfer functions for Adiabatic and Isocurvature initial condition

The CMB temperature fluctuation with the leading order approximation (the integrated Sachs-Wolfe term is neglected) in the Newtonian gauge ($B = F = 0$, $E = 2\Phi$, $A = -2\Psi$) is

$$\frac{\Delta T}{T} \approx \frac{1}{4}\, \delta_\gamma|_r + \Phi|_r \,, \tag{4.138}$$

where the perturbations on the rhs are evaluated at the recombination. We can obtain these perturbations by solving the Einstein and Boltzmann equations with given initial conditions. A projection from a given initial condition to the final CMB temperature fluctuation is called transfer function. In the following subsections, we calculate that the $k$-dependence of the transfer functions for the adiabatic and the isocurvature initial conditions. In particular, we show that the isocurvature transfer function has the additional suppression factor $k_{eq}/k$ compared to the adiabatic one for small scale $k \gg k_{eq}$. Here we basically follow the calculation by Ref. [262, 263].

### Perturbation Equations

For explicit computation, we choose the Newtonian gauge for the scalar metric perturbation (4.2). For simplicity, we consider only photon and CDM fluids, which are denoted in the following equations by subscript $\gamma$ and $m$, respectively. This assumption is valid for the sake of identifying the difference between transfer functions for

adiabatic and isocurvature initial conditions, although baryon and neutrino should be taken into account for accurate description for transfer functions.

The conservation equations for dark matter and photon fluids in Fourier space are

$$\delta_m' = k^2 V_m + 3\Psi, \tag{4.139}$$

$$V_m' = -\mathcal{H} V_m - \Phi, \tag{4.140}$$

$$\delta_\gamma' = \frac{4}{3} k^2 V_\gamma + 4\Psi', \tag{4.141}$$

$$V_\gamma' = -\frac{1}{4}\delta_\gamma - \Phi, \tag{4.142}$$

where ' denotes the time derivative with respect to conformal time $\eta$, $\mathcal{H} \equiv a'/a$, $\delta_a \equiv \delta\rho_a/\rho_a$. Note that $\Phi = \Psi$ since they are perfect fluids. $V_X$ is the peculiar velocity for fluid $X$. These four equation are combined by eliminating $V_X$, and we have

$$\left(a\left(\delta_m' - 3\Phi'\right)\right)' = ak^2\Phi, \tag{4.143}$$

$$\delta_\gamma'' = 4\Phi'' - \frac{k^2}{3}\left(\delta_\gamma + 4\Phi\right). \tag{4.144}$$

The evolution of the metric perturbation is encoded in the Einstein equations. (00) and (ii) components are

$$k^2\Phi + 3\mathcal{H}\left(\Phi' + \mathcal{H}\Phi\right) = -\frac{1}{2M_p^2}a^2\left(\rho_m\delta_m + \rho_\gamma\delta_\gamma\right), \tag{4.145}$$

$$\Phi'' + 3\mathcal{H}\Phi' + \left(2\frac{a''}{a} - \mathcal{H}^2\right)\Phi = \frac{1}{6M_p^2}a^2\rho_\gamma\delta_\gamma. \tag{4.146}$$

Combining with other components, we also find the Poisson equation

$$-k^2\Phi = \frac{3}{2}\mathcal{H}^2\left[\Omega_m\delta_m + \Omega_\gamma\delta_\gamma - 3\mathcal{H}\left(\Omega_m V_m + \frac{4}{3}\Omega_\gamma V_\gamma\right)\right]. \tag{4.147}$$

With the definition of isocurvature (4.20) in Section 4.2

$$\delta_S = \delta_m - \frac{3}{4}\delta_\gamma, \tag{4.148}$$

where we have used $p_\gamma = \rho_\gamma/3$ and $p_m = 0$, we rewrite the differential equations of fluid and metric perturbations in terms of $\Phi$ and $\delta_S$

$$\Phi'' + 3\mathcal{H}\left(1 + c_s^2\right)\Phi' + \left[2\mathcal{H}' + \mathcal{H}^2\left(1 + 3c_s^2\right)\right]\Phi$$
$$+ k^2 c_s^2 \Phi = -\frac{2}{3}\frac{c_s^2}{M_p^2}a^2\rho_m\delta_S, \quad (4.149)$$

$$\frac{1}{3c_s^2}\delta_S'' + \frac{a'}{a}\delta_S' + \frac{k^2 y}{4}\delta_S = -\frac{1}{6}y^2 k^4 \tau_{eq}^2 \Phi, \quad (4.150)$$

where

$$y \equiv \frac{a}{a_{eq}} = \frac{\rho_m}{\rho_\gamma}, \quad \tau_{eq} = \frac{\sqrt{2}}{a_{eq}H_{eq}}, \quad c_s^{-2} \equiv 3\left(1 + \frac{3}{4}y\right). \quad (4.151)$$

In $\eta \to 0$ limit, Eqs. (4.149) and (4.150) admit two linearly independent solutions $\Phi(k, \eta \to 0) = \Phi^i(k)$, $\delta_S(k, \eta \to 0) = 0$, and $\Phi(k, \eta \to 0) = 0$, $\delta_S(k, \eta \to 0) = \delta_S^i(k)$, which corresponds to adiabatic initial condition and isocurvature initial condition, respectively.

## Adiabatic Initial Condition

For large scale perturbations, which enters the horizon later than the recombination. $\delta_S$ remains zero according to Eq. (4.150), and thus Eq. (4.149) is rewritten as

$$\frac{d^2\Phi}{dy^2} + \frac{21y^2 + 54y + 32}{2y(y+1)(3y+4)}\frac{d\Phi}{dy} + \frac{\Phi}{y(y+1)(3y+4)} = 0, \quad (4.152)$$

where is called as Kodama-Sasaki equation. This differential equation can be exactly solved, and we find

$$\Phi(k_l, y \gg 1) = \frac{9}{10}\Phi^i(k_l), \quad (4.153)$$

where the subscript $l$ stands for "super-horizon". For photon energy density $\delta_\gamma$, Eq. (4.141) in the long wavelength limit yields

$$\frac{1}{4}\delta_\gamma - \Phi = const. \quad (4.154)$$

and also Eq. (4.145) gives

$$\delta_\gamma(k_l, \eta \to 0) = -2\Phi(k_l, \eta) = -2\Phi^i(k_l). \tag{4.155}$$

For small scale perturbation, which enter the horizon during the radiation domi-
nated(RD) era, in the early RD limit $\eta \ll \eta_{eq}$, Eq. (4.149) becomes

$$\Phi'' + \frac{4}{\eta}\Phi' + \frac{k^2}{3}\Phi = 0, \tag{4.156}$$

and its solution with the adiabatic initial condition

$$\Phi(k_s, \eta < \eta_{eq}) = \frac{3}{(w\eta)^3}\left(\sin w\eta - w\eta \cos w\eta\right)\Phi^i(k_s), \tag{4.157}$$

where $w = k/\sqrt{3}$. After the perturbation enters the horizon,

$$\Phi(k_s, \eta < \eta_{eq}) \approx -\frac{3\cos w\eta}{(w\eta)^2}\Phi^i(k_s), \tag{4.158}$$

$$\delta_\gamma(k_s, \eta < \eta_{eq}) \approx -\frac{2M_p^2}{\rho_\gamma a^2}\Phi(k_s, \eta) = 6\Phi^i(k_s)\cos w\eta, \tag{4.159}$$

where the subscript $s$ means "sub-horizon", and the second equation is obtained by
the Poisson equation (4.147). Plugging this solution into Eq. (4.143), we find that

$$\delta_m(k_s, \eta < \eta_{eq}) \approx -9\Phi^i(k_s)\left(\ln w\eta + \gamma - \frac{1}{2}\right), \tag{4.160}$$

where $\gamma$ is the Euler Gamma constant. This shows that the dark matter density
perturbation grows logarithmically during the RD era.

Now we should match this with the solutions in the matter dominated (MD) era.
Because the time derivatives of $\Phi$ is negligible compared to the spatial derivatives,
Eq. (4.143) is approximated as

$$\delta_m'' + \mathcal{H}\delta_m' \approx -k^2\Phi \approx \frac{3}{2}\mathcal{H}^2\Omega_m\delta_m, \tag{4.161}$$

where we have used the Poisson equation (4.147). Then, it is rewritten as

$$y(1+y)\frac{d^2\delta_m}{d^2y} + \left(1 + \frac{3}{2}y\right)\frac{d\delta_m}{dy} - \frac{3}{2}\delta_m = 0, \tag{4.162}$$

and its general solution is

$$\delta_m = c_1 \left(1 + \frac{3}{2}y\right) + c_2 \left[\left(1 + \frac{3}{2}y\right) \ln \frac{\sqrt{1+y}+1}{\sqrt{1+y}-1} - 3\sqrt{1+y}\right]. \tag{4.163}$$

Matching this solution with Eq. (4.160) at $y \ll 1$, we find

$$\begin{aligned}
\delta_m(k_s, \eta > \eta_{eq}) &= -9\Phi^i(k)\left(\ln 2w\eta_* + \gamma - \frac{7}{2}\right)\left(1 + \frac{3}{2}y\right) \\
&\quad + 9\Phi^i(k)\left[\left(1 + \frac{3}{2}y\right)\ln \frac{\sqrt{1+y}+1}{\sqrt{1+y}-1} - 3\sqrt{1+y}\right] \quad (4.164) \\
&\rightarrow -\frac{27}{2}y\Phi^i(k)\left(\ln 2w\eta_* + \gamma - \frac{7}{2}\right) \quad \text{when } y \gg 1. \quad (4.165)
\end{aligned}$$

where $\eta_* \equiv \eta_{eq}/\left(\sqrt{2}-1\right) = 2\tau_{eq}$. Note that we have used the results from the Friedman equation

$$\mathcal{H}^2 = \frac{a_{eq}^2 H_{eq}^2}{2}\left(\frac{1}{y} + \frac{1}{y^2}\right), \tag{4.166}$$

$$y = \frac{\eta^2}{(2\tau_{eq})^2} + \frac{\eta}{\tau_{eq}}, \tag{4.167}$$

and Eq. (4.165) corresponds to Eq. (150) in Ref. [263].

Then using Eqs. (4.147) and (4.165), we get

$$\Phi(k_s, \eta > \eta_{eq}) \approx \frac{\ln\left(0.15 k_s \eta_{eq}\right)}{\left(0.27 k_s \eta_{eq}\right)^2}\Phi^i(k_s). \tag{4.168}$$

This shows that the gravitational potential is frozen after the matter-radiation equality. Similarly, we first find the general solution of Eq. (4.144) for sub-horizon modes

$$\delta_\gamma = c_1 \cos w\eta + c_2 \sin w\eta - 4\Phi, \tag{4.169}$$

where we have neglected that time derivatives of $\Phi$. Then matching this with Eq. (4.159), we get

$$\delta_\gamma(k_s, \eta > \eta_{eq}) \approx \left[6\cos\left(w\eta\right) - 4\frac{\ln\left(0.15 k_s \eta_{eq}\right)}{\left(0.27 k_s \eta_{eq}\right)^2}\right]\Phi^i(k_s). \tag{4.170}$$

Now we return factors due to the Silk damping and the acoustic sound speed

$$
\begin{aligned}
\delta_\gamma(k_s, \eta > \eta_{eq}) \quad \approx \quad & \left[ 3^{5/4}\sqrt{4c_s}\cos\left(k_s\int^\eta c_s(\eta')d\eta'\right)e^{-(k_s/k_D)^2} \right. \\
& \left. - \frac{4}{3c_s^2}\frac{\ln(0.15k_s\eta_{eq})}{(0.27k_s\eta_{eq})^2} \right] \Phi^i(k_s),
\end{aligned}
\tag{4.171}
$$

which is Eq. (153) in Ref. [263]. Notice that the the first term is dominant for the scales we are interested in. However, the second term becomes important for very small scales where the diffusion damping is not negligible, $k \gtrsim k_D$.

Finally, the SW term (4.138) becomes

$$
\frac{\Delta T}{T} \quad \approx \quad 
\begin{cases}
6\Phi^i(k)\cos w\eta & \text{if } k > k_{eq} \\[2mm]
\frac{3}{10}\Phi^i(k) & \text{if } k < \eta_r^{-1}.
\end{cases}
\tag{4.172}
$$

Note that

$$
\zeta^i \approx \zeta_R^i = -\Phi^i + \frac{1}{4}\delta_\gamma^i = -\frac{3}{2}\Phi^i.
\tag{4.173}
$$

## Isocurvature initial condition

For large scale perturbations, $\delta_S$ remains constant, and Eq. (4.149) has the solution

$$
\Phi(k_l, \eta) = -\left(\frac{x}{5}\right)\frac{x^2 + 6x + 10}{(x+2)^3}\delta_S^i(k_l),
\tag{4.174}
$$

where $x \equiv \eta/\eta_{eq}$. In the MD era, Eq. (4.174) gives

$$
\Phi(k_l, \eta \gg \eta_{eq}) = -\frac{1}{2}\delta_m(k_l, \eta \gg \eta_{eq}) = \frac{1}{4}\delta_\gamma(k_l, \eta \gg \eta_{eq}) = -\frac{1}{5}\delta_S^i(k_l),
\tag{4.175}
$$

where the last two equations are obtained from Eq. (4.145).

Now, we will see how the perturbations evolve during the RD era, and how they are connected small scale perturbations. In the early RD era, the source term and the last term on the left hand side of Eq. (4.150) is negligible because they are higher

order in $y$. Thus, the solution $\delta_S$ remains constant even inside the horizon. In that case, Eq. (4.149) becomes Eq. (4.156) with the source term $\delta_S/2y\eta_{eq}^2$. Then we find its solution that matches with Eq. (4.174)

$$\Phi(k, \eta < \eta_{eq}) = -\frac{\eta}{\eta_{eq}} \frac{1}{(w\eta)^4} \left[ 1 + \frac{(w\eta)^2}{2} - (\cos w\eta + w\eta \sin w\eta) \right] \delta_S^i(k). \quad (4.176)$$

Furthermore, in the $w\eta \to 0$ limit, we have

$$\Phi(k_l, \eta < \eta_{eq}) \approx -\frac{1}{8}\delta_S^i(k_l) \left( 1 - \frac{(w\eta)^2}{18} \right) y, \quad (4.177)$$

and putting this into Eq. (4.145), we find that

$$\delta_\gamma(k_l, \eta < \eta_{eq}) \approx -\frac{1}{2}\delta_S^i(k_l) \left( 1 - \frac{7}{18}(w\eta)^2 \right) y, \quad (4.178)$$

$$\delta_m(k_l, \eta < \eta_{eq}) \approx \delta_S^i(k_l) \left( 1 - \frac{3}{8}y \right) + \frac{7}{48}\delta_S^i(k)y(w\eta)^2. \quad (4.179)$$

As explained in Section 4.2, we have that $\Phi$ and $\delta_\gamma$ grows like $a$ during the RD era, meanwhile $\delta_m$ decreases.

For sub-horizon modes, Eq. (4.176) becomes

$$\Phi(k_s, \eta < \eta_{eq}) \approx -\frac{y}{(w\eta)^3} \left( \frac{w\eta}{2} - \sin w\eta \right) \delta_S^i(k_s), \quad (4.180)$$

and again plugging this into Eq. (4.145) yields

$$\delta_m(k_s, \eta < \eta_{eq}) \approx -\left( \frac{3}{2}\frac{\sin w\eta}{w\eta}y - 1 \right) \delta_S^i(k_s), \quad (4.181)$$

$$\delta_\gamma(k_s, \eta < \eta_{eq}) \approx -\frac{2\sin w\eta}{w\eta}y\delta_S^i(k_s). \quad (4.182)$$

Matching these with general solutions of perturbations (4.163) and (4.169), and also using Poisson equation (4.147) in the MD era, we get

$$\delta_m(k_s, \eta > \eta_{eq}) \approx \left( 1 + \frac{3}{2}y \right) \delta_S^i(k_s), \quad (4.183)$$

$$\delta_\gamma(k_s, \eta > \eta_{eq}) \approx \left[ -\frac{1}{0.35k_s\eta_{eq}} \sin(w\eta) + 4\frac{1}{(0.8k_s\eta_{eq})^2} \right] \delta_S^i(k_s), \quad (4.184)$$

$$\Phi(k_s, \eta > \eta_{eq}) \approx -\frac{1}{(0.8k_s\eta_{eq})^2}\delta_S^i(k_s), \quad (4.185)$$

Then the SW term becomes

$$\frac{\Delta T}{T} \approx \begin{cases} -\frac{1}{0.35k\eta_{eq}}\delta^i_S(k)\sin(w\eta) & \text{if } k > k_{eq} \\ -\frac{2}{5}\delta^i_S(k) & \text{if } k < \eta_r^{-1}. \end{cases} \tag{4.186}$$

Now we see from Eqs. (4.172) and (4.186) that the isocurvature transfer function has the additional suppression factor $k_{eq}/k$ compared to the adiabatic one for small scale $k > k_{eq}$.

## 4.7 Appendix: Review of Diffeomorphism Invariance

A symmetry in a classical field theory is preserved at the quantum level, if the regulator preserves this symmetry and if the functional measure is invariant under the symmetry transformation. The quantum symmetry is reflected in the transformation of the correlation functions.

For example, consider a scalar field $\sigma$ on a fixed manifold $(\mathcal{M}, g)$. The two point function is

$$\langle \sigma(x)\sigma(y)\rangle_g = \int D\phi e^{iS(\sigma;g)}\sigma(x)\sigma(y) \tag{4.187}$$

The two point function only depends on the metric field $g$ and points $x, y$. Intuitively, the symmetry says for any diffeomorphism $\varphi : \mathcal{M} \mapsto \mathcal{M}$, the metric field and the points changes as

$$g \mapsto \tilde{g} = (\varphi^{-1})^*g, \ x \mapsto \tilde{x} = \varphi(x), \ y \mapsto \tilde{y} = \varphi(y) \tag{4.188}$$

then the two-point function should remain invariant, i.e.

$$\langle \sigma(x)\sigma(y)\rangle_g = \langle \sigma(\tilde{x})\sigma(\tilde{y})\rangle_{\tilde{g}}. \tag{4.189}$$

The Ward identity is the infinitesimal version of this relation.

Let $\varphi = \exp(\epsilon X)$, then

$$
\begin{align}
\tilde{g} &= \exp(-\epsilon X)^* g = g - \epsilon \mathcal{L}_X g + \cdots \tag{4.190} \\
S(\tilde{g}, \sigma) &= S(g, \sigma) - \epsilon \int d^4 x \sqrt{g} \frac{1}{2} T_\sigma^{\mu\nu} \mathcal{L}_X(g)_{\mu\nu} + \cdots \tag{4.191} \\
\sigma(\tilde{x}) &= \sigma(x) + \epsilon \mathcal{L}_X \sigma(x) + \cdots \tag{4.192}
\end{align}
$$

Plugging this into Eq. (4.189) and Taylor expand with respect to $\epsilon$, one get

$$
-i \int d^4 z \sqrt{g} \frac{1}{2} \mathcal{L}_X(g)_{\mu\nu}(z) \langle T_z^{\mu\nu} \sigma_x \sigma_y \rangle_g + \langle \mathcal{L}_X(\sigma)_x \sigma_y \rangle_g + \langle \sigma_x \mathcal{L}_X(\sigma)_y \rangle_g = 0. \tag{4.193}
$$

Or equivalently, using

$$
\mathcal{L}_X(g)_{\mu\nu} = \nabla_\mu X_\nu + \nabla_\nu X_\mu \tag{4.194}
$$

and perform integration by part, we obtain

$$
i \nabla_\mu \langle T_z^{\mu\nu} \sigma_x \sigma_y \rangle_g = \frac{1}{\sqrt{g_x}} \delta^4(x-z) g^{\alpha\nu} \frac{\partial}{\partial x^\alpha} \langle \sigma_x \sigma_y \rangle_g + \frac{1}{\sqrt{g_y}} \delta^4(y-z) g^{\alpha\nu} \frac{\partial}{\partial y^\alpha} \langle \sigma_x \sigma_y \rangle_g \tag{4.195}
$$

which is the Ward identity for the path ordered vacuum expectation value. We can then write down the in-in expectation value Ward identity as

$$
\begin{align}
i \nabla_\mu \langle in | T_z^{\mu\nu+} \sigma_x^+ \sigma_y^+ | in \rangle_g &= \frac{1}{\sqrt{g_x}} \delta^4(x-z) g_x^{\alpha\nu} \frac{\partial}{\partial x^\alpha} \langle in | \sigma_x^+ \sigma_y^+ | in \rangle_g \\
&\quad + \frac{1}{\sqrt{g_y}} \delta^4(y-z) g_y^{\alpha\nu} \frac{\partial}{\partial y^\alpha} \langle in | \sigma_x^+ \sigma_y^+ | in \rangle_g \tag{4.196} \\
i \nabla_\mu \langle in | T_z^{\mu\nu-} \sigma_x^+ \sigma_y^+ | in \rangle_g &= 0 \tag{4.197}
\end{align}
$$

where we kept the external operator inserted on the forward branch. The fact that Eq. (4.197) has no contact term is easy to understand, since $T_z^{\mu\nu-}$ is inserted on the backward time branch of the manifold, it can never *contact* points $x$ and $y$.

