University of California
Santa Barbara

Parity and Primordial Perturbations on the Path to New Physics

A dissertation submitted in partial satisfaction
of the requirements for the degree

Doctor of Philosophy
in
Physics

by

Amara McCune

Committee in charge:

Professor Nathaniel Craig, Chair
Professor Jeffrey Richman
Professor David Berenstein

September 2024

The Dissertation of Amara McCune is approved.

Professor Jeffrey Richman

Professor David Berenstein

Professor Nathaniel Craig, Committee Chair

March 2024

Parity and Primordial Perturbations on the Path to New Physics

by

Amara McCune

To Mom and Dad.

Acknowledgements

I am deeply grateful to my advisor, Nathaniel Craig. Nathaniel is an incredible physicist and an even better human. I thank him for his dedication to my learning and development as both a scientist and a person, as well as unwavering patience and kindness in his approach to mentorship and collaboration. In the down times of my Ph.D., his supporting words and sheer enthusiasm for physics inspired me to keep going. I learned a great deal from Nathaniel about physics, problem-solving, and how to build a friendly and collaborative scientific community. I truly believe I hit the advisor jackpot, and I feel honored to have been his student.

I would like to thank my many collaborators who worked with me on the projects throughout my Ph.D.: Nathaniel Craig, Isabel Garcia Garcia, Giacomo Koszegi, Soubhik Kumar, Lawrence Hall, Claudio Manzari, Quentin Bonnefoy, Christiane Scherb, Lian-Tao Wang, Reza Ebadi, Hanwen Tai, Radha Mastandrea, Dean Robinson, David Dunsky, and Linda Xu. I learned so much through each and every one of you.

About halfway through my Ph.D., I joined my advisor at Berkeley during his sabbatical and ended up staying through the completion of my degree. Thank you, Nathaniel, for allowing me to pursue this rather unusual Ph.D. trajectory, and to everyone at the Berkeley Center for Theoretical Physics and the theory group at Lawrence Berkeley Lab for their incredible hospitality through the course of my stay. I would especially like to thank Simon Knapen and Christian Bauer for their support.

There are also many outstanding mentor figures in my life who have helped me get to where I am today. As a high school student, my teachers Chris Wiemer and Cindy Carter were instrumental in helping me explore my early scientific interests, and were my official supervisors in several of my pursuits to self-study math and compete in high school science fairs. As an undergraduate, Leo Hollberg was the first professor I ever

worked with. Although I quickly learned that experimental research was not my strong suit, I loved working in Leo's lab, and have immensely benefited from his guidance and wisdom. Risa Wechsler, Lukasz Graczykowski, and Roger Blandford supervised my other undergraduate projects, and I am lucky to have learned from such phenomenal researchers. I would also like to thank several professors from both my undergraduate and graduate years that helped me navigate physics, provided advice, or even just offered up some encouraging words that they probably don't remember saying but were so important to me: Carl Wieman, Pat Burchat, Peter Graham, Flip Tanedo, and Raman Sundrum.

A Ph.D. is a roller coaster of an endeavor. I would not have gotten through this time without the mental health resources I have been very fortunate to have access to. I would like to thank the incredible Beth Kanne-Casselman and Lori Goldrich; the work you do is so important, and the completion of this thesis would not have been possible without your support.

Thank you as well to my peers at UCSB and Berkeley, especially especially Sean McBride, Diandian Wang, Alex Potts, Jenny Smith, Vincent Su, and Kevin Langhoff, for being invaluable problem set collaborators, office mates, and friends. I had the pleasure of many interesting physics discussions with Graham van Goffrier and Ian Everbach; thank you both for your encouragement and insights. I am deeply grateful to many wonderful friends who have been a constant source of love and light: Emily Peterson, Yana Boroshok, Preeti Kakani, Rebecca Bromley-Dulfano, Graham van Goffrier, Kat Hu, Fiona Kalensky, Ryte Rutkauskaite, Max Gallant, Lindsay Schmitz, Leila Taleghani, Faisal Asad, Jay Qu, Sierra Killian, Arjun Tambe, Eliza Cornell, Jake Gold, Bethany Suter, Amy Shen, Molly Kaplan, and Aria Chaderjian.