# 4.8    Appendix: ADM formalism and Interaction Hamiltonian

We consider an inflationary model with the inflaton $\phi$ and an extra free massive scalar $\sigma$, where $\sigma$ is only gravitationally coupled with $\phi$.

$$S = \int (dx) \frac{1}{2} M_p^2 R + [-\frac{1}{2} g^{\mu\nu} \partial_\mu \phi \partial_\nu \phi - V(\phi)] + [-\frac{1}{2} g^{\mu\nu} \partial_\mu \sigma \partial_\nu \sigma - U(\sigma)] \qquad (4.198)$$

where $M_p^2 = \frac{1}{8\pi G} = 1$ and $(dx) = d^4 x \sqrt{|\det(g_{\mu\nu})|}$. The metric can be parametrized using ADM formalism [206][19],

$$g_{\mu\nu} = \begin{pmatrix} -N^2 + h_{ij} N^i N^j & h_{ij} N^j \\ h_{ij} N^j & h_{ij} \end{pmatrix}, \quad g^{\mu\nu} = \begin{pmatrix} -N^{-2} & N^i N^{-2} \\ N^i N^{-2} & h^{ij} - N^i N^j N^{-2} \end{pmatrix}, \tag{4.199}$$

where $h_{ij}$ is the metric tensor on the constant time hyper-surface, and $h^{ij}$ is the inverse metric. We use Latin indices $i, j \cdots$ for objects on the 3-dimensional constant time hyper-surface, and we use $h_{ij}$ and $h^{ij}$ to raise and lower the indices. Then the action (4.198) is rewritten as

$$
\begin{aligned}
S \;=\; & \frac{1}{2} \int (dx) \sqrt{h} \left[ N R^{(3)} - 2NV(\phi) - 2NU(\sigma) \right. \\
& + N^{-1} \left( E_{ij} E^{ij} - E^2 \right) + N^{-1} \left( \dot{\phi} - N^i \partial_i \phi \right)^2 - N h^{ij} \partial_i \phi \partial_j \phi \\
& \left. + N^{-1} \left( \dot{\sigma} - N^i \partial_i \sigma \right)^2 - N h^{ij} \partial_i \sigma \partial_j \sigma \right],
\end{aligned}
\qquad (4.200)
$$

where $E_{ij}$ and $E$ are given by

$$E_{ij} \;=\; \frac{1}{2}(\dot{h}_{ij} - \nabla_i^{(3)} N_j - \nabla_j^{(3)} N_i). \tag{4.201}$$

$$E \;=\; E_{ij} h^{ij}. \tag{4.202}$$

---

[19]We use $(-+++)$ sign convention for the metric, and physical time $t$ .

Consider the background solution driven by the inflaton,

$$\phi^{(0)} = \bar{\phi}(t), \quad \sigma^{(0)} = 0, \quad g_{\mu\nu}^{(0)} = \begin{pmatrix} -1 & 0 \\ 0 & a^2(t)\delta_{ij} \end{pmatrix}, \tag{4.203}$$

where they satisfy the background equations of motion

$$3H^2 = \frac{1}{2}\dot{\bar{\phi}}^2 + V(\bar{\phi}) \tag{4.204}$$

$$\dot{H} = -\frac{1}{2}\dot{\bar{\phi}}^2 \tag{4.205}$$

$$\ddot{\bar{\phi}} + 3H\dot{\bar{\phi}} + V'(\bar{\phi}) = 0. \tag{4.206}$$

The action for the perturbations can be obtained by Taylor-expanding the full action around the background solution. However, we may reduce the number of variables by imposing the ADM constraints:

$$0 = \frac{1}{N}[R^{(3)} - \frac{1}{N^2}(E_{ij}E^{ij} - E^2)] - 2NT^{00} \tag{4.207}$$

$$0 = \frac{2}{N}\nabla_i^{(3)}[\frac{1}{N}(E^{ij} - Eh^{ij})] + 2N^j T^{00} + 2T^{0j} \tag{4.208}$$

where

$$T^{\mu\nu} = T_\phi^{\mu\nu} + T_\sigma^{\mu\nu}, \tag{4.209}$$

$$T_\phi^{\mu\nu} = -g^{\mu\nu}\left[\frac{1}{2}(\partial\phi)^2 + V(\phi)\right] + \partial^\mu\phi\partial^\nu\phi, \tag{4.210}$$

$$T_\sigma^{\mu\nu} = -g^{\mu\nu}\left[\frac{1}{2}(\partial\sigma)^2 + U(\sigma)\right] + \partial^\mu\sigma\partial^\nu\sigma, \tag{4.211}$$

and choose a gauge.

One commonly used gauge is the comoving gauge, defined by [20]

$$\delta\phi = 0, \quad \gamma_{ii} = 0, \quad \partial_i\gamma_{ij} = 0 \tag{4.212}$$

where

$$h_{ij} = a^2(t)[e^\Gamma]_{ij}, \quad \Gamma_{ij} = 2\zeta\delta_{ij} + \gamma_{ij} \tag{4.213}$$

---

[20]In this section, Latin indices $i, j$ are raised and lowered by $\delta_{ij}$, and repeated indices are contracted.

The solution of $N$ and $N^i$ is

$$N^{(1,C)} = \frac{\dot{\zeta}}{H}, \quad N_i^{(1,C)} = \partial_i[-\frac{\zeta}{H} + \epsilon\frac{a^2}{\nabla^2}\dot{\zeta}]. \tag{4.214}$$

We find the scalar metric perturbations are

$$\delta g_{\mu\nu}^{(C)} = \begin{pmatrix} -2\frac{\dot{\zeta}}{H} & (-\frac{\zeta}{H} + \epsilon\frac{a^2}{\nabla^2}\dot{\zeta})_{,i} \\ (-\frac{\zeta}{H} + \epsilon\frac{a^2}{\nabla^2}\dot{\zeta})_{,i} & a^2\delta_{ij}2\zeta \end{pmatrix}, \tag{4.215}$$

where $\epsilon \equiv \dot{H}/H^2$. Plugging in the linear metric perturbation back to the action (4.200), we can get the perturbed action action up to cubic order

$$S^{(C)} = S_{\zeta\zeta}^{(C)} + S_{\sigma\sigma}^{(C)} + S_{\gamma\gamma}^{(C)} + S_{\zeta\zeta\zeta}^{(C)} + S_{\zeta\sigma\sigma}^{(C)} + \cdots \tag{4.216}$$

where

$$S_{\zeta\zeta}^{(C)} = \int dt d^3x a_x^3 \epsilon(\dot{\zeta}^2 - (\frac{\nabla}{a}\zeta)^2) \tag{4.217}$$

$$S_{\zeta\sigma\sigma}^{(C)} = \int d^4x a_x^3 [T_\sigma^{ij} a^2\delta_{ij}\zeta + T_\sigma^{0i}(-\frac{\zeta}{H} + \epsilon\frac{a^2}{\nabla^2}\dot{\zeta})_{,i} - T_\sigma^{00}\frac{\dot{\zeta}}{H}]. \tag{4.218}$$

The $\zeta$ cubic interaction and graviton actions can be found in [19].

Another commonly used gauge is the uniform curvature gauge, in which

$$h_{ij} = a^2(t)\,[e^\gamma]_{ij}\,, \quad \gamma_{ii} = 0, \quad \partial_i\gamma_{ij} = 0. \tag{4.219}$$

In this gauge, the inflaton degree of freedom is in $\delta\phi$. However, this degree of freedom can be represented using the gauge-invariant variable

$$\zeta = -\frac{H}{\dot{\phi}}\delta\phi^{(U)} \tag{4.220}$$

In this gauge, the ADM constraint renders

$$N^{(1,U)} = -\epsilon\zeta, \quad N_i^{(1,U)} = \partial_i[\epsilon\frac{a^2}{\nabla^2}\dot{\zeta}] \tag{4.221}$$

We get the linear metric perturbation as

$$\delta g_{\mu\nu}^{(U)} = \begin{pmatrix} 2\epsilon\zeta & \epsilon\frac{a^2}{\nabla^2}\dot{\zeta}_{,i} \\ \epsilon\frac{a^2}{\nabla^2}\dot{\zeta}_{,i} & 0 \end{pmatrix} \qquad (4.222)$$

The free action is the same as in Eq.(4.217), and $\sigma$-$\zeta$ cubic interaction action is

$$S_{\zeta\sigma\sigma}^{(U)} = \int d^4x a_x^3 [T_\sigma^{00}\epsilon\zeta + T_\sigma^{0i}\epsilon\frac{a^2}{\nabla^2}\dot{\zeta}_{,i}]. \qquad (4.223)$$

From these perturbed actions, we can obtain the interaction Hamiltonian. Particularly, note that up to the cubic interaction, $\mathcal{L}_{int} = -\mathcal{H}_{int}$. Thus $S_{\zeta\sigma\sigma} = -\int dt\, H_{\zeta\sigma\sigma}(t)$.

# 4.9 Appendix: Renormalization of Composite Operators

In renormalized perturbation theory, one requires a regulator and renormalization condition. In order to preserve the diffeomorphism invariance, we need to adopt a covariant regulator. Here we choose Pauli-Villars (PV) regulator, following [193, 194]. We will first review PV regularization in subsection 4.9, and renormalize $\sigma^2$ in subsection 4.9. For correlators involving time integrals, we describes the adiabatic expansion of time integral in subsection (4.9).

## Pauli-Villars Regularization

We introduce a set of scalar regulator fields $\chi_n$ for $n = 1, \cdots, s$ with the following free Lagrangian

$$\mathcal{L}_{PV} = \sum_{n=1}^{s} C_n \left( -\frac{1}{2}g^{\mu\nu}\partial_\nu\chi_n\partial_\nu\chi_n - \frac{1}{2}M_n^2\chi_n^2 \right). \qquad (4.224)$$

The number of regulator fields $s$ depends on how many independent divergences one need to remove. In order to eliminate UV divergences up to some even order $2D$,we

must take the $C_n$ and regulator masses $M_n$ to satisfy

$$\sum_{N=0}^{s} C_N^{-1} = 0 \quad , \qquad \sum_{N=0}^{s} C_N^{-1} M_N^2 = 0, \cdots \tag{4.225}$$

$$\sum_{n}^{s} C_n^{-1} M_n^{2D} = -m_\sigma^{2D} \tag{4.226}$$

where we used the notation $M_0^2 = m_\sigma^2$ and $C_0 = 1$, and let $\sigma_0 = \sigma$ and $\sigma_n = \chi_n$. We use $\Lambda$ to represent the set of $M_n$, and the regulator dependence should be removed by counter terms when $M_n$ goes to $\infty$ together.

On a homogeneous FRW background, the physical and regulator scalar field can be quantized as

$$[\sigma_N, \dot{\sigma}_M] = ia^{-3}(t)\delta^3(\vec{x} - \vec{y})\delta_{NM} C_N^{-1} \tag{4.227}$$

with the following mode decomposition

$$\sigma_N(\vec{x}, t) = \int \frac{d^3k}{(2\pi)^3} (a_{N,\vec{k}} u_{N,\vec{k}}(t) + c.c) \tag{4.228}$$

$$[a_{N,\vec{p}}, a_{M,\vec{k}}^\dagger] = (2\pi)^3 C_N^{-1} \delta_{NM} \delta^3(\vec{k} - \vec{p}), \tag{4.229}$$

where $u_{N,\vec{p}}(t)$ satisfies the usual equation of motion

$$\ddot{u}_{N,k} + 3H\dot{u}_{N,k} + \left(\frac{k^2}{a^2} + M_N^2\right) u_{N,k} = 0 \tag{4.230}$$

with the Bunch-Davies initial condition

$$u_{N,k}(t) \to \frac{1}{\sqrt{2k}a(t)} \exp\left(-i \int^t \frac{k}{a(t')} dt'\right) \quad \text{for } t \to -\infty \tag{4.231}$$

and Wronskian conditions[21]

$$u_{N,k}\dot{u}_{N,k}^* - \dot{u}_{N,k}u_{N,k}^* = -i/a^3. \tag{4.232}$$

---

[21] Our treatment here differs from [194] in that the physical scalar field $\phi$ here has no background solution, and the regulator field $\chi_n$ does not mix with $\phi$ by mass term.

Because $M_n \gg H$, Eq.(4.230) possesses the WKB-type solution

$$u_{n,k}(t) = \frac{1}{\sqrt{2\omega_k(t)a^3(t)}} \exp\left(-i\int^t \omega_k(t')dt'\right)\left[1 + \frac{f_1(t)}{\omega_k(t)} + \frac{f_2(t)}{\omega_k^2(t)} + O(\omega_k^{-3})\right],$$
(4.233)

where $\omega_k = \sqrt{k^2/a^2 + M_n^2}$ and $f_i$ are of zeroth order in $\omega_k$. Since we have to regulate up to quadratic divergence in correlator computations, we need to know

$$|u_{n,k}(t)|^2 = \frac{1}{2\omega_k(t)a^3(t)}\left[1 + \frac{2\text{Re}f_1(t)}{\omega_k(t)} + \frac{|f_1(t)|^2 + 2\text{Re}f_2(t)}{\omega_k^2(t)} + O(\omega_k^{-3})\right] \quad (4.234)$$

up to second order. Due to the equation of motion (4.230), $f_1$ should satisfy

$$\frac{d}{dt}\left(\frac{f_1}{k}\right) = \frac{i}{2\omega_k}\left(\dot{H} + 2H^2 + \frac{1}{2}\frac{\left(\dot{H} + 3H^2\right)M_n^2}{\omega_k^2} - \frac{5}{4}\frac{H^2 M_n^4}{\omega_k^4}\right). \quad (4.235)$$

Also, the Wronskian condition (4.232) yields

$$\text{Re}f_1 = 0, \quad |f_1|^2 + 2\text{Re}f_2 = \omega_k\frac{d}{dt}\left(\frac{\text{Im}f_1}{\omega_k}\right). \quad (4.236)$$

Then plugging these two results to Eq.(4.234) gives

$$|u_{n,k}|^2 = \frac{1}{2\omega_k a^3}\left[1 + \frac{\dot{H} + 2H^2}{2\omega_k^2} + \frac{\left(\dot{H} + 3H^2\right)M_n^2}{4\omega_k^4} - \frac{5H^2 M_n^4}{8\omega_k^6} + O(\omega^{-3})\right]. \quad (4.237)$$

**Renormalization of Composite Operator**

The renormalization of composite operators in curved space-time is the same as in flat space-time(see e.g. [5, 6, 193]) , just with new possible counter-terms made from curvature tensor. For an operator of dimension $n$, one need to consider all possible counter-terms of dimension $n$ or less. In our example model with free massive scalar $\sigma$, we renormalize $\sigma^2$ as

$$(\sigma^2)_r = (\sigma + \sum_n \chi_n)^2 + \delta Z_0(\Lambda, m_\sigma) + \delta Z_1(\Lambda, m_\sigma)R, \quad (4.238)$$

where $R$ is the Ricci scalar.

Next, we compute $\delta Z_i$'s divergent part. For example, let us consider the one point function

$$\langle (\sigma^2)_r \rangle = \sum_{N=0}^{s} C_N^{-1} \int \frac{d^3 k}{(2\pi)^3} |u_{N,k}|^2 + \delta Z_0 + \delta Z_1 R. \tag{4.239}$$

In order to determine the counter terms $\delta Z_0$ and $\delta Z_1$, we introduce a comoving scale $Q$ such that $H \ll Q/a \ll M_n$ to break the Fourier space into the UV and the IR sector. Then we use the WKB solution (4.237) for $k \gg Q$. Furthermore, the contribution from the PV fields for $k \ll Q$ vanishes since it is suppressed by $1/M_n$.

$$
\begin{aligned}
\sum_N C_N^{-1} \int \frac{d^3 k}{(2\pi)^3} |u_{N,k}|^2 &= \int^Q \frac{d^3 k}{(2\pi)^3} |u_{0,k}|^2 + \sum_{N=0}^{s} C_N^{-1} \int_Q^{\Lambda_{UV}} \frac{d^3 k}{(2\pi)^3} |u_{i,k}|^2 \quad (4.240) \\
&= \int^Q \frac{d^3 k}{(2\pi)^3} |u_{0,k}|^2 + \frac{1}{48\pi^2} R \left( \ln \frac{a}{2Q} + \frac{10}{12} \right) \\
&\quad - \frac{1}{96\pi^2} R \sum_{N=0}^{s} C_N^{-1} \ln M_N^2 \\
&\quad + \frac{1}{16\pi^2} \sum_{N=0}^{s} C_N^{-1} M_N^2 \ln M_N^2. \quad (4.241)
\end{aligned}
$$

Note that the arbitrary comoving scale $Q$ in the first two terms should cancel each other.

In order to absorb the PV regulator dependence, we need

$$\delta Z_0 = \frac{1}{16\pi^2} \left[ -\sum_N C_N^{-1} M_N^2 \ln M_N^2 + \mu_0^2 \right], \tag{4.242}$$

$$\delta Z_1 = \frac{1}{96\pi^2} \left[ \sum_N C_N^{-1} \ln \frac{M_N^2}{\mu_1^2} \right], \tag{4.243}$$

where $\mu_0$ and $\mu_1$ are unknown mass scales determined by renormalization conditions. We set $\mu_0 = 0$ to have $\langle (\sigma)_r^2 \rangle = 0$ for flat space-time.

## Adiabatic Expansion of Time Integral

In order to compute some correlators using the in-in formalism (4.32), such as two-point function $\langle \sigma^2 \zeta \rangle$, we need to integrate PV field contributions over time. In this subsection, we present how to calculate the time integral of PV fields by adiabatically expanding the integral.

For simplicity, consider a diagram with one internal vertex. Using the WKB solution (4.237) of a PV field, the general form of the time integral is

$$I(k_1, k_2, \cdots, t_f) = \int_{-\infty}^{t_f} dt \, G(k_1, k_2, \cdots; t_f, t) e^{-i \int_t^{t_f} \omega(t') dt'}, \qquad (4.244)$$

where $\omega(t) = \omega_{k_1}(t) + \omega_{k_2}(t) + \cdots$ and $G(k_1, k_2, \cdots; t_f, t) = O(\omega^n)$. Because the integrand is a rapidly oscillatory function, the dominant contribution comes near the final time $t_f$. Thus, using integration by parts we expand the integral with respect to $\omega$:

$$
\begin{aligned}
I(k_1, k_2, \cdots, t_f) &= \frac{G(k_1, k_2, \cdots; t_f, t_f)}{i\omega(t_f)} \\
&\quad - \int_{-\infty}^{t_f} dt \left( \frac{d}{dt} \frac{G(k_1, k_2, \cdots; t_f, t)}{i\omega(t)} \right) e^{-i \int_t^{t_f} \omega(t') dt'} \qquad (4.245) \\
&= \frac{G(k_1, k_2, \cdots; t_f, t_f)}{i\omega(t_f)} - \left( \frac{1}{i\omega(t)} \frac{d}{dt} \frac{G(k_1, k_2, \cdots; t_f, t)}{i\omega(t)} \right) \Bigg|_{t=t_f} \\
&\quad + \left[ \frac{1}{i\omega(t)} \frac{d}{dt} \left( \frac{1}{i\omega(t)} \frac{d}{dt} \frac{G(k_1, k_2, \cdots; t_f, t)}{i\omega(t)} \right) \right] \Bigg|_{t=t_f} \qquad (4.246) \\
&\quad + O(\omega^{n-4}).
\end{aligned}
$$

Note that the mode functions $u_{n,k}$ and $u_{n,k}^*$ appear in pairs because of Wick contraction. Hence, the final result should be written in terms of $|u_{n,k}(t_f)|^2$ and their time derivatives, and we can compute the time integral up to arbitrary order of $\omega$. It is straightforward to generalize this to the cases with any number of internal vertices.

# 4.10    Appendix: Two-Point Function $\left\langle \left(\sigma^2\right)_r \zeta \right\rangle$ in the Uniform Curvature Gauge

In this section, we compute $\langle (\sigma^2)_r \, \zeta \rangle$ using the uniform curvature gauge in the quasi-de Sitter(dS) background, where the slow-roll factor $\epsilon$ is constant. Then we will show that the results in the both gauges are consistent with each other. Particularly, for the massless limit, the next leading order term in the uniform curvature gauge that indeed decays as $p^2/a^2$.

The two-point function is the same as in the comoving gauge except that the counter term contribution appears in the leading order.

$$\widetilde{\langle (\sigma^2)_r \, \zeta \rangle}_p^U = \int d^3x \, e^{-i\vec{p}\cdot\vec{x}} \int^t d^4z \, a^3(t_z) \sum_{N=0}^{n} \left\langle \left[ \sigma_N^2(t,\vec{x})\zeta(t,\vec{0}), \frac{i}{2} \left( T_\sigma^{\mu\nu} \delta g_{\mu\nu}^{(U)} \right)_z \right] \right\rangle$$
$$+ \delta Z_1 \widetilde{\langle R\zeta \rangle}_p, \tag{4.247}$$

where $R$ is the Ricci scalar. After taking non-derivate interaction term $T_\sigma^{00} \delta g_{00}^{(U)}$ only, factoring $\epsilon$ and $\zeta$ out from the integral, we get

$$\widetilde{\langle (\sigma^2)_r \, \zeta \rangle}_p^U = i \left| \zeta_p^o \right|^2 \epsilon \int^t d^4z \, a^3(t_z) \sum_{N=0}^{n} \left\langle \left[ \sigma_N^2(t,\vec{x}), \left( T_\sigma^{00} \right)_z \right] \right\rangle$$
$$+ 24\epsilon H^2 \left| \zeta_p^o \right|^2 \delta Z_1 + O\left( \dot{\epsilon}, \epsilon^2, \frac{p^2}{a^2} \right), \tag{4.248}$$

where we have used the perturbed curvature in the uniform curvature gauge

$$R = 12H^2 - 6\epsilon H^2 + 24\epsilon H^2 \zeta + 4\epsilon H \dot{\zeta} + \cdots, \tag{4.249}$$

where $\cdots$ denotes $O(\dot{\epsilon}, \epsilon^2)$ terms or terms proportional to the equation of motion of $\zeta$.

Since $T_\sigma^{00} = \mathcal{L}_\sigma + \sum_N \left[ \left( \frac{\nabla}{a} \sigma_N \right)^2 + M_N^2 \sigma_N^2 \right]$, together with the identities (4.95),(4.98), and

$$i \int^t d^4z \, a^3(t_z) \left\langle \left[ \sigma_N^2(t,\vec{x}), \sigma_N^2(z) \right] \right\rangle = -2 \frac{\partial}{\partial M_N^2} \left\langle \left( \sigma_N^2 \right)_p \right\rangle, \tag{4.250}$$

with $T_\sigma^{00} = \mathcal{L}_\sigma + \sum_N \left[ \left(\frac{\nabla}{a}\sigma_N\right)^2 + M_N^2 \sigma_N^2 \right]$, we have

$$\langle \widetilde{(\sigma^2)_R \zeta} \rangle_p^U + \frac{1}{H}\frac{d}{dt}\langle (\sigma^2)_R \rangle \widetilde{\langle \zeta\zeta \rangle}_p = \sum_N F_N(t) + O\left(\dot\epsilon, \epsilon^2, \frac{p^2}{a^2}\right), \qquad (4.251)$$

$$F_N(t) = 2\epsilon \left\langle \left(\sigma_N^2\right)_p \right\rangle - \epsilon Z_N^{-1} \frac{k^3}{2\pi^2} |u_{N,k}|^2 \Big|_p^{\Lambda_{UV}}$$
$$-2\epsilon \frac{\partial}{\partial \ln M_N^2} \left\langle \left(\sigma_N^2\right)_p \right\rangle + \frac{1}{H}\frac{d}{dt}\langle \sigma_N^2 \rangle \qquad (4.252)$$

Although the rhs of Eq.(4.247) is well-defined and regulator independent, individual terms are not. Thus, we insert counter terms to have each term regulator independent

$$\sum_N F_N(t) = \epsilon\left(2\left\langle \left(\sigma^2(t)\right)_{r,p} \right\rangle + \frac{p^3}{2\pi^2}|u_p(t)|^2 - 2m_\sigma^2 \frac{\partial}{\partial m_\sigma^2}\left\langle \left(\sigma^2(t)\right)_{r,p} \right\rangle \right)$$
$$+ \frac{1}{H}\frac{d}{dt}\left\langle \left(\sigma^2\right)_r \right\rangle, \qquad (4.253)$$

where we have put the counter terms $\delta Z_0$ and $\delta Z_1 R$ into each one-point function, and the PV field contribution from the third term cancels with those from the other terms. Then, using the relation (4.77) one can find the rhs of Eq.(4.253) is consistent with the result (4.101) in the comoving gauge in the quasi-dS background after explicitly computing renormalized one-point function $\left\langle \left(\sigma^2(t)\right)_{r,p} \right\rangle$. On the other hand, the rhs does not depend on the renormalization as all counter terms cancel. Hence, we can arrive at the same conclusion using the one point function using super-horizon approximation in the dS space-time,

$$\left\langle \left(\sigma^2(t)\right)_{r,p} \right\rangle \approx \int_p^{caH} \frac{d^3k}{(2\pi)^3}|u_k(t)|^2 \approx \int_p^{caH} \frac{d^3k}{(2\pi)^3}\frac{|\Gamma(\nu)|^2}{4\pi Ha^3}\left(\frac{k}{2aH}\right)^{-2\nu}, \qquad (4.254)$$

where the arbitrary constant $c \lesssim O(1)$. Note that the UV boundary of the integral should be a comoving scale in order to to keep the spatial dilatation symmetry.

## Massless Limit

For the massless limit $m_\sigma^2/H^2 \ll \ln p/aH$, we can compute the two-point function explicitly without neglecting any gravitational couplings. We calculate up to the next

leading term here. We decompose Eq. (4.247) as

$$\widetilde{\langle(\sigma^2)_r\,\zeta\rangle}_p^U = I_0(p,t) + \sum_{n=1}^{s} I_n(p,t) + I_{c.t.}(p,t), \tag{4.255}$$

where $I_0$, $I_n$, and $I_{c.t.}$ are the contributions from the physical field $\sigma$, the PV field $\chi_i$ and the counter terms, respectively. Since all the gravitational couplings are $O(\epsilon)$ (See Eq. (4.223)), we may use the mode functions $\zeta_p$ and $u_k$ in the pure dS for $O(\epsilon)$ correction to the two-point function. Then a long but straightforward calculation gives

$$
\begin{aligned}
I_0(p,t) &= \int d^3x\, e^{-i\vec{p}\cdot\vec{x}} \int^t d^4z\, a^3(t_z) \\
&\quad \times \sum_{N=0}^{n} \left\langle \left[ \sigma^2(t,\vec{x})\zeta(t,\vec{0}), \frac{i}{2}\left(T_\sigma^{\mu\nu}\delta g_{\mu\nu}^{(U)}\right)_z \right] \right\rangle \tag{4.256} \\
&= \frac{\epsilon H^2}{4\pi^2}\,|\zeta_p^o|^2 \left[ -\frac{1}{3}\frac{p^3}{a^3 H^3}\frac{\Lambda}{aH} + 2\log\frac{\Lambda}{p} + \frac{5}{3}\frac{p^2}{a^2 H^2}\log\frac{\Lambda}{p} + 1 - \frac{p^2}{a^2 H^2} \right] \\
&\quad + O\left(\frac{p^4}{a^4 H^4}\right) + O(\epsilon^2,\dot\epsilon). \tag{4.257}
\end{aligned}
$$

The PV field contribution $I_n$ requires some more technical explanation. If we write the WKB solution (4.234) as

$$u_{n,k}(t) = \alpha_k(t)e^{-i\int^t w_k(t')dt'}, \tag{4.258}$$

the PV field contribution $I_n$ is written as

$$
\begin{aligned}
I_n(p,t) &= C_n^{-1} \int_Q \frac{d^3k_1}{(2\pi)^3} d^3k_2 \delta^{(3)}(\vec{k}_1 + \vec{k}_2 - \vec{p}) \\
&\quad \times \mathrm{Im}\left[ \int^t dt_z e^{i\int_t^{t_z}\left(\omega_{k_1}(t')+\omega_{k_2}(t')\right)dt'} G_n(k_1,k_2;t,t_z) \right], \tag{4.259}
\end{aligned}
$$

where

$$
G_n(k_1, k_2; t, t_z) = -2a_z^3 \zeta_p(t)\alpha_{k_1}(t)\alpha_{k_2}(t) \left( \sum_i \widehat{O_i} \right) \zeta_p^*(t_z)\alpha_{k_1}^*(t_z)\alpha_{k_2}^*(t_z) \quad (4.260)
$$

$$
\left( \widehat{O_1} \right) = \epsilon \left[ \left( i\omega_{k_1}(t_z) + \partial_{t_z}^{(1)} \right) \left( i\omega_{k_2}(t_z) + \partial_{t_z}^{(2)} \right) \right.
$$
$$
\left. - \frac{\vec{k}_1 \cdot \vec{k}_2}{a_z^2} + M_n^2 \right], \quad (4.261)
$$

$$
\left( \widehat{O_2} \right) = \left[ \frac{\vec{k}_2 \cdot \vec{p}}{a_z^2} \left( i\omega_{k_1}(t_z) + \partial_{t_z}^{(1)} \right) + \frac{\vec{k}_3 \cdot \vec{p}}{a_z^2} \left( i\omega_{k_2}(t_z) + \partial_{t_z}^{(2)} \right) \right]
$$
$$
\times \left( \epsilon \frac{a_z^2}{p^2} \partial_{t_z}^\zeta \right), \quad (4.262)
$$

where $\partial_{t_z}^{(i)}$ and $\partial_{t_z}^\zeta$ denotes the time derivative with respect to $\alpha_{k_i}^*(t_z)$ and $\zeta_p^*(t_z)$, respectively, and $\left( \widehat{O_1} \right)$ and $\left( \widehat{O_1} \right)$ correspond to the (00) and the (i0) components of the gravitational couplings, respectively. Notice that $\alpha_k = O\left(\omega^{-1/2}\right)$ and $G(k_1, k_2; t, t_z) = O(\omega^0)$, and thus $I_n$ has quadratic divergences superficially. However, the quadratic divergences arising from $\left( \widehat{O_1} \right)$ vanish in the $M_n \to \infty$ limit. Effectively, the integral (4.259) is linearly divergent. That means we have to adiabatically expand the integral to the second order. Similarly, the integral of the two-point function in the comoving gauge is quadratic divergent, and thus one need to expand the integral to the third order. This makes the computation easier in the uniform curvature gauge. Using

$$
|\alpha_k(t)|^2 = \frac{1}{2\omega_k a^3} \left[ 1 + \beta_2(k, t) + O(\omega_k^{-3}) \right], \quad (4.263)
$$

$$
\alpha_k(t)\dot{\alpha}_k^*(t) = \frac{1}{2\omega_k a^3} \left[ \gamma_0(k, t) - i\omega_k \beta_2(k, t) + O(\omega_k^{-2}) \right], \quad (4.264)
$$

$$
\alpha_k(t)\ddot{\alpha}_k^*(t) = \frac{1}{2\omega_k a^3} \left[ -3iH - 2i\gamma_0(k, t) + i\frac{k^2/a^2}{\omega_k^2}H + O(\omega_k^{-1}) \right], \quad (4.265)
$$

where

$$
\gamma_0(k, t) = -\frac{3}{2}H + \frac{1}{2}\frac{k^2/a^2}{\omega_k^2}H, \quad (4.266)
$$

$$
\beta_2(k, t) = \frac{\dot{H} + 2H^2}{2\omega_k^2} + \frac{\left( \dot{H} + 3H^2 \right) M_n^2}{4\omega_k^4} - \frac{5H^2 M_n^4}{8\omega_k^4}, \quad (4.267)
$$

which are obtained by combining Eq. (4.237) with Eq. (4.258), the integral (4.259)

becomes

$$
\begin{aligned}
I_n(p,t) &= \frac{C_n^{-1}}{4\pi^2}\epsilon H^2 \left|\zeta_p^o\right|^2 \left[ -\frac{1}{3}\frac{p^3}{a^3 H^3}\frac{\Lambda}{aH} + 2\log\frac{2\Lambda}{aM_n} + \frac{5}{3}\frac{p^2}{a^2 H^2}\log\frac{2\Lambda}{aM_n} - \frac{5}{3} \right. \\
&\quad \left. -\frac{25}{18}\frac{p^2}{a^2 H^2} + O\left(\frac{p^4}{a^4 H^4}\right) \right] + O(\epsilon^2,\dot\epsilon).
\end{aligned}
\tag{4.268}
$$

Note that all $\Lambda$ dependent terms in $I_0 + \sum_n I_n$ vanishes by the PV field normalization

conditions (4.226).

Putting Eqs. (4.257) and (4.268) together into Eq. (4.255), we have

$$
\begin{aligned}
\widetilde{\langle(\sigma^2)_R\zeta\rangle}_p^U &= \frac{1}{4\pi^2}\epsilon H^2\left|\zeta_p^o\right|^2 \left[ 2\log\frac{a\mu_1}{2p} + \frac{5}{3}\frac{p^2}{a^2 H^2}\log\frac{a\mu_1}{2p} + \frac{8}{3} \right. \\
&\quad \left. + \frac{7}{18}\frac{p^2}{a^2 H^2} + O\left(\frac{p^4}{a^4 H^4}\right) \right] + O(\epsilon^2,\dot\epsilon).
\end{aligned}
\tag{4.269}
$$

We still need to compute one-point function $\frac{d}{dt}\langle(\sigma^2)_r\rangle$ up to $O(\epsilon)$ in order to compare

the results in both gauges. Because mode functions for a massless scalar field are

$O(\epsilon^0)$, we need $O(\epsilon)$ correction on it. In a quasi-dS background, we take an ansatz

for the mode function

$$
u_k(t) = \left(\frac{1}{\sqrt{2k}a(t)} + i\frac{H(t)}{\sqrt{2k^3}}\right)e^{i\frac{k}{a(t)H(t)}} + \frac{\epsilon(t)}{\sqrt{2k}a(t)}f_k(t)e^{i\frac{k}{a(t)H(t)}},
\tag{4.270}
$$

where $f_k(t) = O(\epsilon^0)$ so that it recovers the dS solution in the $\epsilon \to 0$ limit. Applying

this to the differential equation

$$
\ddot u_k(t) + 3H\dot u_k(t) + \frac{k^2}{a^2}u_k(t) = 0,
\tag{4.271}
$$

we get

$$
\ddot f_k + \left(H(t) - 2i\frac{k}{a(t)}\right)\dot f_k - H(t)^2 f_k = 3H(t)^2 - 2i\frac{k}{a(t)}H(t) - 2\frac{k^2}{a(t)^2} + O(\epsilon),
\tag{4.272}
$$

whose solution is

$$
\begin{aligned}
f_k(t) &= -\frac{3}{2} + iq + \frac{i}{2}\frac{1}{q} + \left(1 - \frac{i}{q}\right)e^{-2iq}Ei(2iq) \\
&\quad + c_1\left(1 + \frac{i}{q}\right) + c_2\left(1 - \frac{i}{q}\right)e^{-2iq}
\end{aligned}
\tag{4.273}
$$

where $q = \frac{k}{a(t)H(t)}$, and $Ei$ is the exponential integral function

$$Ei(z) = -\int_{-z}^{\infty} \frac{e^{-t}}{t} dt \qquad (4.274)$$

$$Ei(\pm ix \to \infty) \to \pm i\pi + e^{\pm ix} \left( \frac{0!}{(\pm ix)} + \frac{1!}{(\pm ix)^2} + \frac{2!}{(\pm ix)^3} + \cdots \right). \qquad (4.275)$$

Matching this solution with the Bunch-Davies initial condition (4.231) and the Wronskian condition (4.232) respectively give

$$c_2 = -i\pi \text{ and } c_1 = \frac{1}{2}. \qquad (4.276)$$

Then the mode function with $O(\epsilon)$ correction in a quasi-dS space-time becomes

$$
\begin{aligned}
u_k(t) = & \left( \frac{1}{\sqrt{2ka}} + i\frac{H}{\sqrt{2k^3}} \right) e^{i\frac{k}{aH}} \\
& + \frac{\epsilon}{\sqrt{2ka}} \left[ -1 + i\frac{k}{aH} + i\frac{aH}{k} + \left( 1 - i\frac{aH}{k} \right) \left( -i\pi + Ei(2i\frac{k}{aH}) \right) e^{-2i\frac{k}{aH}} \right] e^{i\frac{k}{aH}} \\
& + O(\epsilon^2, \dot{\epsilon}).
\end{aligned} \qquad (4.277)
$$

Now we calculate the one-point function using this mode function as shown in Subsection 4.9, and we get

$$\frac{d}{dt} \left\langle (\sigma^2)_r \right\rangle = \frac{H^3}{4\pi^2} + \frac{\epsilon H^3}{2\pi^2} \left( \log \frac{H}{\mu_1} + \frac{1}{6} - \gamma \right) + O(\epsilon^2, \dot{\epsilon}). \qquad (4.278)$$

Finally, we find

$$
\begin{aligned}
& \frac{1}{H} \frac{d}{dt} \left\langle (\sigma^2)_r \right\rangle \widetilde{\langle \zeta\zeta \rangle}_p + \widetilde{\langle (\sigma^2)_r \zeta \rangle}_p^U \\
& = \frac{H^2(t)}{4\pi^2} |\zeta_p(t)|^2 + \frac{\epsilon H^2}{2\pi^2} |\zeta_p^o|^2 \left[ \log \frac{aH}{2p} + \frac{3}{2} - \gamma \right] \\
& + \frac{\epsilon H^2}{4\pi^2} |\zeta_p^o|^2 \frac{p^2}{a^2 H^2} \left[ \frac{13}{18} - 2\gamma + \frac{5}{3} \log \frac{a\mu_1}{2p} + 2 \log \frac{H}{\mu_1} \right] \\
& + O\left( \epsilon^2, \dot{\epsilon}, \frac{p^4}{a^4 H^4} \right).
\end{aligned} \qquad (4.279)
$$

The non-$p^2/a^2$-suppressed terms are rewritten as

$$\frac{H^2(t)}{4\pi^2}\left|\zeta_p^o\right|^2 + \frac{\epsilon H^2}{2\pi^2}\left|\zeta_p^o\right|^2\left[\log\frac{aH}{2p} + \frac{3}{2} - \gamma\right]$$

$$\approx \frac{H^2(t)}{4\pi^2}\left|\zeta_p^o\right|^2\left(1 + 2\epsilon\log\frac{aH}{p}\right) \tag{4.280}$$

$$\approx \frac{H^2(t)}{4\pi^2}\left|\zeta_p^o\right|^2\left(\frac{p}{aH}\right)^{-2\epsilon} \tag{4.281}$$

$$\approx \frac{H_*^2}{4\pi^2}\left|\zeta_p^o\right|^2. \tag{4.282}$$

As expected, this is the result (4.101) in the comoving gauge. The other terms are suppressed by the factor $p^2/a^2$. This explicitly proves that the next leading terms for the two-point function $\langle(\sigma^2)_r\,\zeta\rangle$ are $O(p^2/a^2)$.

# Part II

# Particle Production under Electromagnetic Fields

# Chapter 5

# Introduction

In this part of the thesis, we investigate the vacuum $e^{\pm}$ pair production (the Schwinger effect) and we suggest a toy model that may be responsible to the intermittent strong emissions, called nanoshots, from the Crab pulsar. The following sections in this chapter present short reviews for non-perturbative calculations for the pair production under strong electromagnetic fields. Unlike the previous part, we shall use the $(+, -, -, -)$ metric signature with the natural units $c = \hbar = \varepsilon_0 = 1$.

## 5.1    Operator Approach

Under an strong external electric field, the quantum vacuum is unstable because the virtual electron-positron pairs gain energy and momentum from the external field, and once they obtain energy more than the pair creation threshold $2m$, they finally become real pairs. The decay rate is calculated by Schwinger in 1951 in the language of quantum electrodynamics [264]. Here we present the calculation of the vacuum decay rate under the homogeneous electric and magnetic fields [264–266]. In the following derivation, we shall assume that the electric field $\vec{E}$ and the magnetic field

$\vec{B}$ are parallel. [1]

The QED effective action under an external vector potential $A_\mu$ is given by

$$e^{iW[A]} = \int \mathcal{D}\psi \mathcal{D}\bar{\psi} \exp\left[i \int dx^4 \mathcal{L}_{QED}\right], \tag{5.5}$$

$$\mathcal{L}_{QED} = \bar{\psi}\left(i\gamma^\mu \partial_\mu - e\gamma^\mu A_\mu - m\right)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \text{gauge-fixing term}, \tag{5.6}$$

where $\gamma^\mu$ are the Dirac matrices, and $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. Then the amplitude of the vacuum to vacuum transition is

$$|\langle \text{out}, 0|0, \text{in}\rangle|^2 = 2\text{Im}W[A], \tag{5.7}$$

where we have neglected the fluctuations of $A_\mu$. Formally the effective action is written as

$$W[A] = -i\text{Tr}\ln\left[\frac{i\gamma^\mu \partial_\mu - e\gamma^\mu A_\mu - m + i\epsilon}{i\gamma^\mu \partial_\mu - m + i\epsilon}\right]. \tag{5.8}$$

$$= -\frac{i}{2}\text{Tr}\ln\left[\frac{(i\partial - eA)^2 + \frac{e}{2}\sigma F - m^2 + i\epsilon}{-\partial^2 - m^2 + i\epsilon}\right]. \tag{5.9}$$

---

[1]This is because we can almost always choose a frame in such a way that the electric and magnetic fields are parallel for general uniform static electromagnetic fields. For example, under the Lorentz boost the fields transform as

$$\vec{E}' = \gamma\left(\vec{E} + \vec{\beta} \times \vec{B}\right) - \frac{\gamma^2}{\gamma + 1}\vec{\beta}\left(\vec{\beta} \cdot \vec{E}\right), \tag{5.1}$$

$$\vec{B}' = \gamma\left(\vec{B} - \vec{\beta} \times \vec{E}\right) - \frac{\gamma^2}{\gamma + 1}\vec{\beta}\left(\vec{\beta} \cdot \vec{B}\right). \tag{5.2}$$

Then we can find $\vec{\beta}$ such that $\vec{E}' \times \vec{B}' = 0$. Notice that the solution is not unique because even if $\vec{E} \parallel \vec{B}$, we can still boost the frame along the $\vec{E}$ (or $\vec{B}$) direction while keeping they are parallel. Thus we choose the boost direction is perpendicular to the E and B directions, i.e., $\vec{\beta} = \alpha\left(\vec{E} \times \vec{B}\right)$. From the above equations, we have

$$\vec{E}' \times \vec{B}' = \gamma\left(\vec{E} \times \vec{B}\right)\left(1 - \alpha\left(E^2 + B^2\right) - \alpha^2\left(\vec{E} \times \vec{B}\right)^2\right) = 0, \tag{5.3}$$

which gives

$$\frac{\alpha}{1 + \alpha^2\left(\vec{E} \times \vec{B}\right)^2} = \frac{1}{E^2 + B^2}. \tag{5.4}$$

Note that there is one exceptional case such that $\vec{E} \perp \vec{B}$ and $E = B$.

where $\sigma^{\mu\nu} \equiv \frac{i}{2}[\gamma^\mu, \gamma^\nu]$. It is convenient to express the log in form of a Frullani integral:

$$W[A] \equiv \int d^4x \, \mathcal{L}_{eff} = \frac{i}{2} \int_0^\infty \frac{ds}{s} e^{-is[m^2-i\epsilon]} \text{Tr} \left\langle x \left| e^{is\left[(i\partial-eA)^2+\frac{e}{2}\sigma F-m^2\right]} - e^{-is\partial^2} \right| x \right\rangle,$$
(5.10)

where $|x\rangle$ is the eigenvector of the coordinate operator $\hat{X}|x\rangle = x|x\rangle$ normalized by

$$\langle x|x'\rangle = \delta^4(x-x'), \qquad \int d^4x \, |x\rangle\langle x| = I.$$

We can calculate the exponential term in Eq. (5.10) explicitly when the electric and magnetic fields are time-independent and uniform, and we obtain

$$\left\langle x \left| e^{is\left[(i\partial-eA)^2+\frac{e}{2}\sigma F-m^2\right]} \right| x \right\rangle = -\frac{ie^2}{4\pi^2} EB \coth(seE) \cot(seB).$$
(5.11)

When we put this back into Eq. (5.10), it is necessary to subtract the small $s$-divergence terms, which corresponds to the terms in the $F \to 0$ limit

$$-\frac{i}{s^2}\left[1 + \frac{e^2s^2}{3}\left(E^2 - B^2\right)\right].$$
(5.12)

Finally we get

$$\mathcal{L}_{eff} = \frac{1}{8\pi^2} \int_0^\infty \frac{ds}{s^3} e^{-is[m^2-i\epsilon]} \left[e^2s^2 EB \coth(seE) \cot(seB) - 1 - \frac{e^2s^2}{3}\left(E^2 - B^2\right)\right],$$
(5.13)

and we find the decay rate of the vacuum per unit volume

$$\frac{\Gamma}{V} = 2\text{Im}\mathcal{L}_{eff} = \frac{1}{4\pi^2}\text{Im} \int_0^\infty \frac{ds}{s^3} e^{-is[m^2-i\epsilon]}$$
(5.14)

$$\times \left[e^2s^2 EB \cot(seB) \coth(seE) - 1 - \frac{e^2s^2}{3}\left(E^2 - B^2\right)\right]$$
(5.15)

This integral can be calculated by taking a integration contour to enclose the negative imaginary axis and picking up the contribution of the poles at $s = n\frac{\pi}{eB}$ and $-in\frac{\pi}{eE}$, $n = 1, 2, \cdots$. Then we find the decay rate

$$\frac{\Gamma}{V} = \frac{e^2 EB}{4\pi^2} \sum_{n=1}^\infty \frac{1}{n} \coth\left(n\frac{\pi B}{E}\right) e^{-n\pi m^2/eE}.$$
(5.16)

## 5.2  Quantum Vlasov Equation

Although the operator approach is an elegant way to calculate the decay rate, it is limited to the time-independent uniform electromagnetic fields, and the back-reaction is ignored. On the other hand, the quantum-Vlasov equation [267], which consists of a set of non-linear differential equations, can naturally employ time-dependent fields with back-reactions. Here we derive the Quantum-Vlasov for a charged fermion in uniform time-dependent electric fields $E(t)\hat{z}$.