I would not be where I am today without the love and support of my wonderful parents, Denise Walder and Roark McCune. They made sure I grew up in a world of curiosity and learning, immersing our house in library books, art projects, museum visits,

musical instruments, and the nature of southwest Wisconsin. I am so grateful for these memories, as well as their unwavering support and encouragement to forge my own path. When I started reading popular science books on relativity and string theory at twelve, my dad began earmarking articles in science magazines and tracking down textbooks on supersymmetry that were far too advanced for my level of knowledge at the time (although, a decade later, I finally understood enough to crack one open). When I tried my hand at a research project in astrophysics at sixteen, my mom was there with me at 6am the day of my first science fair, helping me glue the last pieces of construction paper onto my tri-fold poster board. Through college and graduate school, they have always been there to support me at the points I felt like quitting. Thank you for everything you do, I love you both so much.

Thank you as well to my wonderful extended family. My cousin, Sam Walder, is one of my closest friends and an absolutely incredible human being. He has been my cheerleader and inspiration throughout my Ph.D. Thank you as well to the Marks family for welcoming me into your lives and home with open arms. I would not be completing this thesis without the love and support of Jacob Marks, whose brilliance is a constant source of inspiration, and his presence a continual source of laughter and light. Thank you for being by my side through the ups and downs of Ph.D. life, being my best friend and lifelong teammate, and always seeing the best in me, even when I couldn't see it myself.

Finally, I would like to acknowledge the generous help and support of several organizations throughout my education. The Jack Kent Cooke Foundation is an incredible non-profit that made my undergraduate education financially possible through the the Cooke College Scholarship, and supported my graduate education through the Cooke Graduate Scholarship. Their commitment to supporting students from lower-income backgrounds through college and graduate school changes so many lives, and my own

is no exception. I would also like to acknowledge the support of the DOE Office of Science Graduate Student Research Fellowship, the Graduate Opportunity Fellowship, the Brython-Davis Fellowship, and the Broida-Hirschfelder Fellowship. Thank you for believing in me and allowing me to pursue my passion.

Curriculum Vitæ

Amara McCune

Education

2024	Ph.D. in Physics (Expected), University of California, Santa Barbara.
2021	M.A. in Physics, University of California, Santa Barbara.
2018	B.S. in Physics and Mathematics, Stanford University.

Publications

"P not PQ"
N. Craig, I. Garcia Garcia, G. Kozsegi, A. McCune
JHEP **09**, 130 (2021), [arXiv:2012.13416 [hep-ph]]

"The Muon Smasher's Guide"
H. Al Ali *et al.* including A. McCune
Rept. Prog. Phys. **85**, no.8, 084201 (2022), [arXiv:2103.14043 [hep-ph]]

"Gravitational Waves from Stochastic Scalar Fluctuations"
R. Ebadi, S. Kumar, A. McCune, H. Tai, L.T. Wang
Phys. Rev. D **109**, no.8, 083519 (2024), [arXiv:2307.01248 [astro-ph.CO]]

"Dark Matter in a Mirror Solution to the Strong CP Problem"
Q. Bonnefoy, L.J. Hall, C.A. Manzari, A. McCune, C. Scherb
Phys. Rev. D **109**, no.5, 055045 (2024), [arXiv:2311.00702 [hep-ph]]

"An Effective Cosmologiacl Collider"
N. Craig, S. Kumar, A. McCune
JHEP **07**, 108 (2024), [arXiv:2401.10976 [hep-ph]]