The fermion field is quantized as

$$\psi(x) = \sum_{s=-1,1} \int \frac{d^3p}{(2\pi)^3} \left[ a_{\vec{p},s} u_{\vec{p},s}(t) e^{i\vec{p}\cdot\vec{x}} + b_{\vec{p},s}^\dagger v_{-\vec{p},s}(t) e^{-i\vec{p}\cdot\vec{x}} \right], \tag{5.17}$$

with the commutation relation

$$\left\{ a_{\vec{p},s}, a_{\vec{p'},s'}^\dagger \right\} = \left\{ b_{\vec{p},s}, b_{\vec{p'},s'}^\dagger \right\} = (2\pi)^3 \, \delta^{(3)} \left( \vec{p} - \vec{p'} \right) \delta_{ss'}, \tag{5.18}$$

and the wave function normalization

$$u_{\vec{p},s}^\dagger(t) u_{\vec{p},s'}(t) = v_{\vec{p},s}^\dagger(t) v_{\vec{p},s'}(t) = \delta_{ss'}, \tag{5.19}$$

where the mode functions[2] are written in the gauge $A_0 = 0$ as [268, 269]

$$u_{\vec{p},s}(t) = (D+m)\,\phi_{\vec{p}}^+(t)\varepsilon_s, \tag{5.20}$$

$$v_{\vec{p},s}(t) = (D+m)\,\phi_{\vec{p}}^-(t)\varepsilon_s, \tag{5.21}$$

and

$$D \equiv i\gamma^0\partial_t - \gamma^1 p_x - \gamma^2 p_y - \gamma^3(p_z - eA_z(t)), \tag{5.22}$$

$$\varepsilon_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \qquad \varepsilon_{-1} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \end{pmatrix}. \tag{5.23}$$

---

[2]We have used the Dirac representation for spinors.

The scalar functions $\phi^\pm$ are positive and negative modes that correspond to mode functions for particle and anti-particle, respectively, and they satisfy

$$\left(\partial_t^2 + p_x^2 + p_y^2 + (p_z - eA_z(t))^2 + m^2 + ieE\right)\phi_{\vec{p}}^\pm(t) = 0. \tag{5.24}$$

Note that the field normalization (5.19) is insufficient to determine $\phi_{\vec{p}}^\pm$. Thus we put extra information to fix the positive and negative modes following the adiabatic prescription. We use an WKB (adiabatic) mode function

$$\phi_{\vec{p}}^\pm(t) = \frac{1}{\sqrt{2\omega\left(\omega \mp \pi_z\right)}}e^{\mp i\Theta}, \tag{5.25}$$

$$\Theta(t, t_0) = \int_{t_0}^t \omega(t', \vec{p})dt', \tag{5.26}$$

$$\omega^2(t, \vec{p}) = p_x^2 + p_y^2 + (p_z - eA_z(t))^2 + m^2, \tag{5.27}$$

$$\pi_z(t, p_z) = p_z - eA_z(t), \tag{5.28}$$

to choose the positive and negative mode functions for a given time $t$. In the asymptotic limit, $t \to -\infty$(or $\infty$), they correspond to $\pm$ mode functions for the "in" (or "out") vacuum. Also, when $E \to 0$, the mode functions are equivalent with those in the usual vacuum solutions. However, this choice is not unique, and even the parametrization itself is arbitrary as long as mode function recovers to a vacuum mode function as external fields vanish. We denote mode functions and creation and annihilation operators associated with the "in" vacuum by superscript "in". Otherwise, they are associated with the WKB (adiabatic) vacuum $|0; t\rangle$. Using the adiabatic mode functions, we can write field $\psi$ as

$$\psi(x) = \sum_{s=-1,1} \int \frac{d^3p}{(2\pi)^3}\left[a_{\vec{p},s}^{in}u_{\vec{p},s}^{in}(t)e^{i\vec{p}\cdot\vec{x}} + b_{\vec{p},s}^{in\dagger}v_{-\vec{p},s}^{in}(t)e^{-i\vec{p}\cdot\vec{x}}\right] \tag{5.29}$$

$$= \sum_{s=-1,1} \int \frac{d^3p}{(2\pi)^3}\left[a_{\vec{p},s}(t)u_{\vec{p},s}(t)e^{i\vec{p}\cdot\vec{x}} + b_{\vec{p},s}^\dagger(t)v_{-\vec{p},s}(t)e^{-i\vec{p}\cdot\vec{x}}\right]. \tag{5.30}$$

The mode functions are related by the Bogoliubov coefficients

$$u^{in}_{\vec{p},s}(t) = \alpha_{\vec{p}}(t)u_{\vec{p},s}(t) + \beta^*_{\vec{p}}(t)v_{\vec{p},s}(t), \tag{5.31}$$

$$a_{\vec{p}}(t) = \alpha_{\vec{p}}(t)a^{in}_{\vec{p}} + \beta_{\vec{p}}(t)b^{in\dagger}_{-\vec{p}}, \tag{5.32}$$

$$b^{\dagger}_{-\vec{p}}(t) = \alpha^*_{\vec{p}}(t)b^{in\dagger}_{-\vec{p}} + \beta^*_{\vec{p}}(t)a^{in}_{\vec{p}}, \tag{5.33}$$

where the spin indexes are suppressed because spins do not affect the relation when $B = 0$. The field normalization (5.19) and the adiabatic vacuum choice automatically yield the relation

$$\left|\alpha^2_{\vec{p}}(t)\right| + \left|\beta^2_{\vec{p}}(t)\right| = 1. \tag{5.34}$$

Furthermore, the field equation (5.24) gives

$$\dot{\alpha}_{\vec{p}}(t) = \frac{eE\epsilon_\perp}{2\omega^2}\beta_{\vec{p}}(t)e^{2i\Theta}, \tag{5.35}$$

$$\dot{\beta}_{\vec{p}}(t) = -\frac{eE\epsilon_\perp}{2\omega^2}\alpha_{\vec{p}}(t)e^{-2i\Theta}, \tag{5.36}$$

where $\epsilon_\perp \equiv \sqrt{p_x^2 + p_y^2 + m^2}$.

Now we define the particle (momentum) distribution function as

$$n_{\vec{k}}(t) = \frac{1}{V}\sum_s \left\langle 0; in \left| a^{\dagger}_{\vec{k},s}(t)a_{\vec{k},s}(t)\right| 0; in \right\rangle = 2\left|\beta_k(t)\right|^2, \tag{5.37}$$

where we have assumed $\lim_{t\to-\infty} n_{\vec{k}}(t) = 0$. Using the evolution equations (5.35) and (5.36) of $\alpha$ and $\beta$, we have

$$\dot{n}_{\vec{k}}(t) = \frac{eE(t)\epsilon_\perp(t)}{\omega^2(t,\vec{k})}\int_{t_0}^t dt' \frac{eE(t')\epsilon_\perp(t')}{\omega^2(t',\vec{k})}\left(1 - n_{\vec{k}}(t')\right)\cos 2\Theta(t,t'), \tag{5.38}$$

which is called "Quantum Vlasov" equation [196,267,270]. We can rewrite this integro-differential equation in a set of ordinary differential equations, which is more convenient for numerical computation

$$\dot{n}_{\vec{k}}(t) = \frac{eE(t)\epsilon_\perp(t)}{\omega^2(t,\vec{k})}f_{\vec{k}}(t), \tag{5.39}$$

$$\dot{f}_{\vec{k}}(t) = \frac{eE(t)\epsilon_\perp(t)}{\omega^2(t,\vec{k})}(1 - n_{\vec{k}}(t)) - 2\omega(t,\vec{k})g_{\vec{k}}(t), \tag{5.40}$$

$$\dot{g}_{\vec{k}}(t) = 2\omega(t,\vec{k})f_{\vec{k}}(t). \tag{5.41}$$

From the result (5.38), we can calculate the induced current

$$j_z(t) \;=\; e\left\langle \bar{\psi}\gamma^3\psi \right\rangle \tag{5.42}$$

$$=\; 2e\sum_s \int \frac{d^3k}{(2\pi)^3}\frac{k_z - eA_z(t)}{\omega(t,\vec{k})}n_{\vec{k}}(t) \tag{5.43}$$

$$+\frac{2}{E(t)}\sum_s \int \frac{d^3k}{(2\pi)^3}\omega(t,\vec{k})\frac{d}{dt}n_{\vec{k}}(t)$$

$$\equiv\; j_{cond}(t) + j_{pol}(t). \tag{5.44}$$

With Eq. (5.38), the current equation (5.43) and the Maxwell's equation

$$\ddot{A}_z(t) = -\dot{E}(t) = j_z(t), \tag{5.45}$$

we have a closed nonlinear integro-differential equation that includes back-reactions of produced pairs. Numerical and analytic behaviors of solutions of this equation has been studied extensively in literature (See, e.g., [266, 267, 270–279]). These numerical and analytic investigations show that in the Markovian limit [270] (when $E$ varies slow enough and the particle occupation number is negligible), the particle production rate is well approximated in forms of

$$\frac{d}{dt}n_{\vec{k}}(t) \;=\; \left|\beta_{\vec{k}}(\infty)\right|^2 |eE(t)|\,\delta(p_z), \tag{5.46}$$

$$\equiv\; S(\vec{k};t) \tag{5.47}$$

where $\vec{p}$ is the kinetic momentum $\vec{p} = \vec{k} - e\vec{A}$, and $\beta_{\vec{k}}(\infty)$ is the Bogoliubov coefficient relating the "in" and "out" vacua, for a slowly varying electric field

$$\left|\beta_{\vec{k}}(\infty)\right|^2 \approx \exp\left[-\pi\frac{k_x^2 + k_y^2 + m^2}{eE}\right]. \tag{5.48}$$

Eq. (5.46) is often written in the terms of kinetic momenta

$$\left(\frac{\partial}{\partial t} \pm eE\frac{\partial}{\partial p_z}\right)f^{\pm}(\vec{p},t) = S(\vec{k},t), \tag{5.49}$$

where $f^{\pm}(\vec{p}, t) = n_{\vec{p} \mp e\vec{A}}(t)$, which looks more like a Boltzmann-Vlasov equation. Also from (5.38) we get the pair production rate

$$\frac{d}{dt} N_{pair} = 2 \int \frac{d^3 k}{(2\pi)^3} \dot{n}_{\vec{k}}(t) \approx \frac{2}{(2\pi)^3} \left( eE \right)^2 \exp \left[ -\pi \frac{m^2}{eE} \right], \qquad (5.50)$$

where the factor 2 is from the spin degrees of freedom. This result is consistent with Eq. (5.16).

# Chapter 6

# Nanoshots from the Crab and Schwinger Sparks

The highest brightness temperature ever measured are from nanoshots from the Crab pulsar which could be the signature of bursts of vacuum pair production. If so, this would be the first time that the astronomical Schwinger effect has been observed. In this work, we consider a toy model called a "Schwinger spark chamber" producing a short-period emission, which may be responsible for the nanoshots. This model supposes a growing electric field in a vacuum of a finite volume. When the field reaches a threshold just below the Schwinger limit, it turns on a current of $e^{\pm}$ pairs swiftly, and this rapid transition can generate a radio pulse. Furthermore, using this model we show that the Schinwger spark chamber leads to a narrow electromagnetic pulses, which indicates that the Schwinger effect can be a feasible origin of the nanoshots.

This work was performed in collaboration with Albert Stebbins.

## 6.1   Introduction

Quantum electrodynamics (QED) has been very successful, and it has predicted many novel quantum effects, such as Casimir effect, Lamb shift, photon splitting and nonlinear Compton scattering, that have been experimentally tested [280–288]. The Schwinger effect (the vacuum $e^{\pm}$ pair production) is also one of long-standing predictions of QED since 1930s [264, 289, 290]. The effect predicts that the vacuum is unstable and decays into electron-positron pairs in the presence of strong electric field $E$ with the decay rate given by $\Gamma \sim e^{-\pi E_c/E}$ [264–266, 291, 292]. However, all attempts to observe this fundamental, non-perturbative effect have failed due to the experimental difficulties in generating an critical field strength of $E_c = \frac{m^2}{e}\frac{c^3}{\hbar} \sim 10^{16} V/cm$[1] although the recent rapid development of laser technology may achieve a direct observation of the Schwinger effect in near future [293, 294].

On the other hand, it is possible that compact astronomical objects possess extremely strong magnetic fields. In particular, pulsars, rapidly rotating and highly magnetized neutron stars, are believed to have a magnetic field in the range from $\sim 10^8$ to $10^{15}G$ near their surfaces. Because such a strong magnetic field may induce electric fields close to or even beyond $E_c$ naturally, pulsars are often considered as a ideal laboratory to test the Schwinger effect. In spite of such a optimistic prospect, the Schwinger effect has not been sucessfully employed yet in modeling the magnetosphere structure of a pulsar [295][2]. For instance, in the simple magnetic dipole model [298–300] , an induced electric field from a rotating magnetic field is suppressed by $R_s\Omega/c$, where $R_s$ is the pulsar radius, and $\Omega$ is the angular velocity of a pulsar. Thus the electric field $E$ rarely becomes close to $E_c$ in the magnetic dipole model.

---

[1]Equivalently, $B_c \approx 4 \times 10^{13}G$ in CGS units

[2]There were a couple of attempts to utilize the Schwinger effect to explain regular pulses of pulsars [296] and to obtain electron-positron plasma from strong intrinsic magnetic field of pulsars [297].

Moreover, the force-free approximation, $\vec{E} = -\left(\vec{\Omega} \times \vec{r}\right) \times \vec{B}$, inevitably blocks the Schwinger effect [301] (See also [302–305] for a review).

Nonetheless, such a strong electric field close to $E_c$ is believed to be involved in some phenomena of pulsars, such as Giant Radio Pulses (GPs). GPs have been observed from a variety of pulsars including the Crab pulsar. The GPs are at least thousand times brighter than regular pulses and have duration $\sim 5\mu s$. They consist of several distinct groups of micro-bursts of width $\sim 1\mu s$, and a micro-bust contains short-lived, relatively narrow-band "nanoshots" whose width $\lesssim 1\,ns^3$. The observed peak flux density $S_\nu$ of nanoshots goes up to $150\,kJy$ in a band $\nu = 8.5 \pm 0.2\,GHz$. Because the Crab pulsar is known to be at a distance $D_{crab} \approx 2.2\,kpc$ from Earth [308], the corresponding radio bolometric luminosity is

$$\dot{\mathcal{E}} \approx 4\pi D_{crab}^2 S_\nu \delta\nu \gtrsim 10^2 L_\odot. \tag{6.1}$$

Moreover, the nanoshot duration $\delta t$ suggests the diameter of the emission volume to be $l \lesssim 30\,cm$, and it indicates a peak brightness temperature

$$kT_\nu^{peak} \simeq \frac{1}{2}\left(\frac{1}{\nu}\right)^2 S_\nu \left(\frac{D_{Crab}}{2l}\right)^2 \gtrsim 10^{23}\,erg, \tag{6.2}$$

where $l \sim \delta t \approx 30\,cm$. This number, millions of times the Planck energy, requires macroscopic coherent emission by large number of charged particles moving together rather than a sum of microscopic single particle emission. Hence the propagating electric field at emission is

$$\delta E \approx \sqrt{\frac{8\pi\dot{\mathcal{E}}\delta t}{l^3}} \gtrsim 6 \times 10^{14}\,V/cm, \tag{6.3}$$

which is about a order of magnitude smaller than $E_c$ and also similar to the "typical" magnetic field at the pulsar surface. This implies that nanoshots can be associated

---

[3] According to [306, 307], some unresolved nanoshots have duration $\delta t \lesssim 0.4\,ns$.

with electric fields close to $E_c$, and also the similar magnitude of $\delta E$ and $B_{crab}$ hints the possibilities of some generation mechanism of the electric fields.

Furthermore, the Schwinger effect naturally produces a short-period emission (a spark) because it is a threshold phenomenon. The Schwinger effect is not effective initially when the electric field $E$ is far below $E_c$ such that $\Gamma$ is too small to affect $E$. Growing gradually, $E$ eventually reaches a limiting value $E_{lim}$ just below $E_c$, at which point the pair current suddenly overwhelms the increasing $E$, and the field cannot increase beyond $E_{lim}$. If the Schwinger effect takes place over a large enough volume, then the pair current can actually reverse the field creating a persistent pair plasma, and $E$ will oscillate and dissipate [266, 267]. However for small volumes, one does not have such oscillations, and the Schwinger effect merely limits the magnitude of the field excursion. Consequently, this causes a change of the current density once for a short time, which produces a spark, which we call a "Schwinger spark".

This work presented in the following orders. In Sec. 6.2, we present a toy model for generating a spark using the Schwinger effect and estimate the luminosity and the width of a spark, and we derive a consistency relation. In Sec. 6.3, we apply the result from the toy model to nanoshots from the Crab pulsar and discuss the possibilities of strong electric fields. Finally, we conclude in Sec. 6.4. we are going to use the natural units $\hbar = c = \varepsilon_0 = 1$ in the following sections.

## 6.2   Toy Model: Schwinger Spark Chamber

Suppose that there is a parallel plate vacuum capacitor connected to a power supply. Because we have its application to a pulsar in mind, we assume a uniform magnetic field $\vec{B}$ also although the result would not be altered much without it. The magnetic field $\vec{B}$ is along the direction of the electric field $\vec{E}$ from charges on the plates.

Whether it is aligned or anti-aligned does not matter in the following consideration. Furthermore, because the diameter of the plates are larger than the distance $w$ between them, the electric field in the capacitor is almost uniform.

As the power supply begins to charge the capacitor, the electric field inside the capacitor grows as well. In spite of continuous piling charges up on both plates, the electric field cannot grows to infinity due to the production of electron and positron pairs when it reaches close to the Schwinger limit $E_c$. Once the pair production starts to yield a current inside the capacitor. The induced current grows very rapidly and becomes saturated at a value just enough to neutralize the current from the power supply. The rapid emergence of the current from pair production at the transition emits a narrow radio pulse, which looks just like a spark.

This model can be translated into a simple one-dimensional problem[4], and the behavior of the electric field inside the capacitor described by the differential equation

$$-\frac{d}{dt}E(t) = j_{pp}(t) + j_{ext}(t) \equiv j_{tot}(t), \tag{6.4}$$

where $E$ is the electric field inside the capacitor, $j_{ext}$ is the current density generated from the power supply, and $j_{pp}$ is the induced one by the pair production given by

$$j_{pp}(t) \approx 2eS\left[E, B\right]w = 2\frac{e^3 EB}{(2\pi)^2}we^{-\pi\frac{m^2}{eE}}. \tag{6.5}$$

We can see that $E$ becomes saturated to $E_*$ at some point $t_*$ unless $j_{ext}$ grows very quickly to the infinity. For a steady current $j_{ext} = const$, we explicitly find that

$$E_* \approx \begin{cases} \frac{|j_{ext}|}{\xi}E_c & \text{for } |j_{ext}| \gg \xi \\ \frac{\pi E_c}{\log\frac{\pi\xi}{|j_{ext}|}} & \text{for } |j_{ext}| \ll \xi \end{cases}, \quad t_* \approx \begin{cases} E_c/\xi & \text{for } |j_{ext}| \gg \xi \\ \frac{\pi E_c}{|j_{ext}|\log\frac{\pi\xi}{|j_{ext}|}} & \text{for } |j_{ext}| \ll \xi \end{cases}, \tag{6.6}$$

where

$$\xi \equiv 2w\frac{e^3 E_c B}{(2\pi)^2}. \tag{6.7}$$

---

[4]The required conditions are discussed in Appendix. 6.5

Phenomenologically it may be difficult to achieve $E_* > E_c$, which requires an enormously large external current density $|j_{ext}| \gg \xi$ [5]. On the other hand, a weak external current makes $E$ increase gradually and reach below $E_c$ if the external current persists sufficiently long $t > t_*$. More precisely, for the typical parameters of the consideration, $w = 30\,cm$ and $B \sim 0.1 \times B_c$, the electric field $E_*$ becomes approximately

$$E_* \sim 0.05 E_c \times O(1) \equiv E_{lim}, \tag{6.8}$$

The $O(1)$ factor is from the logarithmic dependence of the external current $j_{ext}$, and hence we expect that the typical value $E_*$ is $\sim 0.05 E_c$.[6]

Through the transition from vacuum to the electron-positron plasma, there is a sudden change of the total current density $j_{tot}$, which eventually gives rise to a spark. Now we estimate the peak of $dj_{tot}/dt$ and its width $\delta t_*$ near the transition. Firstly, taking a derivative of Eq. (6.4) with respect to $t$, we have

$$\frac{d}{dt} j_{tot} = \frac{j_{tot}}{E} \left( j_{tot} - j_{ext} \right) \left( 1 + \pi \frac{E_c}{E} \right) + \frac{d}{dt} j_{ext}. \tag{6.11}$$

Neglecting $\frac{d}{dt} j_{ext}$ and treating $E$ as a constant at the transition, we can solve Eq. (6.11) explicitly and find

$$j_{tot}(t) = \frac{j_{ext}^*}{1 + \exp\left(t - t_*\right)/\delta t_*}, \tag{6.12}$$

---

[5]For $B = 0.1 B_c$ and $w = 30\,cm$, $\xi \simeq 3.4 \times 10^{51} e \cdot cm^{-2} s^{-1}$

[6]For example, near a surface of a pulsar, the typical current density may be estimated from the Goldreich-Julian density [301]

$$j_{ext} \quad \sim \quad e\, n_{GJ} \sim B/P, \tag{6.9}$$

where $B$ is the magnetic field and and $P$ is the rotation period. For the Crab pulsar, the peak surface magnetic field $B_p \sim 3 \times 10^{13}\,G$ and the rotation period $P \sim 33\,msec$,

$$E_* \sim 0.046 \times E_c. \tag{6.10}$$

Also, for a typical current density in a tabletop experiment $j_{ext} \sim 10^{19} e\,cm^{-2} s^{-1}$,

$$E_* \sim 0.042 \times E_c.$$

where $j_{ext}^*$ is the external current density at the transition, and

$$\delta t_* \equiv \frac{1}{\pi} \left| \frac{E_*^2}{E_c j_{ext}^*} \right|. \tag{6.13}$$

The peak value of $\left| \frac{d}{dt} j_{tot} \right|$ near the transition is estimated as

$$\left| \frac{d}{dt} j_{tot} \right|_{t_{peak}} \approx \left| \frac{j_{ext}^{*2}}{4E_*} \left( 1 + \pi \frac{E_c}{E_*} \right) \right| \equiv \left| \dot{j}_{tot}^* \right|. \tag{6.14}$$

As shown in Fig. 6.1, $\dot{j}_{tot}$ clearly behaves as a sharp pulse at the transition, which can be interpreted as a spark.