Abstract

Parity and Primordial Perturbations on the Path to New Physics

by

Amara McCune

The interplay between particle physics models and their cosmological implications provides a crucial set of constraints for model-builders. As upcoming experiments set out to determine the particle nature of dark matter, detect primordial gravitational waves (GWs), and map primordial non-Gaussianities (PNGs), a unique opportunity arises to search for physics beyond the Standard Model (BSM) at scales that may be as high as 10^{13} GeV. This thesis explores these three examples of cosmological aspects of model-building. We begin with a review of the problems that motivate the search for BSM physics, touching on both the naturalness problems and the inflationary paradigm. Then, we introduce a parity-based solution to the strong CP problem, motivating the question of whether a model containing a heavy mirror copy of the SM gauge group may contain a viable candidate for dark matter. We then turn to GWs, introducing a new source of primordial GWs from stochastic fluctuations of scalar fields in the early universe. Finally, we turn to effective field theories (EFTs) in de-Sitter space, introducing the problem in the context of cosmological collider physics and initiating a systematic understanding of EFT construction in inflationary spacetimes. In turn, we illustrate the promise of near-future cosmological probes to provide insights into the underlying physics of our universe.

Permissions and Attributions

1. The content of chapter 2 is the result of a collaboration with Nathaniel Craig, Isabel Garcia Garcia, and Giacomo Koszegi. This work previously appeared in the Journal of High Energy Physics (JHEP 09 (2021), 130).

2. The content of chapter 3 is the result of a collaboration with Quentin Bonnefoy, Lawrence Hall, Claudio Andrea Manzari, and Christiane Scherb. This work previously appeared in Physical Review D (Phys.Rev.D 109 (2024) 5, 055045).

3. The content of chapter 4 is the result of a collaboration with Reza Ebadi, Soubhik Kumar, Lian-Tao Wang, and Hanwen Tai. This work previously appeared in Physical Review D (Phys.Rev.D 109 (2024) 8, 083519).

4. The content of chapter 5 is the result of a collaboration with Nathaniel Craig and Soubhik Kumar. This work previously appeared in the Journal of High Energy Physics (JHEP 07 (2024), 108).

Preface

We begin this thesis with an extensive pedagogical introduction to the mathematical foundations of the Standard Model (SM), focusing on understanding quantum field theory through a group-theoretic lens. We then review concepts in effective field theory (EFT), building up the techniques of Wilsonian renormalization and emphasizing the origin of operator redundancies in enumerating a minimal EFT basis. We then go into the ΛCDM model of cosmology, detailing the formalism of working with cosmological perturbations and introducing the inflationary paradigm. Finally, we go over several key problems that motivate the search for physics beyond the Standard Model (BSM), including the naturalness problems, dark matter, and the microscopic considerations of inflation. Readers familiar with this material may wish to proceed to chapter 2.

Although the strong CP problem is not the central feature of this thesis, I believe chapter 2 provides an important overview of parity solutions to the strong CP problem and their phenomenological features. The following chapter builds upon this work, presenting a model that extends these ideas to feature a mirror copy of the entire SM gauge group. This first half of the thesis underscores the importance of parity symmetry in model-building, showing how models based in parity can solve key SM puzzles and lead to a rich phenomenology that may be visible to near-future experiments.

The next half of the thesis, as the title suggests, focuses on primordial perturbations. Chapter 4 considers a new source of gravitational waves, while chapter 5 then turns

towards interactions of heavy particles during inflation. Both works were inspired by the promise of cosmological collider physics, and seek to connect the considerations of early-universe model building to cosmological observables that may be visible in an upcoming suite of experiments.

Units Throughout this thesis, we will employ natural units in which $\hbar = c = 1$ unless otherwise stated. Chapter 4 will explicitly write factors of H, the Hubble parameter, while chapter 5 will not, in order to simplify notation. These factors can always be restored via dimensional analysis.

Contents

Chapter 1

A Tale of Two Standard Models

Over the past century, humanity has constructed two crowning achievements of science: The Standard Model (SM) of particle physics and the ΛCDM model of cosmology, representing our current best understanding of the large-scale structure and fundamental building blocks of our universe. Our observations span nearly fifty orders of magnitude, from the smallest experimentally confirmed particles to the size of the observable universe itself.