From these we can find the luminosity and the duration of an induced electromagnetic pulse from $\dot{j}_{tot}$ at the transition. The Poynting vector from a localized and linear current is given by (See Appendix 6.6 for a derivation)

$$\vec{S} \simeq \frac{\vec{Q} \cdot \vec{Q} - \left( \hat{r} \cdot \vec{Q} \right)^2}{4\pi r^2} \hat{r}, \tag{6.15}$$

where

$$\vec{Q}(\hat{r}_0, \delta t) = \int_V d^3 \delta x' \, \dot{\vec{j}} \left( \vec{x}_0' + \delta \vec{x}', t_0' + \delta t - \hat{r}_0 \cdot \delta x' \right). \tag{6.16}$$

Suppose that $\dot{\vec{j}}_{tot}$ is uniform and also varies coherently over the volume of the Schwinger chamber. Putting Eq. (6.12) into Eq. (6.16) we obtain at the transition

$$\left| \vec{Q}(\hat{r}_0, \delta t) \right| = \frac{j_{ext}^* w^2 \sinh \frac{w}{2\delta t_*}}{\cosh \frac{w}{2\delta t_*} + \cosh \frac{\delta t}{\delta t_*}}, \tag{6.17}$$

where we have treated the chamber as a cubic box with side $w$ and $\hat{r}_0$ is aligned with one of the sides to simplify the integral and the estimation of the luminosity. From this, the maximum value of $Q$ and the width of the pulse $\Delta t$[7] are estimated as

$$Q_{max} \approx j_{ext}^* w^2 \tanh \frac{w}{4\delta t_*}, \tag{6.18}$$

$$\Delta t \approx 2\delta t_* \cosh^{-1} \left[ 2 + \cosh \frac{w}{2\delta t_*} \right]. \tag{6.19}$$
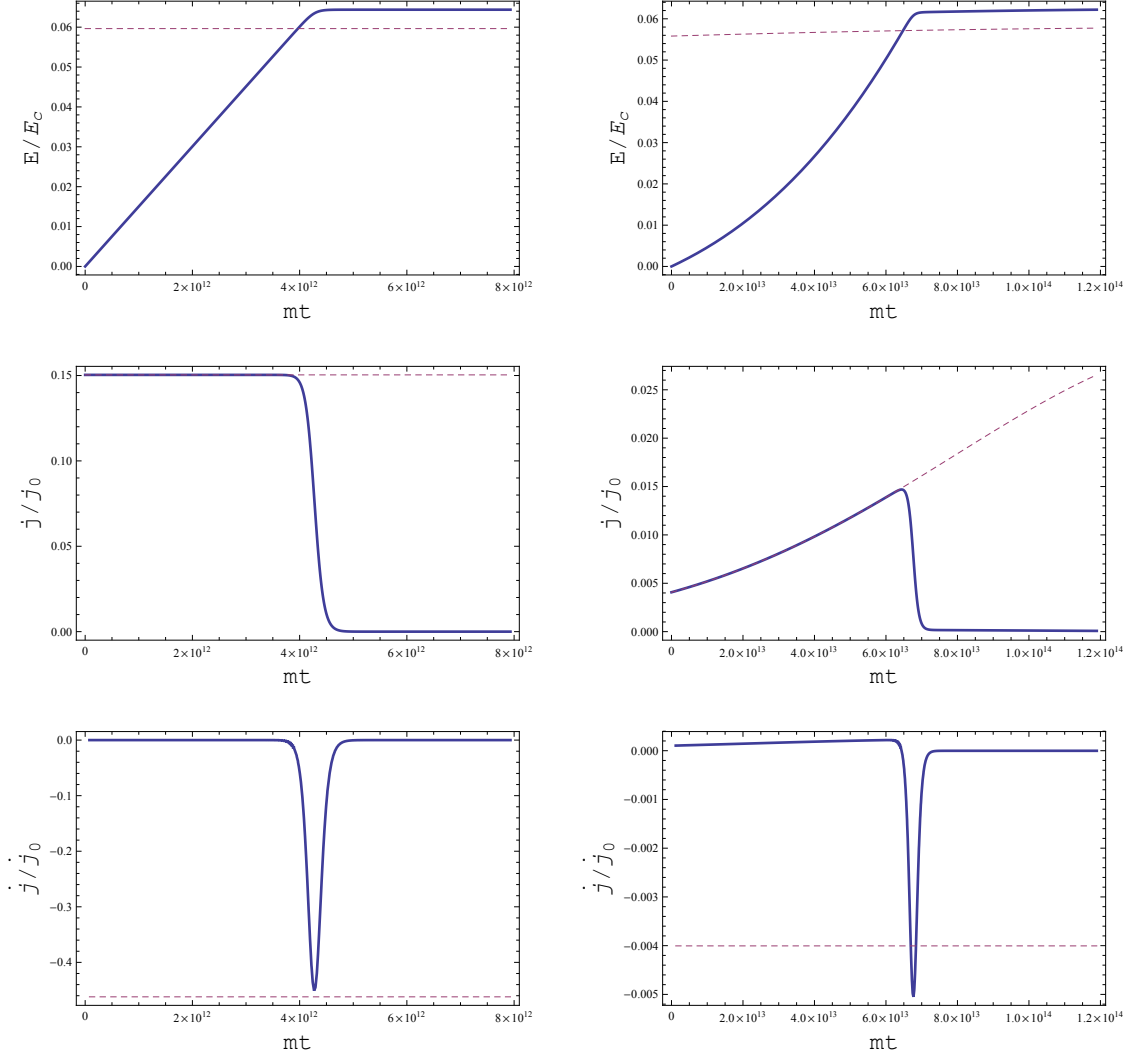
[7]Full width at half maximum

Figure 6.1: These plots shows how $\left\{ E, j_{tot}, \dot{j}_{tot} \right\}$ varies over time for a constant current (left) and a growing current (right), where $j_0 = 10^{-13} m E_c$ and $\dot{j}_0 = 10^{-25} m^2 E_c$. In each plot, the dashed lines denotes the expected $E_*$ from Eq. (6.8), the external current $j_{ext}$, and the peak of $\dot{j}_{tot}$ from Eq. (6.14). At the transition $t = t_*$, the field $E$ becomes saturated to $E_*$, the current $j$ suddenly drops to 0, and $\dot{j}_{tot}$ shows a pulse-like behavior.

Using Eqs. (6.13) and (6.18) we obtain

$$
L_{max} \;=\; \int d\Omega \vec{S}_{max} \cdot \vec{n} \tag{6.20}
$$

$$
\simeq \; \frac{2}{3}\left[\frac{1}{\pi}\frac{E_*^2}{E_c \delta t_*}w^2 \tanh\frac{w}{4\delta t_*}\right]^2. \tag{6.21}
$$

Notice that there are only two independent model parameters among $\{E_*, w, \delta t_*\}$ if the change of $j_{ext}$ during the transition is negligible. For example, we can obtain $\delta t_*$ from $E_*$ and $w$ using Eq. (6.6). Hence, for given values of $E_*$ and $w$, Eq. (6.21) and Eq. (6.19) determines the maximum luminosity and the duration of a electromagnetic pulse induced by the transition, $L_{max}$ and $\Delta t$. Note also that the toy model analysis has been done in the particular frame where the external electric field and magnetic field are parallel. Generally, we can boost a frame to have the external electric field and the magnetic field parallel, but it yields a boosting factor, which is neglected in this analysis. In the next section, we are going to apply the relation to nanoshots from the Crab pulsar.

## 6.3 Nanoshots from Schwinger Sparks

In Ref. [309], nanoshots from the Crab pulsar have been observed with a peak flux up to $S_\nu \simeq 150\,kJy$ in a band $\nu = 8.5 \pm 0.2\,GHz$. Its typical flux $S_\nu$ is $\sim 10\,kJy$ and its duration $\Delta t$ is $\sim 1\,ns$. The peak luminosity obtained from $L_{max} \simeq 4\pi D_{crab}^2 S_\nu \delta\nu$. Using (6.19) and (6.21) with the observation of the peak flux density and the duration of a nanoshot, $S_\nu$ and $\Delta t$ , we can contrain the parameters of the toy model, the electric field at the transition and the width of the Schwinger chamber, $E_*$ and $w$, as shown in Fig. 6.2.

Fig. (6.2) shows that the Schwinger spark can cover quite wide ranges of the peak density flux and the duration, $S_\nu$ and $\Delta t$, including the ranges of the typical
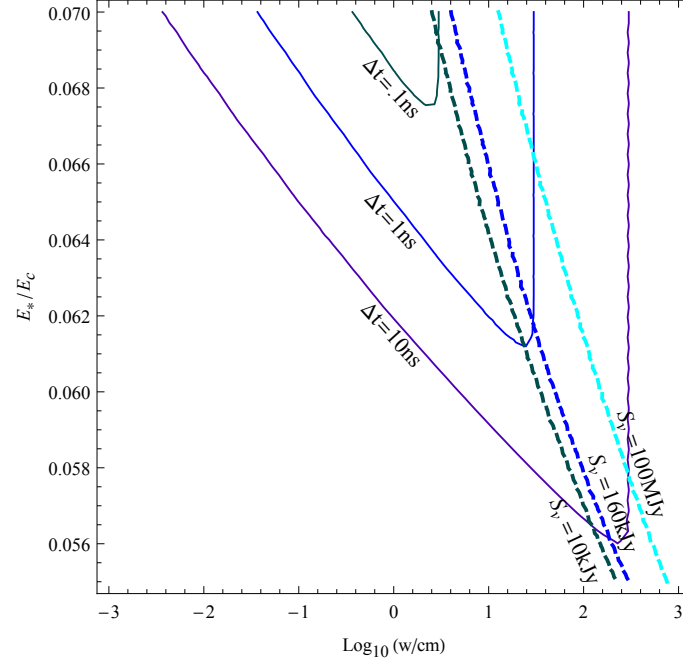
Figure 6.2: The peak flux density and the duration of an induced EM pulse from the Scwhinger spark. The solid lines represent constraints from the durations of pulses, and the dotted lines denote those from the peak flux densities.

nanoshots. This indicates that the Schwinger spark can indeed be an origin of the nanoshots from the Crab pulsars. Moreover, for $S_\nu \sim 40 - 160\, kJy$ and $\Delta t \sim 1\, ns$ implies that $\delta t_* \sim \Delta t \sim w$.

At this point, we should mention that the generation mechanism for such large external coherent electric fields inducing the Schwinger sparks within an appropriate size of volumes is still in question. However, the similarity in magnitude of $E_*$ and $B_{crab}$ itself is suggestive that nanoshots are generated near the neutron star surface and related to transients in the electromagnetic (EM) field of the neutron star. While such EM fields are not part of normal pulsar modeling [301, 303–305], it would not be surprising to find this phenomena in young pulsars such as the Crab. Highly magnetized and rapidly rotating neutron stars are out of thermodynamic balance with their environment and will utilize all available channels to equilibrate; shedding

energy, angular momentum and magnetic field. The initial magnetic field frozen into the neutron star material could be quite convoluted with large field variations on fairly small scales since high multipole fields do not contribute significantly to rotational energy loss and are unconstrained by observations. Small magnetic field loops may be shaken or if loosely tethered even become detached. Similar surface phenomena power stellar flares and and soft gamma repeaters.

## 6.4   Conclusion

In this work, we have considered a possibility that the Schwinger effect is responsible for the nanoshot phenomena in the emission from the Crab pulsar. We build a toy model using the Schwinger effect that leads to a short-period electomagnetic pulse. Moreover, we show that this model can explain nanoshots from the Crab, which suggests that the Schwinger effect gives rise to the nanoshots. If so, they are the first direct evidence of the Schwinger effect. Furthermore, it implies the existence of extremely high energy $e^{\pm}$ pairs as well. Produced $e^{\pm}$ pairs within an emission volume gain energy $\sim \frac{1}{2} e E_* w \approx 10 \, PeV$ as they escape the volume. That can make our model more interesting as the pairs may trigger the pair cascades in the pulsar magnetosphere [302, 310, 311] or provide a novel method to examine BSM physics. We postpone the study of the high energy $e^{\pm}$ pairs for our future work.

## 6.5   Appendix: Setup for a Schwinger Spark Chamber

We consider a system with inhomogeneous background the electromagnetic fields in the frame that $\vec{E} \parallel \vec{B}$. The fields are non-vanishing within a cylindrical region

with radius $r$ and width $w$. Both electric and magnetic fields are aligned along the symmetrical axis of the cylinder, and due to the external charge and current configuration, the electric field outside the region is negligible, $E_{out} \ll m^2/e$ , i.e,

$$E(t,\vec{x}) = \begin{cases} E(t) & \text{if } \vec{x} \in \text{Cylinder} \\ 0 & \text{otherwise} \end{cases}, \tag{6.22}$$

and the magnetic field lines are diverging smoothly.

Then the total pair production rate within the region is given by

$$\mathcal{N}_{tot} = S\,[E,B]\,V_{cyl}, \tag{6.23}$$

$$V_{cyl} \equiv \pi r^2 w, \tag{6.24}$$

where $S \approx \frac{eEB}{(2\pi)^2}e^{-\pi\frac{m^2}{eE}}$ is a pair production rate per unit volume [264–266]. Because we assume that $E$ and $B$ are sufficiently large such that most of pairs in the region are relativistic[8] and they move along the magnetic field lines, the current density generated by pair production is [264–266, 291, 292]

$$j_{pp} \approx 2eS\,[E,B]\,w = 2w\frac{e^3EB}{(2\pi)^2}e^{-\pi\frac{m^2}{eE}}. \tag{6.26}$$

Note that we have taken $E$, and $B$ are positive. This induced current back-reacts to the electric field in the region. As long as the width of the region is not too narrow, $w \gtrsim r$, from the Maxwell's equation the back-reaction due to the induced current is given by

$$\frac{d}{dt}E = -j_{pp} - j_{ext}, \tag{6.27}$$

---

[8]This assumption implies that the polarization current is always sub-dominant compared to the conducting current. To be precise, the polarization and the conducting current currents are

$$j_{pol} = \frac{2e^2Bm}{(2\pi)^2}e^{-\pi\frac{m^2}{eE}}, \quad j_{cond} = \frac{2e^3EB}{(2\pi)^2}e^{-\pi\frac{m^2}{eE}}\,\delta t, \tag{6.25}$$

where $\delta t$ is the average time pairs staying in the region, $\delta t \sim w$. The assumption of the relativistic current is valid if $eE\delta t \gtrsim m$, which guarantees that the conducting current is larger than the polarization current, $j_{pp} \approx j_{cond}$. Moreover, the condition $eE\delta t \gtrsim m$ is well satisfied for the typical values we consider, $E \sim 0.05E_c$ and $w \sim 30\,cm$.

where $j_{ext} < 0$ is the external current that sources the electric field energy in the region. Notice that the homogeneous solution of Eq. 6.27 is a monotonically decaying function unlike the usual oscillatory solution [266, 267, 270, 273]. This is due to the finiteness of the field configuration, which gives the conditions for the approximation,

$$\left|\frac{E}{\dot{j}_{pp}}\right| \gg w \implies \frac{1}{w^2} \gg 2\frac{e^3 B}{(2\pi)^2}e^{-\pi\frac{m^2}{eE}}, \tag{6.28}$$

$$\left|\frac{B}{2\pi\dot{j}_{pp}}\right| \gg r \implies \frac{1}{rw} \gg 2\frac{e^3 E}{(2\pi)}e^{-\pi\frac{m^2}{eE}}. \tag{6.29}$$

The first condition is to prevent the plasma oscillation, and the second one to restrict the induced magnetic field by the pair production to be weaker than the external magnetic field $B$. For the typical values in consideration $E \sim 0.05 E_c$ and $B \sim 0.1 B_c$, the conditions are

$$w \ll 4 \times 10^4 \, cm, \quad r \ll 8 \times 10^4 \, cm, \tag{6.30}$$

which shows the typical width for nanoshots $w \sim 30 \, cm$ is well within the range.

## 6.6   Appendix: Jefimenko's Equation

A general solution to Maxwell's equations is given by Jefimenko's equations [312]

$$\vec{E}(x) = \int d^3\vec{x}' \left(\frac{\rho(\vec{x}', t_{ret})}{|\vec{x} - \vec{x}'|^3} + \frac{\dot{\rho}(\vec{x}', t_{ret})}{|\vec{x} - \vec{x}'|^2}\right)(\vec{x} - \vec{x}') - \frac{\dot{\vec{j}}(\vec{x}', t_{ret})}{|\vec{x} - \vec{x}'|}, \tag{6.31}$$

$$\vec{B}(x) = \int d^3\vec{x}' \left(\frac{\vec{j}(\vec{x}', t_{ret})}{|\vec{x} - \vec{x}'|^3} + \frac{\dot{\vec{j}}(\vec{x}', t_{ret})}{|\vec{x} - \vec{x}'|^2}\right) \times (\vec{x} - \vec{x}'), \tag{6.32}$$

where we have dropped the homogeneous solutions, and $t_{ret} = t - |\vec{x} - \vec{x}'|$. Using the continuity equation

$$\dot{\rho} + \nabla \cdot \vec{j} = 0 \tag{6.33}$$

we can integrate by parts to find

$$
\vec{E}(x) \;=\; \int d^3 x' \left[ \left( \frac{\rho\left(\vec{x}',t_{ret}\right) + \dot{\vec{j}}\left(\vec{x}',t_{ret}\right)\cdot\left(\vec{x}-\vec{x}'\right) + 2\vec{j}\left(\vec{x}',t_{ret}\right)\cdot\frac{(\vec{x}-\vec{x}')}{|\vec{x}-\vec{x}'|}}{|\vec{x}-\vec{x}'|^3} \right)\left(\vec{x}-\vec{x}'\right) \right.
$$
$$
\left. -\frac{\vec{j}\left(\vec{x}',t_{ret}\right)}{|\vec{x}-\vec{x}'|^2} - \frac{\dot{\vec{j}}\left(\vec{x}',t_{ret}\right)}{|\vec{x}-\vec{x}'|} \right]. \tag{6.34}
$$

Then the far-field behavior is given by

$$
\vec{E}(x) \;=\; \int d^3 x' \frac{\dot{\vec{j}}\left(\vec{x}',t_{ret}\right)\cdot\left(-I+\hat{r}\otimes\hat{r}\right)}{|\vec{x}-\vec{x}'|} + O\left[\frac{1}{|\vec{x}-\vec{x}'|^2}\right] \tag{6.35}
$$

$$
\vec{B}(x) \;=\; \int d^3\vec{x}' \frac{\dot{\vec{j}}\left(\vec{x}',t_{ret}\right)}{|\vec{x}-\vec{x}'|}\times\hat{r}, \tag{6.36}
$$

where $\hat{r} = \left(\vec{x}-\vec{x}'\right)/|\vec{x}-\vec{x}'|$ is the radial unit vector.

Suppose the charges and currents are confined in a small region around a fiducial source event, $\left(\vec{x}_0',t_0'\right)$ and we wish to find the electric and magnetic field in small region around a distant fiducial observation event $\left(\vec{x}_0,t_0\right)$. We assume that $\left(\vec{x}_0,t_0\right)$ is in the forward light cone of $\left(\vec{x}_0',t_0'\right)$, i.e., $c\left(t_0-t_0'\right)=|\vec{x}_0-\vec{x}_0'|$. Define $r_0 \equiv |\vec{x}_0-\vec{x}_0'|\quad \hat{r}_0 \equiv \frac{\vec{x}_0-\vec{x}_0'}{|\vec{x}_0-\vec{x}_0'|}$ then we see that

$$
\vec{E}(x_0+\delta x) \;=\; \frac{1}{r_0}\int d^3\delta x'\, \dot{\vec{j}}\left(\vec{x}_0'+\delta\vec{x}',t_0'+\delta t-\hat{r}_0\cdot\left(\delta x-\delta x'\right)\right)\cdot\left(-I+\hat{r}_0\otimes\hat{r}_0\right)
$$
$$
+ O\left[\frac{1}{r_0^2}\right], \tag{6.37}
$$
$$
\vec{B}(x_0+\delta x) \;=\; \frac{1}{r_0}\int d^3\vec{x}'\, \dot{\vec{j}}\left(\vec{x}_0'+\delta\vec{x}',t_0'+\delta t-\hat{r}_0\cdot\left(\delta x-\delta x'\right)\right)\times\hat{r}_0 + O\left[\frac{1}{r_0^2}\right] \tag{6.38}
$$

where we have used $t_{ret} \approx t_0' + \delta t - \hat{r}_0\left(\delta x - \delta x'\right)$. Note that in this approximation

$$
\vec{E}(\vec{x}_0+\hat{r}_0\delta t, t_0+\delta t) \;=\; \vec{E}(\vec{x}_0,t_0) + O\left[\frac{1}{r_0^2}\right], \tag{6.39}
$$
$$
\vec{B}(\vec{x}_0+\hat{r}_0\delta t, t_0+\delta t) \;=\; \vec{B}(\vec{x}_0,t_0) + O\left[\frac{1}{r_0^2}\right], \tag{6.40}
$$

so this is an outgoing wave traveling in the $\hat{r}_0$ direction at the speed of light $c$. Furthermore, translating the observer perpendicular to the line-of-sight does not change the fields:

$$\vec{E}(\vec{x}_0 + \delta\vec{x}_\perp, t_0) = \vec{E}(\vec{x}_0, t_0) + O\left[\frac{1}{r_0^2}\right], \tag{6.41}$$

$$\vec{B}(\vec{x}_0 + \delta\vec{x}_\perp, t_0) = \vec{B}(\vec{x}_0, t_0) + O\left[\frac{1}{r_0^2}\right], \tag{6.42}$$

when $\delta x_\perp \cdot \hat{r}_0 = 0$. Without loss of generality, we can set $\delta\vec{x} = 0$. Moreover noting the integrals are the same in the expressions for the electric and magnetic fields we can write this solution more compactly

$$\vec{E}(\vec{x}_0, t_0 + \delta t) = \frac{1}{r_0}\left(-I + \hat{r}_0 \otimes \hat{r}_0\right) \cdot \vec{Q}(\hat{r}_0, \delta t) + O\left[\frac{1}{r_0^2}\right], \tag{6.43}$$

$$\vec{B}(\vec{x}_0, t_0 + \delta t) = \frac{1}{r_0}\vec{Q}(\hat{r}_0, \delta t) \times \hat{r}_0 + O\left[\frac{1}{r_0^2}\right], \tag{6.44}$$

where

$$\vec{Q}(\hat{r}_0, \delta t) = \int d^3\delta x' \, \dot{\vec{j}}\left(\vec{x}_0' + \delta\vec{x}', t_0' + \delta t - \hat{r}_0 \cdot \delta x'\right). \tag{6.45}$$

The Poynting vector is

$$\vec{S} = \frac{1}{4\pi}\vec{E} \times \vec{B} = \frac{\hat{r}_0}{4\pi r_0^2}\left(\left(\vec{Q} \cdot \vec{Q}\right) - \left(\hat{r}_0 \cdot \vec{Q}\right)^2\right). \tag{6.46}$$

which is of course directed directly away from the source and is zero if $\vec{Q}$ is pointed directly away or toward the source.

# Bibliography

[1]   S. Weinberg, *Cosmology*. Oxford University Press, USA, Apr., 2008.

[2]   D. H. Lyth and A. Riotto, "Particle physics models of inflation and the cosmological density perturbation," *Phys.Rept.* **314** (1999) 1–146, arXiv:hep-ph/9807278 [hep-ph].

[3]   A. R. Liddle and D. Lyth, "Cosmological inflation and large scale structure,".

[4]   A. Riotto, "Inflation and the theory of cosmological perturbations," arXiv:hep-ph/0210162 [hep-ph].

[5]   B. S. DeWitt, "Quantum Field Theory in Curved Space-Time," *Phys.Rept.* **19** (1975) 295–357.

[6]   N. Birrell and P. Davies, *Quantum Fields in Curved Space*. Cambridge Monographs on Mathematical Physics. Cambridge University Press, 1984.

[7]   L. E. Parker, "Quantum field theory in curved spacetime,".

[8]   A. A. Starobinsky, "A New Type of Isotropic Cosmological Models Without Singularity," *Phys.Lett.* **B91** (1980) 99–102.

[9]   A. H. Guth, "The Inflationary Universe: A Possible Solution to the Horizon and Flatness Problems," *Phys.Rev.* **D23** (1981) 347–356.

[10]  A. D. Linde, "A New Inflationary Universe Scenario: A Possible Solution of the Horizon, Flatness, Homogeneity, Isotropy and Primordial Monopole Problems," *Phys.Lett.* **B108** (1982) 389–393.

[11] A. Albrecht and P. J. Steinhardt, "Cosmology for Grand Unified Theories with Radiatively Induced Symmetry Breaking," *Phys.Rev.Lett.* **48** (1982) 1220–1223.

[12] A. D. Linde, "Monopoles as big as a universe and the universe inside a bubble," *Int.J.Mod.Phys.* **D5** (1996) 845–867.

[13] A. Vilenkin, "Topological inflation," *Phys. Rev. Lett.* **72** (May, 1994) 3137–3140. `http://link.aps.org/doi/10.1103/PhysRevLett.72.3137`.

[14] P. J. Steinhardt, "The inflation debate: Is the theory at the heart of modern cosmology deeply flawed?," *Sci.Am.* **304N4** (2011) 18–25.

[15] A. Ijjas, P. J. Steinhardt, and A. Loeb, "Inflationary paradigm in trouble after Planck2013," *Phys.Lett.* **B723** (2013) 261–266, `arXiv:1304.2785 [astro-ph.CO]`.

[16] A. Ijjas, P. J. Steinhardt, and A. Loeb, "Inflationary schism after Planck2013," `arXiv:1402.6980 [astro-ph.CO]`.

[17] S. Fulling, "REMARKS ON POSITIVE FREQUENCY AND HAMILTONIANS IN EXPANDING UNIVERSES," *Gen.Rel.Grav.* **10** (1979) 807–824.

[18] D. J. Chung and H. Yoo, "Isocurvature Perturbations and Non-Gaussianity of Gravitationally Produced Nonthermal Dark Matter," *Phys.Rev.* **D87** (2013) 023516, `arXiv:1110.5931 [astro-ph.CO]`.

[19] J. M. Maldacena, "Non-Gaussian features of primordial fluctuations in single field inflationary models," *JHEP* **0305** (2003) 013, `arXiv:astro-ph/0210603 [astro-ph]`.

[20] D. Seery and J. E. Lidsey, "Primordial non-gaussianities in single field inflation," *JCAP* **0506** (2005) 003, `arXiv:astro-ph/0503692`.

[21] V. Acquaviva, N. Bartolo, S. Matarrese, and A. Riotto, "Second order cosmological perturbations from inflation," *Nucl.Phys.* **B667** (2003) 119–148, `arXiv:astro-ph/0209156 [astro-ph]`.

[22] P. Creminelli, "On non-Gaussianities in single-field inflation," *JCAP* **0310** (2003) 003, `arXiv:astro-ph/0306122 [astro-ph]`.

[23] A. P. Yadav and B. D. Wandelt, "Evidence of Primordial Non-Gaussianity (f(NL)) in the Wilkinson Microwave Anisotropy Probe 3-Year Data at 2.8sigma," *Phys.Rev.Lett.* **100** (2008) 181301, `arXiv:0712.1148 [astro-ph]`.

[24] **WMAP Collaboration** Collaboration, E. Komatsu *et al.*, "Five-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation," *Astrophys.J.Suppl.* **180** (2009) 330–376, `arXiv:0803.0547 [astro-ph]`.

[25] J. Fergusson and E. Shellard, "The shape of primordial non-Gaussianity and the CMB bispectrum," *Phys.Rev.* **D80** (2009) 043510, `arXiv:0812.3413 [astro-ph]`.

[26] K. M. Smith, L. Senatore, and M. Zaldarriaga, "Optimal limits on $f_{NL}^{local}$ from WMAP 5-year data," *JCAP* **0909** (2009) 006, `arXiv:0901.2572 [astro-ph]`.

[27] A. P. Yadav and B. D. Wandelt, "Primordial Non-Gaussianity in the Cosmic Microwave Background," *Adv.Astron.* **2010** (2010) 565248, `arXiv:1006.0275 [astro-ph.CO]`.

[28] E. Komatsu, "Hunting for Primordial Non-Gaussianity in the Cosmic Microwave Background," *Class.Quant.Grav.* **27** (2010) 124010, `arXiv:1003.6097 [astro-ph.CO]`.

[29] A. Curto, E. Martinez-Gonzalez, R. Barreiro, and M. Hobson, "Constraints on general primordial non-Gaussianity using wavelets for the Wilkinson Microwave anisotropy probe 7-year data," `arXiv:1105.6106 [astro-ph.CO]`.

[30] M. LoVerde, A. Miller, S. Shandera, and L. Verde, "Effects of Scale-Dependent Non-Gaussianity on Cosmological Structures," *JCAP* **0804** (2008) 014, `arXiv:0711.4126 [astro-ph]`.

[31] A. Slosar, C. Hirata, U. Seljak, S. Ho, and N. Padmanabhan, "Constraints on local primordial non-Gaussianity from large scale structure," *JCAP* **0808** (2008) 031, `arXiv:0805.3580 [astro-ph]`.

[32]  N. Afshordi and A. J. Tolley, "Primordial non-gaussianity, statistics of collapsed objects, and the Integrated Sachs-Wolfe effect," *Phys.Rev.* **D78** (2008) 123507, `arXiv:0806.1046 [astro-ph]`.

[33]  S. Matarrese and L. Verde, "The effect of primordial non-Gaussianity on halo bias," *Astrophys.J.* **677** (2008) L77–L80, `arXiv:0801.4826 [astro-ph]`.

[34]  E. Sefusatti and E. Komatsu, "The bispectrum of galaxies from high-redshift galaxy surveys: Primordial non-Gaussianity and non-linear galaxy bias," *Phys.Rev.* **D76** (2007) 083004, `arXiv:0705.0343 [astro-ph]`.

[35]  N. Dalal, O. Dore, D. Huterer, and A. Shirokov, "The imprints of primordial non-gaussianities on large-scale structure: scale dependent bias and abundance of virialized objects," *Phys.Rev.* **D77** (2008) 123514, `arXiv:0710.4560 [astro-ph]`.

[36]  S. Shandera, N. Dalal, and D. Huterer, "A generalized local ansatz and its effect on halo bias," *JCAP* **1103** (2011) 017, `arXiv:1010.3722 [astro-ph.CO]`.

[37]  K. M. Smith and M. LoVerde, "Local stochastic non-Gaussianity and N-body simulations," `arXiv:1010.0055 [astro-ph.CO]`.

[38]  G. D'Amico, M. Musso, J. Norena, and A. Paranjape, "An Improved Calculation of the Non-Gaussian Halo Mass Function," *JCAP* **1102** (2011) 001, `arXiv:1005.1203 [astro-ph.CO]`.

[39]  D. Tseliakhovich, C. Hirata, and A. Slosar, "Non-Gaussianity and large-scale structure in a two-field inflationary model," *Phys.Rev.* **D82** (2010) 043531, `arXiv:1004.3302 [astro-ph.CO]`.

[40]  A. D. Linde and V. F. Mukhanov, "Nongaussian isocurvature perturbations from inflation," *Phys. Rev.* **D56** (1997) 535–539, `arXiv:astro-ph/9610219`.

[41]  D. H. Lyth, C. Ungarelli, and D. Wands, "The primordial density perturbation in the curvaton scenario," *Phys. Rev.* **D67** (2003) 023503, `arXiv:astro-ph/0208055`.

[42]  M. Zaldarriaga, "Non-Gaussianities in models with a varying inflaton decay rate," *Phys. Rev.* **D69** (2004) 043508, `arXiv:astro-ph/0306006`.

[43]  Q.-G. Huang, "Curvaton with Polynomial Potential," *JCAP* **0811** (2008) 005,
      `arXiv:0808.1793 [hep-th]`.

[44]  J. Kumar, L. Leblond, and A. Rajaraman, "Scale Dependent Local
      Non-Gaussianity from Loops," *JCAP* **1004** (2010) 024, `arXiv:0909.2040`
      `[astro-ph.CO]`.

[45]  C. T. Byrnes and K.-Y. Choi, "Review of local non-Gaussianity from
      multi-field inflation," *Adv.Astron.* **2010** (2010) 724525, `arXiv:1002.3110`
      `[astro-ph.CO]`.

[46]  N. Barnaby, "Nongaussianity from Particle Production During Inflation,"
      *Adv.Astron.* **2010** (2010) 156180, `arXiv:1010.5507 [astro-ph.CO]`.

[47]  J.-O. Gong and H. M. Lee, "Large non-Gaussianity in non-minimal inflation,"
      `arXiv:1105.0073 [hep-ph]`.

[48]  V. Demozzi, A. Linde, and V. Mukhanov, "Supercurvaton," *JCAP* **1104**
      (2011) 013, `arXiv:1012.0549 [hep-th]`.

[49]  C. M. Peterson and M. Tegmark, "Non-Gaussianity in Two-Field Inflation,"
      *Phys.Rev.* **D84** (2011) 023520, `arXiv:1011.6675 [astro-ph.CO]`.

[50]  T. Suyama, T. Takahashi, M. Yamaguchi, and S. Yokoyama, "On
      Classification of Models of Large Local-Type Non-Gaussianity," *JCAP* **1012**
      (2010) 030, `arXiv:1009.1979 [astro-ph.CO]`.

[51]  D. Langlois and A. Lepidi, "General treatment of isocurvature perturbations
      and non-Gaussianities," *JCAP* **1101** (2011) 008, `arXiv:1007.5498`
      `[astro-ph.CO]`.

[52]  C. T. Byrnes, K. Enqvist, and T. Takahashi, "Scale-dependence of
      Non-Gaussianity in the Curvaton Model," *JCAP* **1009** (2010) 026,
      `arXiv:1007.5148 [astro-ph.CO]`.

[53]  C. Burgess, M. Cicoli, M. Gomez-Reino, F. Quevedo, G. Tasinato, *et al.*,
      "Non-standard primordial fluctuations and nongaussianity in string inflation,"
      *JHEP* **1008** (2010) 045, `arXiv:1005.4840 [hep-th]`.

[54] A. Mazumdar and J. Rocher, "Particle physics models of inflation and curvaton scenarios," *Phys.Rept.* **497** (2011) 85–215, `arXiv:1001.0993 [hep-ph]`.

[55] X. Chen and Y. Wang, "Quasi-Single Field Inflation and Non-Gaussianities," *JCAP* **1004** (2010) 027, `arXiv:0911.3380 [hep-th]`.

[56] L. Alabidi, K. Malik, C. T. Byrnes, and K.-Y. Choi, "How the curvaton scenario, modulated reheating and an inhomogeneous end of inflation are related," *JCAP* **1011** (2010) 037, `arXiv:1002.1700 [astro-ph.CO]`.

[57] T. Takahashi, M. Yamaguchi, J. Yokoyama, and S. Yokoyama, "Gravitino Dark Matter and Non-Gaussianity," *Phys.Lett.* **B678** (2009) 15–19, `arXiv:0905.0240 [astro-ph.CO]`.

[58] Q.-G. Huang and Y. Wang, "Curvaton Dynamics and the Non-Linearity Parameters in Curvaton Model," *JCAP* **0809** (2008) 025, `arXiv:0808.1168 [hep-th]`.

[59] M. Kawasaki, K. Nakayama, T. Sekiguchi, T. Suyama, and F. Takahashi, "Non-Gaussianity from isocurvature perturbations," *JCAP* **0811** (2008) 019, `arXiv:0808.0009 [astro-ph]`.

[60] A. Chambers and A. Rajantie, "Non-Gaussianity from massless preheating," *JCAP* **0808** (2008) 002, `arXiv:0805.4795 [astro-ph]`.

[61] S. Yokoyama, T. Suyama, and T. Tanaka, "Primordial Non-Gaussianity in Multi-Scalar Slow-Roll Inflation," *JCAP* **0707** (2007) 013, `arXiv:0705.3178 [astro-ph]`.

[62] M. Sasaki, J. Valiviita, and D. Wands, "Non-Gaussianity of the primordial perturbation in the curvaton model," *Phys.Rev.* **D74** (2006) 103003, `arXiv:astro-ph/0607627 [astro-ph]`.

[63] N. Barnaby and J. M. Cline, "Nongaussian and nonscale-invariant perturbations from tachyonic preheating in hybrid inflation," *Phys.Rev.* **D73** (2006) 106012, `arXiv:astro-ph/0601481 [astro-ph]`.

[64] H. R. Cogollo, Y. Rodriguez, and C. A. Valenzuela-Toledo, "On the Issue of the zeta Series Convergence and Loop Corrections in the Generation of

Observable Primordial Non-Gaussianity in Slow-Roll Inflation. Part I: The Bispectrum," *JCAP* **0808** (2008) 029, `arXiv:0806.1546 [astro-ph]`.

[65] D. J. Chung, E. W. Kolb, and A. Riotto, "Superheavy dark matter," *Phys.Rev.* **D59** (1999) 023501, `arXiv:hep-ph/9802238 [hep-ph]`. In *Venice 1999, Neutrino telescopes, vol. 2* 217-237.

[66] V. Kuzmin and I. Tkachev, "Matter creation via vacuum fluctuations in the early universe and observed ultra-high energy cosmic ray events," *Phys. Rev.* **D59** (1999) 123006, `arXiv:hep-ph/9809547`.

[67] D. J. Chung, E. W. Kolb, A. Riotto, and L. Senatore, "Isocurvature constraints on gravitationally produced superheavy dark matter," *Phys.Rev.* **D72** (2005) 023511, `arXiv:astro-ph/0411468 [astro-ph]`.

[68] D. J. Chung, E. W. Kolb, and A. Riotto, "Nonthermal supermassive dark matter," *Phys.Rev.Lett.* **81** (1998) 4048–4051, `arXiv:hep-ph/9805473 [hep-ph]`.

[69] J. R. Ellis, J. L. Lopez, and D. V. Nanopoulos, "Confinement of fractional charges yields integer charged relics in string models," *Phys.Lett.* **B247** (1990) 257.

[70] K. Benakli, J. R. Ellis, and D. V. Nanopoulos, "Natural candidates for superheavy dark matter in string and M theory," *Phys. Rev.* **D59** (1999) 047301, `arXiv:hep-ph/9803333`.

[71] K. Hamaguchi, Y. Nomura, and T. Yanagida, "Longlived superheavy dark matter with discrete gauge symmetries," *Phys.Rev.* **D59** (1999) 063507, `arXiv:hep-ph/9809426 [hep-ph]`.

[72] K. Hamaguchi, K. Izawa, Y. Nomura, and T. Yanagida, "Longlived superheavy particles in dynamical supersymmetry breaking models in supergravity," *Phys.Rev.* **D60** (1999) 125009, `arXiv:hep-ph/9903207 [hep-ph]`.

[73] C. Coriano, A. E. Faraggi, and M. Plumacher, "Stable superstring relics and ultrahigh-energy cosmic rays," *Nucl.Phys.* **B614** (2001) 233–253, `arXiv:hep-ph/0107053 [hep-ph]`.

[74] A. Kusenko and M. E. Shaposhnikov, "Supersymmetric Q balls as dark matter," *Phys.Lett.* **B418** (1998) 46–54, `arXiv:hep-ph/9709492 [hep-ph]`.

[75] T. Han, T. Yanagida, and R.-J. Zhang, "Adjoint messengers and perturbative unification at the string scale," *Phys.Rev.* **D58** (1998) 095011, `arXiv:hep-ph/9804228 [hep-ph]`.

[76] G. Dvali, "Infrared hierarchy, thermal brane inflation and superstrings as superheavy dark matter," *Phys.Lett.* **B459** (1999) 489–496, `arXiv:hep-ph/9905204 [hep-ph]`.

[77] H.-C. Cheng, K. T. Matchev, and M. Schmaltz, "Radiative corrections to Kaluza-Klein masses," *Phys.Rev.* **D66** (2002) 036005, `arXiv:hep-ph/0204342 [hep-ph]`.

[78] G. Shiu and L.-T. Wang, "D matter," *Phys.Rev.* **D69** (2004) 126007, `arXiv:hep-ph/0311228 [hep-ph]`.

[79] V. Berezinsky, M. Kachelriess, and M. Solberg, "Supersymmetric superheavy dark matter," *Phys.Rev.* **D78** (2008) 123535, `arXiv:0810.3012 [hep-ph]`.

[80] T. W. Kephart and Q. Shafi, "Family unification, exotic states and magnetic monopoles," *Phys.Lett.* **B520** (2001) 313–316, `arXiv:hep-ph/0105237 [hep-ph]`.

[81] T. W. Kephart, C.-A. Lee, and Q. Shafi, "Family unification, exotic states and light magnetic monopoles," *JHEP* **0701** (2007) 088, `arXiv:hep-ph/0602055 [hep-ph]`.

[82] T. Bunch and P. Davies, "Quantum Field Theory in de Sitter Space: Renormalization by Point Splitting," *Proc.Roy.Soc.Lond.* **A360** (1978) 117–134.

[83] N. Birrell and P. Davies, "QUANTUM FIELDS IN CURVED SPACE,".

[84] L. Parker, "Quantized fields and particle creation in expanding universes. 1.," *Phys.Rev.* **183** (1969) 1057–1068.

[85] D. J. Chung, A. Notari, and A. Riotto, "Minimal theoretical uncertainties in inflationary predictions," *JCAP* **0310** (2003) 012, `arXiv:hep-ph/0305074 [hep-ph]`.

[86] D. J. Chung, "Classical inflation field induced creation of superheavy dark matter," *Phys.Rev.* **D67** (2003) 083514, `arXiv:hep-ph/9809489 [hep-ph]`.

[87] D. H. Lyth, "The curvature perturbation in a box," *JCAP* **0712** (2007) 016, `arXiv:0707.0361 [astro-ph]`.

[88] R. Tabensky and A. H. Taub *Commun.Math.Phys.* **29** (1973) 61.

[89] D. Wands, K. A. Malik, D. H. Lyth, and A. R. Liddle, "A New approach to the evolution of cosmological perturbations on large scales," *Phys.Rev.* **D62** (2000) 043527, `arXiv:astro-ph/0003278 [astro-ph]`.

[90] C. Hikage, K. Koyama, T. Matsubara, T. Takahashi, and M. Yamaguchi, "Limits on Isocurvature Perturbations from Non-Gaussianity in WMAP Temperature Anisotropy," *Mon.Not.Roy.Astron.Soc.* **398** (2009) 2188–2198, `arXiv:0812.3500 [astro-ph]`.

[91] I. Sollom, A. Challinor, and M. P. Hobson, "Cold Dark Matter Isocurvature Perturbations: Constraints and Model Selection," *Phys.Rev.* **D79** (2009) 123521, `arXiv:0903.5257 [astro-ph.CO]`.

[92] J. Valiviita and T. Giannantonio, "Constraints on primordial isocurvature perturbations and spatial curvature by Bayesian model selection," *Phys.Rev.* **D80** (2009) 123516, `arXiv:0909.5190 [astro-ph.CO]`.

[93] S. Weinberg, "Quantum contributions to cosmological correlations," *Phys. Rev.* **D72** (2005) 043514, `arXiv:hep-th/0506236`.

[94] D. J. H. Chung, H. Yoo, and P. Zhou, "Quadratic Isocurvature Cross-Correlation, Ward Identity, and Dark Matter," *Phys.Rev.* **D87** no. 12, (2013) 123502, `arXiv:1303.6024 [astro-ph.CO]`.

[95] K. Hinterbichler, L. Hui, and J. Khoury, "Conformal Symmetries of Adiabatic Modes in Cosmology," *JCAP* **1208** (2012) 017, `arXiv:1203.6351 [hep-th]`.

[96] **WMAP Collaboration** Collaboration, E. Komatsu *et al.*, "Seven-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation," *Astrophys.J.Suppl.* **192** (2011) 18, `arXiv:1001.4538 [astro-ph.CO]`.

[97] R. Bean, J. Dunkley, and E. Pierpaoli, "Constraining Isocurvature Initial Conditions with WMAP 3- year data," *Phys. Rev.* **D74** (2006) 063503, `arXiv:astro-ph/0606685`.

[98] A. P. Yadav, E. Komatsu, and B. D. Wandelt, "Fast Estimator of Primordial Non-Gaussianity from Temperature and Polarization Anisotropies in the Cosmic Microwave Background," *Astrophys.J.* **664** (2007) 680–686, `arXiv:astro-ph/0701921 [astro-ph]`. PhD Thesis (Advisor: Alfonso Arag on-Salamanca).

[99] L. Verde and S. Matarrese, "Detectability of the effect of Inflationary non-Gaussianity on halo bias," *Astrophys.J.* **706** (2009) L91–L95, `arXiv:0909.3224 [astro-ph.CO]`.

[100] D. J. H. Chung, P. Crotty, E. W. Kolb, and A. Riotto, "On the gravitational production of superheavy dark matter," *Phys. Rev.* **D64** (2001) 043503, `arXiv:hep-ph/0104100`.

[101] A. R. Liddle and S. M. Leach, "How long before the end of inflation were observable perturbations produced?," *Phys. Rev.* **D68** (2003) 103503, `arXiv:astro-ph/0305263`.

[102] B. J. Carr, J. Gilbert, and J. E. Lidsey, "Black hole relics and inflation: Limits on blue perturbation spectra," *Phys.Rev.* **D50** (1994) 4853–4867, `arXiv:astro-ph/9405027 [astro-ph]`.

[103] M. Berkooz, D. J. Chung, and T. Volansky, "Constraining modular inflation in the MSSM from giant Q-ball formation," *Phys.Rev.* **D73** (2006) 063526, `arXiv:hep-ph/0507218 [hep-ph]`.

[104] R. Easther, R. Flauger, and J. B. Gilmore, "Delayed Reheating and the Breakdown of Coherent Oscillations," *JCAP* **1104** (2011) 027, `arXiv:1003.3011 [astro-ph.CO]`.

[105] B. S. Acharya, P. Kumar, K. Bobkov, G. Kane, J. Shao, and S. Watson, "Non-thermal Dark Matter and the Moduli Problem in String Frameworks," *JHEP* **0806** (2008) 064, arXiv:0804.0863 [hep-ph].

[106] B. S. Acharya, G. Kane, S. Watson, and P. Kumar, "A Non-thermal WIMP Miracle," *Phys.Rev.* **D80** (2009) 083529, arXiv:0908.2430 [astro-ph.CO].

[107] **Planck Collaboration** Collaboration, P. Ade *et al.*, "Planck 2013 results. I. Overview of products and scientific results," arXiv:1303.5062 [astro-ph.CO].

[108] **Planck Collaboration** Collaboration, P. Ade *et al.*, "Planck 2013 results. XVI. Cosmological parameters," arXiv:1303.5076 [astro-ph.CO].

[109] **Planck Collaboration** Collaboration, P. Ade *et al.*, "Planck 2013 results. XXII. Constraints on inflation," arXiv:1303.5082 [astro-ph.CO].

[110] **Planck Collaboration** Collaboration, P. Ade *et al.*, "Planck 2013 results. XXIII. Isotropy and Statistics of the CMB," arXiv:1303.5083 [astro-ph.CO].

[111] **Planck Collaboration** Collaboration, P. Ade *et al.*, "Planck 2013 Results. XXIV. Constraints on primordial non-Gaussianity," arXiv:1303.5084 [astro-ph.CO].

[112] G. Hinshaw, D. Larson, E. Komatsu, D. Spergel, C. Bennett, *et al.*, "Nine-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Parameter Results," arXiv:1212.5226 [astro-ph.CO].

[113] **QUaD collaboration** Collaboration, . M. Brown *et al.*, "Improved measurements of the temperature and polarization of the CMB from QUaD," *Astrophys.J.* **705** (2009) 978–999, arXiv:0906.1003 [astro-ph.CO].

[114] C. Reichardt, P. Ade, J. Bock, J. R. Bond, J. Brevik, *et al.*, "High resolution CMB power spectrum from the complete ACBAR data set," *Astrophys.J.* **694** (2009) 1200–1219, arXiv:0801.1491 [astro-ph].

[115] **ACT Collaboration** Collaboration, J. Fowler *et al.*, "The Atacama Cosmology Telescope: A Measurement of the 600 ell 8000 Cosmic Microwave

Background Power Spectrum at 148 GHz," *Astrophys.J.* **722** (2010) 1148–1161, `arXiv:1001.2934 [astro-ph.CO]`.

[116] M. Lueker, C. Reichardt, K. Schaffer, O. Zahn, P. Ade, *et al.*, "Measurements of Secondary Cosmic Microwave Background Anisotropies with the South Pole Telescope," *Astrophys.J.* **719** (2010) 1045–1066, `arXiv:0912.4317 [astro-ph.CO]`.

[117] W. J. Percival, S. Cole, D. J. Eisenstein, R. C. Nichol, J. A. Peacock, *et al.*, "Measuring the Baryon Acoustic Oscillation scale using the SDSS and 2dFGRS," *Mon.Not.Roy.Astron.Soc.* **381** (2007) 1053–1066, `arXiv:0705.3323 [astro-ph]`.

[118] **SDSS Collaboration** Collaboration, D. J. Eisenstein *et al.*, "Detection of the baryon acoustic peak in the large-scale correlation function of SDSS luminous red galaxies," *Astrophys.J.* **633** (2005) 560–574, `arXiv:astro-ph/0501171 [astro-ph]`.

[119] K. Sato, "First Order Phase Transition of a Vacuum and Expansion of the Universe," *Mon.Not.Roy.Astron.Soc.* **195** (1981) 467–479.

[120] V. F. Mukhanov and G. Chibisov, "Quantum Fluctuation and Nonsingular Universe. (In Russian)," *JETP Lett.* **33** (1981) 532–535.

[121] S. Hawking and I. Moss, "FLUCTUATIONS IN THE INFLATIONARY UNIVERSE," *Nucl.Phys.* **B224** (1983) 180.

[122] A. H. Guth and S. Pi, "Fluctuations in the New Inflationary Universe," *Phys.Rev.Lett.* **49** (1982) 1110–1113.

[123] A. A. Starobinsky, "Dynamics of Phase Transition in the New Inflationary Universe Scenario and Generation of Perturbations," *Phys.Lett.* **B117** (1982) 175–178.

[124] J. M. Bardeen, P. J. Steinhardt, and M. S. Turner, "Spontaneous Creation of Almost Scale - Free Density Perturbations in an Inflationary Universe," *Phys.Rev.* **D28** (1983) 679.

[125] J. Silk and M. S. Turner, "Double Inflation," *Phys.Rev.* **D35** (1987) 419.

[126] D. Polarski and A. A. Starobinsky, "Isocurvature perturbations in multiple inflationary models," *Phys.Rev.* **D50** (1994) 6123–6129, `arXiv:astro-ph/9404061 [astro-ph]`.

[127] D. Langlois, "Correlated adiabatic and isocurvature perturbations from double inflation," *Phys.Rev.* **D59** (1999) 123512, `arXiv:astro-ph/9906080 [astro-ph]`.

[128] M. Yamaguchi, "Density fluctuations and primordial black holes formation in natural double inflation in supergravity," *Phys.Rev.* **D64** (2001) 063503, `arXiv:hep-ph/0105001 [hep-ph]`.

[129] D. H. Lyth and D. Wands, "Generating the curvature perturbation without an inflaton," *Phys.Lett.* **B524** (2002) 5–14, `arXiv:hep-ph/0110002 [hep-ph]`.

[130] K. Enqvist and M. S. Sloth, "Adiabatic CMB perturbations in pre - big bang string cosmology," *Nucl.Phys.* **B626** (2002) 395–409, `arXiv:hep-ph/0109214 [hep-ph]`.

[131] T. Moroi and T. Takahashi, "Effects of cosmological moduli fields on cosmic microwave background," *Phys.Lett.* **B522** (2001) 215–221, `arXiv:hep-ph/0110096 [hep-ph]`.

[132] D. Langlois and T. Takahashi, "Density Perturbations from Modulated Decay of the Curvaton," `arXiv:1301.3319 [astro-ph.CO]`.

[133] K. Enqvist, D. G. Figueroa, and R. N. Lerner, "Curvaton Decay by Resonant Production of the Standard Model Higgs," *JCAP* **1301** (2013) 040, `arXiv:1211.5028 [astro-ph.CO]`.

[134] K. Harigaya, M. Ibe, M. Kawasaki, and T. T. Yanagida, "Non-Gaussianity from Attractor Curvaton," `arXiv:1211.3535 [hep-ph]`.

[135] S. Enomoto, K. Kohri, and T. Matsuda, "Non-Gaussianity in the unified curvaton mechanism : The generalized curvaton mechanism that comprehends modulation at the transition," `arXiv:1210.7118 [hep-ph]`.

[136] K. Enqvist, R. N. Lerner, O. Taanila, and A. Tranberg, "Spectator field dynamics in de Sitter and curvaton initial conditions," *JCAP* **1210** (2012) 052, `arXiv:1205.5446 [astro-ph.CO]`.

[137] K. Dimopoulos, K. Kohri, D. H. Lyth, and T. Matsuda, "The inflating curvaton," *JCAP* **1203** (2012) 022, `arXiv:1110.2951 [astro-ph.CO]`.

[138] C. Lin and Y. Wang, "Quadra-Spectrum and Quint-Spectrum from Inflation and Curvaton Models," *JCAP* **1007** (2010) 011, `arXiv:1004.0461 [astro-ph.CO]`.

[139] H. Assadullahi, J. Valiviita, and D. Wands, "Primordial non-Gaussianity from two curvaton decays," *Phys.Rev.* **D76** (2007) 103003, `arXiv:0708.0223 [hep-ph]`.

[140] T. Moroi and T. Takahashi, "Cosmic density perturbations from late decaying scalar condensations," *Phys.Rev.* **D66** (2002) 063501, `arXiv:hep-ph/0206026 [hep-ph]`.

[141] N. Bartolo and A. R. Liddle, "The Simplest curvaton model," *Phys.Rev.* **D65** (2002) 121301, `arXiv:astro-ph/0203076 [astro-ph]`.

[142] D. Seckel and M. S. Turner, "Isothermal Density Perturbations in an Axion Dominated Inflationary Universe," *Phys.Rev.* **D32** (1985) 3178.

[143] J. Preskill, M. B. Wise, and F. Wilczek, "Cosmology of the Invisible Axion," *Phys.Lett.* **B120** (1983) 127–132.

[144] L. Abbott and P. Sikivie, "A Cosmological Bound on the Invisible Axion," *Phys.Lett.* **B120** (1983) 133–136.

[145] M. Dine and W. Fischler, "The Not So Harmless Axion," *Phys.Lett.* **B120** (1983) 137–141.

[146] P. J. Steinhardt and M. S. Turner, "Saving the Invisible Axion," *Phys.Lett.* **B129** (1983) 51.

[147] M. S. Turner, "Cosmic and Local Mass Density of Invisible Axions," *Phys.Rev.* **D33** (1986) 889–896.

[148] E. W. Kolb and M. S. Turner, "The Early universe," *Front.Phys.* **69** (1990) 1–547.

[149] P. Fox, A. Pierce, and S. D. Thomas, "Probing a QCD string axion with precision cosmological measurements," arXiv:hep-th/0409059 [hep-th].

[150] M. Beltran, J. Garcia-Bellido, and J. Lesgourgues, "Isocurvature bounds on axions revisited," *Phys.Rev.* **D75** (2007) 103507, arXiv:hep-ph/0606107 [hep-ph].

[151] M. P. Hertzberg, M. Tegmark, and F. Wilczek, "Axion Cosmology and the Energy Scale of Inflation," *Phys.Rev.* **D78** (2008) 083507, arXiv:0807.1726 [astro-ph].

[152] N. Bartolo, S. Matarrese, and A. Riotto, "Nongaussianity from inflation," *Phys.Rev.* **D65** (2002) 103505, arXiv:hep-ph/0112261 [hep-ph].

[153] D. Langlois, F. Vernizzi, and D. Wands, "Non-linear isocurvature perturbations and non-Gaussianities," *JCAP* **0812** (2008) 004, arXiv:0809.4646 [astro-ph].

[154] L. Kofman, G. R. Blumenthal, H. Hodges, and J. R. Primack, "GENERATION OF NONFLAT AND NONGAUSSIAN PERTURBATIONS FROM INFLATION," *ASP Conf.Ser.* **15** (1991) 339–351.

[155] B. Geyer, D. Robaschik, and J. Eilers, "Target mass corrections for virtual Compton scattering at twist-2 and generalized, non-forward Wandzura-Wilczek and Callan-Gross relations," *Nucl.Phys.* **B704** (2005) 279–331, arXiv:hep-ph/0407300 [hep-ph].

[156] F. Ferrer, S. Rasanen, and J. Valiviita, "Correlated isocurvature perturbations from mixed inflaton-curvaton decay," *JCAP* **0410** (2004) 010, arXiv:astro-ph/0407300 [astro-ph].

[157] L. Boubekeur and D. Lyth, "Detecting a small perturbation through its non-Gaussianity," *Phys.Rev.* **D73** (2006) 021301, arXiv:astro-ph/0504046 [astro-ph].

[158] J. Barbon and C. Hoyos-Badajoz, "Dynamical Higgs potentials with a landscape," *Phys.Rev.* **D73** (2006) 126002, arXiv:hep-th/0602285 [hep-th].

[159] D. H. Lyth, "Non-gaussianity and cosmic uncertainty in curvaton-type models," *JCAP* **0606** (2006) 015, `arXiv:astro-ph/0602285` `[astro-ph]`.

[160] K. Koyama, S. Mizuno, F. Vernizzi, and D. Wands, "Non-Gaussianities from ekpyrotic collapse with multiple fields," *JCAP* **0711** (2007) 024, `arXiv:0708.4321` `[hep-th]`.

[161] Z. Lalak, D. Langlois, S. Pokorski, and K. Turzynski, "Curvature and isocurvature perturbations in two-field inflation," *JCAP* **0707** (2007) 014, `arXiv:0704.0212` `[hep-th]`.

[162] M.-x. Huang, G. Shiu, and B. Underwood, "Multifield DBI Inflation and Non-Gaussianities," *Phys.Rev.* **D77** (2008) 023511, `arXiv:0709.3299` `[hep-th]`.

[163] J.-L. Lehners, "Ekpyrotic and Cyclic Cosmology," *Phys.Rept.* **465** (2008) 223–263, `arXiv:0806.1245` `[astro-ph]`.

[164] M. Beltrán, "Isocurvature, non-Gaussianity, and the curvaton model," *Physical Review D* **78** no. 2, (2008) 023530.

[165] M. Kawasaki, K. Nakayama, and F. Takahashi, "Non-Gaussianity from Baryon Asymmetry," *JCAP* **0901** (2009) 002, `arXiv:0809.2242` `[hep-ph]`.

[166] D. Langlois and L. Sorbo, "Primordial perturbations and non-Gaussianities from modulated trapping," *JCAP* **0908** (2009) 014, `arXiv:0906.1813` `[astro-ph.CO]`.

[167] X. Chen, "Primordial Non-Gaussianities from Inflation Models," *Adv.Astron.* **2010** (2010) 638979, `arXiv:1002.1416` `[astro-ph.CO]`.

[168] D. Langlois and T. Takahashi, "Primordial Trispectrum from Isocurvature Fluctuations," *JCAP* **1102** (2011) 020, `arXiv:1012.4885` `[astro-ph.CO]`.

[169] D. Mulryne, S. Orani, and A. Rajantie, "Non-Gaussianity from the hybrid potential," *Phys.Rev.* **D84** (2011) 123527, `arXiv:1107.4739` `[hep-th]`.

[170] A. De Simone, H. Perrier, and A. Riotto, "Non-Gaussianities from the Standard Model Higgs," *JCAP* **1301** (2013) 037, `arXiv:1210.6618` `[hep-ph]`.

[171] K. Enqvist and S. Rusak, "Modulated preheating and isocurvature perturbations," *JCAP* **1303** (2013) 017, `arXiv:1210.2192 [astro-ph.CO]`.

[172] M. Kawasaki, T. Kobayashi, and F. Takahashi, "Non-Gaussianity from Curvatons Revisited," *Phys.Rev.* **D84** (2011) 123506, `arXiv:1107.6011 [astro-ph.CO]`.

[173] M. Kawasaki and K. Nakayama, "Axions : Theory and Cosmological Role," `arXiv:1301.1123 [hep-ph]`.

[174] S. Nurmi, C. T. Byrnes, and G. Tasinato, "A non-Gaussian landscape," `arXiv:1301.3128 [astro-ph.CO]`.

[175] M. Kawasaki, F. Takahashi, and T. Yanagida, "Gravitino overproduction in inflaton decay," *Phys.Lett.* **B638** (2006) 8–12, `arXiv:hep-ph/0603265 [hep-ph]`.

[176] M. Kawasaki, F. Takahashi, and T. Yanagida, "The Gravitino-overproduction problem in inflationary universe," *Phys.Rev.* **D74** (2006) 043519, `arXiv:hep-ph/0605297 [hep-ph]`.

[177] J. Pradler and F. D. Steffen, "Constraints on the Reheating Temperature in Gravitino Dark Matter Scenarios," *Phys.Lett.* **B648** (2007) 224–235, `arXiv:hep-ph/0612291 [hep-ph]`.

[178] J. Pradler and F. D. Steffen, "Thermal gravitino production and collider tests of leptogenesis," *Phys.Rev.* **D75** (2007) 023509, `arXiv:hep-ph/0608344 [hep-ph]`.

[179] M. Endo, F. Takahashi, and T. Yanagida, "Anomaly-induced inflaton decay and gravitino-overproduction problem," *Phys.Lett.* **B658** (2008) 236–240, `arXiv:hep-ph/0701042 [hep-ph]`.

[180] M. Endo, F. Takahashi, and T. Yanagida, "Inflaton Decay in Supergravity," *Phys.Rev.* **D76** (2007) 083509, `arXiv:0706.0986 [hep-ph]`.

[181] T. Takahashi, M. Yamaguchi, and S. Yokoyama, "Primordial Non-Gaussianity in Models with Dark Matter Isocurvature Fluctuations," *Phys.Rev.* **D80** (2009) 063524, `arXiv:0907.3052 [astro-ph.CO]`.

[182] A. A. Starobinsky, "Multicomponent de Sitter (Inflationary) Stages and the Generation of Perturbations," *JETP Lett.* **42** (1985) 152–155.

[183] 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** (1996) 71–78, `arXiv:astro-ph/9507001 [astro-ph]`.

[184] M. Sasaki and T. Tanaka, "Superhorizon scale dynamics of multiscalar inflation," *Prog.Theor.Phys.* **99** (1998) 763–782, `arXiv:gr-qc/9801017 [gr-qc]`.

[185] D. J. Chung, L. L. Everett, H. Yoo, and P. Zhou, "Gravitational Fermion Production in Inflationary Cosmology," *Phys.Lett.* **B712** (2012) 147–154, `arXiv:1109.2524 [astro-ph.CO]`.

[186] V. A. Kuzmin and I. I. Tkachev, "Ultrahigh-energy cosmic rays and inflation relics," *Phys.Rept.* **320** (1999) 199–221, `arXiv:hep-ph/9903542 [hep-ph]`.

[187] B. S. DeWitt, "Quantum Field Theory in Curved Space-Time," *Phys.Rept.* **19** (1975) 295–357.

[188] P. Peebles, *The Large-scale Structure of the Universe*. Princeton series in physics. PRINCETON University Press, 1980.

[189] G. Efstathiou and J. R. Bond, "Isocurvature cold dark matter fluctuations," *Monthly Notices of the Royal Astronomical Society* **218** (Jan., 1986) 103–121.

[190] D. Green, M. Lewandowski, L. Senatore, E. Silverstein, and M. Zaldarriaga, "Anomalous Dimensions and Non-Gaussianity," `arXiv:1301.2630 [hep-th]`.

[191] N. Bartolo, E. Komatsu, S. Matarrese, and A. Riotto, "Non-Gaussianity from inflation: Theory and observations," *Phys.Rept.* **402** (2004) 103–266, `arXiv:astro-ph/0406398 [astro-ph]`.

[192] M. Kawasaki, K. Nakayama, T. Sekiguchi, T. Suyama, and F. Takahashi, "A General Analysis of Non-Gaussianity from Isocurvature Perturbations," *JCAP* **0901** (2009) 042, `arXiv:0810.0208 [astro-ph]`.

[193] J. Zinn-Justin, *Quantum Field Theory and Critical Phenomena*. International Series of Monographs on Physics. Clarendon Press, 2002.

[194] S. Weinberg, "Ultraviolet Divergences in Cosmological Correlations," *Phys.Rev.* **D83** (2011) 063508, `arXiv:1011.1630 [hep-th]`.

[195] S. Weinberg, "Quantum contributions to cosmological correlations," *Physical Review D* **72** no. 4, (2005) 043514.

[196] E. A. Calzetta and B.-L. Hu, *Nonequilibrium quantum field theory.* Cambridge Univ Pr, 2008.

[197] P. Sikivie, "Axion Cosmology," *Lect.Notes Phys.* **741** (2008) 19–50, `arXiv:astro-ph/0610440 [astro-ph]`.

[198] E. Calzetta and B. Hu, *Nonequilibrium Quantum Field Theory.* Cambridge Monographs on Mathematical Physics. Cambridge University Press, 2008.

[199] R. Bean, J. Dunkley, and E. Pierpaoli, "Constraining isocurvature initial conditions with WMAP 3-year data," *Physical Review D* **74** no. 6, (2006) 063503.

[200] E. Komatsu, J. Dunkley, M. R. Nolta, C. L. Bennett, B. Gold, G. Hinshaw, N. Jarosik, D. Larson, M. Limon, L. Page, D. N. Spergel, M. Halpern, R. S. Hill, A. Kogut, S. S. Meyer, G. S. Tucker, J. L. Weiland, E. Wollack, and E. L. Wright, "Five-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation," *arXiv* **astro-ph** (Mar., 2008) .

[201] E. Komatsu, K. M. Smith, J. Dunkley, C. L. Bennett, B. Gold, G. Hinshaw, N. Jarosik, D. Larson, M. R. Nolta, L. Page, D. N. Spergel, M. Halpern, R. S. Hill, A. Kogut, M. Limon, S. S. Meyer, N. Odegard, G. S. Tucker, J. L. Weiland, E. Wollack, and E. L. Wright, "SEVEN-YEAR WILKINSON MICROWAVE ANISOTROPY PROBE (WMAP1) OBSERVATIONS: COSMOLOGICAL INTERPRETATION," *arXiv* **astro-ph.CO** (Jan., 2010) .

[202] D. Larson, J. Dunkley, G. Hinshaw, E. Komatsu, M. Nolta, C. Bennett, B. Gold, M. Halpern, R. Hill, and N. Jarosik, "Seven-year Wilkinson Microwave Anisotropy Probe (WMAP) observations: power spectra and WMAP-derived parameters," *The Astrophysical Journal Supplement Series* **192** (2011) 16.

[203] E. Komatsu and D. N. Spergel, "Acoustic signatures in the primary microwave background bispectrum," *Phys. Rev.* **D63** (2001) 063002, `arXiv:astro-ph/0005036`.

[204] C. Hikage, M. Kawasaki, T. Sekiguchi, and T. Takahashi, "CMB constraint on non-Gaussianity in isocurvature perturbations," `arXiv:1211.1095 [astro-ph.CO]`.

[205] C. Hikage, M. Kawasaki, T. Sekiguchi, and T. Takahashi, "Extended analysis of CMB constraints on non-Gaussianity in isocurvature perturbations," *JCAP* **1303** (2013) 020, `arXiv:1212.6001 [astro-ph.CO]`.

[206] R. Arnowitt, S. Deser, and C. Misner, "The dynamics of general relativity," in *Gravitation: An Introduction to Current Research*, L. Witten, ed., ch. 7, pp. 227–265. Wiley, 1962. `arXiv:0405109 [gr-qc]`.

[207] P. Crotty, J. Garcia-Bellido, J. Lesgourgues, and A. Riazuelo, "Bounds on isocurvature perturbations from cosmic microwave background and large scale structure data," *Physical Review Letters* **91** no. 17, (2003) 171301.

[208] J. Väliviita and T. Giannantonio, "Constraints on primordial isocurvature perturbations and spatial curvature by Bayesian model selection," *Physical Review D* **80** no. 12, (Dec., 2009) .

[209] P. Collaboration, "XXII. Constraints on inflation,".

[210] H. Assadullahi, H. Firouzjahi, M. H. Namjoo, and D. Wands, "Modulated curvaton decay," `arXiv:1301.3439 [hep-th]`.

[211] S. Weinberg, "Adiabatic modes in cosmology," *Phys.Rev.* **D67** (2003) 123504, `arXiv:astro-ph/0302326 [astro-ph]`.

[212] C. Cheung, P. Creminelli, A. L. Fitzpatrick, J. Kaplan, and L. Senatore, "The Effective Field Theory of Inflation," *JHEP* **0803** (2008) 014, `arXiv:0709.0293 [hep-th]`.

[213] J. M. Maldacena and G. L. Pimentel, "On graviton non-Gaussianities during inflation," *JHEP* **1109** (2011) 045, `arXiv:1104.2846 [hep-th]`.

[214] V. Assassi, D. Baumann, and D. Green, "On Soft Limits of Inflationary Correlation Functions," *JCAP* **1211** (2012) 047, `arXiv:1204.4207 [hep-th]`.

[215] M. Li and Y. Wang, "Consistency Relations for Non-Gaussianity," *JCAP* **0809** (2008) 018, `arXiv:0807.3058 [hep-th]`.

[216] I. Antoniadis, P. O. Mazur, and E. Mottola, "Conformal Invariance, Dark Energy, and CMB Non-Gaussianity," *JCAP* **1209** (2012) 024, `arXiv:1103.4164 [gr-qc]`.

[217] A. Kehagias and A. Riotto, "Operator Product Expansion of Inflationary Correlators and Conformal Symmetry of de Sitter," *Nucl.Phys.* **B864** (2012) 492–529, `arXiv:1205.1523 [hep-th]`.

[218] P. Creminelli, A. Joyce, J. Khoury, and M. Simonovic, "Consistency Relations for the Conformal Mechanism," `arXiv:1212.3329 [hep-th]`.

[219] P. Creminelli, J. Norena, and M. Simonovic, "Conformal consistency relations for single-field inflation," *JCAP* **1207** (2012) 052, `arXiv:1203.4595 [hep-th]`.

[220] L. Senatore and M. Zaldarriaga, "A Note on the Consistency Condition of Primordial Fluctuations," *JCAP* **1208** (2012) 001, `arXiv:1203.6884 [astro-ph.CO]`.

[221] N. S. Sugiyama, "Consistency Relation for multifield inflation scenario with all loop contributions," *JCAP* **1205** (2012) 032, `arXiv:1201.4048 [gr-qc]`.

[222] V. Assassi, D. Baumann, and D. Green, "Symmetries and Loops in Inflation," `arXiv:1210.7792 [hep-th]`.

[223] P. Creminelli, "Conformal invariance of scalar perturbations in inflation," *Phys.Rev.* **D85** (2012) 041302, `arXiv:1108.0874 [hep-th]`.

[224] P. Creminelli, G. D'Amico, M. Musso, and J. Norena, "The (not so) squeezed limit of the primordial 3-point function," *JCAP* **1111** (2011) 038, `arXiv:1106.1462 [astro-ph.CO]`.

[225] A. Kehagias and A. Riotto, "Symmetries and Consistency Relations in the Large Scale Structure of the Universe," `arXiv:1302.0130 [astro-ph.CO]`.

[226] H. Li, J. Liu, J.-Q. Xia, and Y.-F. Cai, "Cold Dark Matter Isocurvature Perturbations: Cosmological Constraints and Applications," *Phys.Rev.* **D83** (2011) 123517, `arXiv:1012.2511 [astro-ph.CO]`.

[227] A. H. Guth and S. Pi, "Fluctuations in the New Inflationary Universe," *Phys.Rev.Lett.* **49** (1982) 1110–1113.

[228] A. D. Linde, "Generation of Isothermal Density Perturbations in the Inflationary Universe," *Phys.Lett.* **B158** (1985) 375–380.

[229] C. Kiefer, J. Lesgourgues, D. Polarski, and A. A. Starobinsky, "The Coherence of primordial fluctuations produced during inflation," *Class.Quant.Grav.* **15** (1998) L67–L72, `arXiv:gr-qc/9806066 [gr-qc]`.

[230] A. D. Linde, "GENERATION OF ISOTHERMAL DENSITY PERTURBATIONS IN THE INFLATIONARY UNIVERSE," *JETP Lett.* **40** (1984) 1333–1336.

[231] M. Bucher, J. Dunkley, P. Ferreira, K. Moodley, and C. Skordis, "The Initial conditions of the universe: How much isocurvature is allowed?," *Phys.Rev.Lett.* **93** (2004) 081301, `arXiv:astro-ph/0401417 [astro-ph]`.

[232] B. A. Bassett, S. Tsujikawa, and D. Wands, "Inflation dynamics and reheating," *Rev.Mod.Phys.* **78** (2006) 537–589, `arXiv:astro-ph/0507632 [astro-ph]`.

[233] J. M. Bardeen, "Gauge Invariant Cosmological Perturbations," *Phys.Rev.* **D22** (1980) 1882–1905.

[234] V. F. Mukhanov, H. Feldman, and R. H. Brandenberger, "Theory of cosmological perturbations. Part 1. Classical perturbations. Part 2. Quantum theory of perturbations. Part 3. Extensions," *Phys.Rept.* **215** (1992) 203–333.

[235] S. Weinberg, *Cosmology.* 2008.

[236] H. Kodama and M. Sasaki, "Cosmological Perturbation Theory," *Prog.Theor.Phys.Suppl.* **78** (1984) 1–166.

[237] M. Bruni, S. Matarrese, S. Mollerach, and S. Sonego, "Perturbations of space-time: Gauge transformations and gauge invariance at second order and

beyond," *Class.Quant.Grav.* **14** (1997) 2585–2606, `arXiv:gr-qc/9609040 [gr-qc]`.

[238] T. Prokopec and J. Weenink, "Uniqueness of the gauge invariant action for cosmological perturbations," *JCAP* **1212** (2012) 031, `arXiv:1209.1701 [gr-qc]`.

[239] F. Arroja and T. Tanaka, "A note on the role of the boundary terms for the non-Gaussianity in general k-inflation," *JCAP* **1105** (2011) 005, `arXiv:1103.1102 [astro-ph.CO]`.

[240] D. Langlois and F. Vernizzi, "A geometrical approach to nonlinear perturbations in relativistic cosmology," *Class.Quant.Grav.* **27** (2010) 124007, `arXiv:1003.3270 [astro-ph.CO]`.

[241] G. Rigopoulos, "Gauge invariance and non-Gaussianity in Inflation," *Phys.Rev.* **D84** (2011) 021301, `arXiv:1104.0292 [astro-ph.CO]`.

[242] K. A. Malik and D. Wands, "Cosmological perturbations," *Phys.Rept.* **475** (2009) 1–51, `arXiv:0809.4944 [astro-ph]`.

[243] G. Ellis and M. Bruni, "COVARIANT AND GAUGE INVARIANT APPROACH TO COSMOLOGICAL DENSITY FLUCTUATIONS," *Phys.Rev.* **D40** (1989) 1804–1818.

[244] J.-c. Hwang, "Perturbations of the Robertson-Walker space - Multicomponent sources and generalized gravity," *Astrophys.J.* **375** (1991) 443–462.

[245] S. Fulling, "ASPECTS OF QUANTUM FIELD THEORY IN CURVED SPACE-TIME," *London Math.Soc.Student Texts* **17** (1989) 1–315.

[246] G. Horowitz and R. M. Wald, "QUANTUM STRESS ENERGY IN NEARLY CONFORMALLY FLAT SPACE-TIMES," *Phys.Rev.* **D21** (1980) 1462–1465.

[247] B. Hu and E. Verdaguer, "Stochastic gravity: Theory and applications," *Living Rev.Rel.* **7** (2004) 3, `arXiv:gr-qc/0307032 [gr-qc]`.

[248] F. Finelli, G. Marozzi, G. Vacca, and G. Venturi, "The Impact of ultraviolet regularization on the spectrum of curvature perturbations during inflation," *Phys.Rev.* **D76** (2007) 103528, `arXiv:0707.1416 [hep-th]`.

[249] T. Prokopec, N. C. Tsamis, and R. P. Woodard, "Two loop stress-energy tensor for inflationary scalar electrodynamics," *Phys.Rev.* **D78** (2008) 043523, `arXiv:0802.3673 [gr-qc]`.

[250] G. Perez-Nadal, A. Roura, and E. Verdaguer, "Stress tensor fluctuations in de Sitter spacetime," *JCAP* **1005** (2010) 036, `arXiv:0911.4870 [gr-qc]`.

[251] J. Baacke, L. Covi, and N. Kevlishvili, "Coupled scalar fields in a flat FRW universe: Renormalisation," *JCAP* **1008** (2010) 026, `arXiv:1006.2376 [hep-ph]`.

[252] L. Ford, S. Miao, K.-W. Ng, R. Woodard, and C.-H. Wu, "Quantum Stress Tensor Fluctuations of a Conformal Field and Inflationary Cosmology," *Phys.Rev.* **D82** (2010) 043501, `arXiv:1005.4530 [gr-qc]`.

[253] I. Agullo, J. Navarro-Salas, G. J. Olmo, and L. Parker, "Remarks on the renormalization of primordial cosmological perturbations," *Phys.Rev.* **D84** (2011) 107304, `arXiv:1108.0949 [gr-qc]`.

[254] C.-H. Wu, J.-T. Hsiang, L. Ford, and K.-W. Ng, "Gravity Waves from Quantum Stress Tensor Fluctuations in Inflation," *Phys.Rev.* **D84** (2011) 103515, `arXiv:1105.1155 [gr-qc]`.

[255] D. Lyth, "Axions and inflation: Sitting in the vacuum," *Phys.Rev.* **D45** (1992) 3394–3404.

[256] D. J. Chung, "Superheavy dark matter," `arXiv:hep-ph/9808323 [hep-ph]`.

[257] R. Peccei and H. R. Quinn, "CP Conservation in the Presence of Instantons," *Phys.Rev.Lett.* **38** (1977) 1440–1443.

[258] S. Weinberg, "A New Light Boson?," *Phys.Rev.Lett.* **40** (1978) 223–226.

[259] F. Wilczek, "Problem of Strong p and t Invariance in the Presence of Instantons," *Phys.Rev.Lett.* **40** (1978) 279–282.

[260] K. Griest and M. Kamionkowski, "Unitarity Limits on the Mass and Radius of Dark Matter Particles," *Phys.Rev.Lett.* **64** (1990) 615.

[261] e. Kuster, Markus, e. Raffelt, Georg, and e. Beltran, Berta, "Axions: Theory, cosmology, and experimental searches. Proceedings, 1st Joint ILIAS-CERN-CAST axion training, Geneva, Switzerland, November 30-December 2, 2005," *Lect.Notes Phys.* **741** (2008) pp.1–258.

[262] S. Dodelson, "Modern cosmology,".

[263] V. F. Mukhanov, "CMB-slow, or how to estimate cosmological parameters by hand," *Int.J.Theor.Phys.* **43** (2004) 623–668, `arXiv:astro-ph/0303072 [astro-ph]`.

[264] J. S. Schwinger, "On gauge invariance and vacuum polarization," *Phys.Rev.* **82** (1951) 664–679.

[265] V. Baier, V. Katkov, and V. Strakhovenko, "An Operator Approach to Quantum Electrodynamics in External Field. 2. Electron Loops," *Zh.Eksp.Teor.Fiz.* **68** (1975) 405–420.

[266] R. Ruffini, G. Vereshchagin, and S.-S. Xue, "Electron-positron pairs in physics and astrophysics: from heavy nuclei to black holes," *Phys.Rept.* **487** (2010) 1–140, `arXiv:0910.0974 [astro-ph.HE]`.

[267] Y. Kluger, J. Eisenberg, B. Svetitsky, F. Cooper, and E. Mottola, "Fermion pair production in a strong electric field," *Phys.Rev.* **D45** (1992) 4659–4671.

[268] A. Nikishov, "Pair production by a constant external field," *Zh.Eksp.Teor.Fiz.* **57** (1969) 1210–1216.

[269] M. Soffel, B. Muller, and W. Greiner, "STABILITY AND DECAY OF THE DIRAC VACUUM IN EXTERNAL GAUGE FIELDS," *Phys.Rept.* **85** (1982) 51–122.

[270] Y. Kluger, E. Mottola, and J. M. Eisenberg, "The Quantum Vlasov equation and its Markov limit," *Phys.Rev.* **D58** (1998) 125015, `arXiv:hep-ph/9803372 [hep-ph]`.

[271] S. P. Kim and D. N. Page, "Schwinger pair production in electric and magnetic fields," *Phys.Rev.* **D73** (2006) 065020, `arXiv:hep-th/0301132 [hep-th]`.

[272] S. P. Kim and D. N. Page, "Improved Approximations for Fermion Pair Production in Inhomogeneous Electric Fields," *Phys.Rev.* **D75** (2007) 045013, `arXiv:hep-th/0701047` `[hep-th]`.

[273] R. Ruffini, G. Vereshchagin, and S.-S. Xue, "Vacuum polarization and plasma oscillations," *Phys.Lett.* **A371** (2007) 399–405, `arXiv:0706.4363` `[astro-ph]`.

[274] F. Hebenstreit, R. Alkofer, and H. Gies, "Pair Production Beyond the Schwinger Formula in Time-Dependent Electric Fields," *Phys.Rev.* **D78** (2008) 061701, `arXiv:0807.2785` `[hep-ph]`.

[275] B. Mihaila, F. Cooper, and J. F. Dawson, "Backreaction and Particle Production in (3+1)-dimensional QED," *Phys.Rev.* **D80** (2009) 014010, `arXiv:0905.1360` `[hep-ph]`.

[276] W.-B. Han, R. Ruffini, and S.-S. Xue, "Electron-positron pair oscillation in spatially inhomogeneous electric fields and radiation," *Phys.Lett.* **B691** (2010) 99–104, `arXiv:1004.0309` `[hep-ph]`.

[277] M. Orthaber, F. Hebenstreit, and R. Alkofer, "Momentum Spectra for Dynamically Assisted Schwinger Pair Production," *Phys.Lett.* **B698** (2011) 80–85, `arXiv:1102.2182` `[hep-ph]`.

[278] F. Hebenstreit, J. Berges, and D. Gelfand, "Simulating fermion production in 1+1 dimensional QED," *Phys.Rev.* **D87** (2013) 105006, `arXiv:1302.5537` `[hep-ph]`.

[279] V. Kasper, F. Hebenstreit, and J. Berges, "Fermion production from real-time lattice gauge theory in the classical-statistical regime," *Phys.Rev.* **D90** (2014) 025016, `arXiv:1403.4849` `[hep-ph]`.

[280] H. Casimir, "On the Attraction Between Two Perfectly Conducting Plates," *Indag.Math.* **10** (1948) 261–263.

[281] W. E. Lamb and R. C. Retherford, "Fine Structure of the Hydrogen Atom. Part I," *Phys.Rev.* **79** (1950) 549–572.

[282] **E144 Collaboration** Collaboration, C. Bula *et al.*, "Observation of nonlinear effects in Compton scattering," *Phys.Rev.Lett.* **76** (1996) 3116–3119.

[283] S. L. Adler, "Photon splitting and photon dispersion in a strong magnetic field," *Annals Phys.* **67** (1971) 599–647.

[284] S. Z. Akhmadaliev, G. Y. Kezerashvili, S. Klimenko, V. Malyshev, A. Maslennikov, *et al.*, "Delbruck scattering at energies of 140-450 MeV," *Phys.Rev.* **C58** (1998) 2844–2850.

[285] S. Akhmadaliev, G. Kezerashvili, S. Klimenko, R. Lee, V. Malyshev, *et al.*, "Experimental investigation of high-energy photon splitting in atomic fields," *Phys.Rev.Lett.* **89** (2002) 061802, `arXiv:hep-ex/0111084 [hep-ex]`.

[286] M. Bordag, U. Mohideen, and V. Mostepanenko, "New developments in the Casimir effect," *Phys.Rept.* **353** (2001) 1–205, `arXiv:quant-ph/0106045 [quant-ph]`.

[287] S. G. Karshenboim, "Precision physics of simple atoms: QED tests, nuclear structure and fundamental constants," *Phys.Rept.* **422** (2005) 1–63, `arXiv:hep-ph/0509010 [hep-ph]`.

[288] R. Jaffe, "The Casimir effect and the quantum vacuum," *Phys.Rev.* **D72** (2005) 021301, `arXiv:hep-th/0503158 [hep-th]`.

[289] F. Sauter, "Uber das Verhalten eines Elektrons im homogenen elektrischen Feld nach der relativistischen Theorie Diracs," *Z.Phys.* **69** (1931) 742–764.

[290] W. Heisenberg and H. Euler, "Consequences of Dirac's theory of positrons," *Z.Phys.* **98** (1936) 714–732, `arXiv:physics/0605038 [physics]`.

[291] J. Schwinger, "The Theory of Quantized Fields. VI," *Phys.Rev.* **94** (1954) 1362–1384.

[292] J. Schwinger, "The Theory of Quantized Fields. 5," *Phys.Rev.* **93** (1954) 615–628.

[293] XFEL, "XFEL," 2014. `http://xfel.desy.de/`.

[294] ELI, "ELI," 2014. `http://www.extreme-light-infrastructure.eu/`.

[295] J. Daugherty and I. Lerche, "Theory of Pair Production in Strong Electric and Magnetic Fields and Its Applicability to Pulsars," *Phys.Rev.* **D14** (1976) 340–355.

[296] L. Parker and J. Tiomno, "PAIR PRODUCING ELECTRIC FIELDS AND PULSARS,".

[297] Y. Luo and P. Ji, "Pair production induced by quantum electrodynamic vacuum polarization in pulsars," *Monthly Notices of the Royal Astronomical Society* **420** no. 2, (2012) 1673–1683. http://dx.doi.org/10.1111/j.1365-2966.2011.20158.x.

[298] F. PACINI, "Energy Emission from a Neutron Star," *Nature* **216** no. 5115, (Nov., 1967) 567–568. http://dx.doi.org/10.1038/216567a0.

[299] F. PACINI, "Rotating Neutron Stars, Pulsars and Supernova Remnants," *Nature* **219** no. 5150, (July, 1968) 145–146. http://dx.doi.org/10.1038/219145a0.

[300] J. E. GUNN and J. P. OSTRIKER, "Magnetic Dipole Radiation from Pulsars," *Nature* **221** no. 5179, (Feb., 1969) 454–456. http://dx.doi.org/10.1038/221454a0.

[301] P. Goldreich and W. H. Julian, "Pulsar electrodynamics," *Astrophys.J.* **157** (1969) 869.

[302] M. Ruderman and P. Sutherland, "Theory of pulsars: Polar caps, sparks, and coherent microwave radiation," *Astrophys.J.* **196** (1975) 51.

[303] F. C. Michel, "Theory of pulsar magnetospheres," *Rev.Mod.Phys.* **54** (1982) 1–66.

[304] S. L. Shapiro and S. A. Teukolsky, *Black Holes, White Dwarfs and Neutron Stars: The Physics of Compact Objects.* Wiley, 2008.

[305] V. S. Beskin, "MHD Flows in Compact Astrophysical Objects,".

[306] J. Cordes, N. Bhat, T. Hankins, M. McLaughlin, and J. Kern, "The brightest pulses in the universe: multifrequency observations of the crab pulsar's giant

pulses," *Astrophys.J.* **612** (2004) 375–388, `arXiv:astro-ph/0304495`
`[astro-ph]`.

[307] T. Hankins and J. Eilek, "Radio Emission Signatures in the Crab Pulsar,"
*Astrophys.J.* (2007) , `arXiv:0708.2505 [astro-ph]`.

[308] J. Taylor, R. Manchester, and A. Lyne, "Catalog of 558 pulsars,"
*Astrophys.J.Suppl.* **88** (1993) 529–568.

[309] A. Jessner, M. V. Popov, V. I. Kondratiev, Y. Y. Kovalev, D. Graham, *et al.*,
"Giant Pulses with Nanosecond Time Resolution detected from the Crab
Pulsar at 8.5 and 15.1 GHz," *Astron.Astrophys.* **524** (2011) A60,
`arXiv:1008.3992 [astro-ph.HE]`.

[310] J. K. Daugherty and A. K. Harding, "Electromagnetic cascades in pulsars,"
*apj* **252** (Jan., 1982) 337–347.

[311] A. V. Gurevich and I. N. Istomin, "Electron-positron plasma generation in a
pulsar magnetosphere," *Zhurnal Eksperimentalnoi i Teoreticheskoi Fiziki* **89**
(July, 1985) 3–21.

[312] J. Jackson, *Classical Electrodynamics.* Wiley, 1999.