These models are astonishingly precise in the phenomena they are able to capture, while utilizing relatively simple and compact mathematical structures. This itself, as captured by the above quote, is remarkable. Part of what makes this comprehension possible is a separation of scales. We don't need to know about quantum mechanics to compute the trajectory of a ball thrown up in the air on Earth, a common problem in introductory physics courses. Similarly, computations of scattering cross sections at particle accelerators do not typically consider the effects of gravity, as its influence at this scale is negligible. In other words, these degrees of freedom decouple from one another.

There are cases, however, in which we cannot consider these regimes to be independent. Black holes are a classic example; matter that falls into a black hole is subject

to extreme spacetime curvature while being compressed toward a singularity. Relativistic particles in strong gravitational fields require the considerations of both relativistic quantum field theory (QFT) and general relativity. Other astrophysical objects, such as neutron stars and white dwarfs, also present such environments. A particularly interesting intersection of these domains lies in the beginning of the universe itself. Signatures of these phenomena are therefore crucial testing grounds for new ideas in high-energy physics, exploring the behavior of matter at higher energies or curved spacetimes at which our two "standard models" fail to accurately converge.

Aside from the omission of gravity – incorporating only three of the four known fundamental forces – there are many reasons to suspect that the SM is not the final story. Much of the observed energy density in the universe also seems to remain unknown, with baryonic matter comprising only about 5 percent of the universe's mass-energy budget. The rest is wrapped up in dark matter and dark energy, whose constituents and underlying mechanisms remain unknown. Although the inflationary paradigm provides a compelling explanation as to why the universe is isotropic and homogeneous on large scales, the specific microphysics by which inflation occurs remains unknown.

The SM itself, however, also presents its puzzles. For example, why is the Higgs mass measured to be 125 GeV, despite the possibility of substantial quantum corrections? Why is there no apparent CP violation in the strong sector, despite the fact that it is not a priori prohibited? What sets the value of the vacuum energy density, which is predicted by standard methods in QFT to be some 50 or more orders of magnitude higher than its measured value? This collection of puzzles, known as the naturalness problems, seek to explain the existence of a large separation of scales, which historically has indicated the presence of a new symmetry or degree of freedom.

Finally, there are the many questions that one may expect to be answered in a Theory of Everything. In particular, we can ask what the origin of the SM's structure is.

The SM's free parameters, namely its masses and mixing angles, have no known origin, but could be explained by a more encompassing theory. The quarks and leptons are arranged in a particular fashion, with each containing three generations, two families, and a distinct, hierarchical structure in their mixing matrix. Where this hierarchy comes from constitutes the SM's flavor puzzle and remains an active area of research.

1.1 The Standard Model of Particle Physics

The first standard model we encounter is the one that is more often known as the Standard Model (SM); the cosmological standard model is more often referred to as ΛCDM model of cosmology. The SM of particle physics is made of up 17 fundamental particles and 3 fundamental forces — the electromagnetic, strong, and weak forces — under which these particles are charged. There is a surprising amount of structure involved: There are 6 quarks and 6 leptons that can each be grouped into 3 generations that interact in the same way under the electromagnetic and strong forces. Most of the SM's free parameters relate to this structure — such as the quark masses and mixing angles — and have we have no verified explanation of where it comes from.

1.1.1 From Particles to Fields and Back Again

The idea of a particle as a tiny, indivisible element of matter dates back to the 6th century BC. It wasn't until the early 1900s that, following centuries of observations and modeling of both the wave-like and particle-like properties of light, the idea of discrete quanta was proposed by Max Planck as a resolution to "ultraviolet catastrophe" of black-body radiation.

The advent of quantum field theory (QFT) in the subsequent decades formally introduced the idea that a particle is not the most fundamental object comprising matter, but

rather an excited state of an underlying quantum field. In this picture, one can quantize a classical field theory by promoting the field variables to operators with canonical commutation relations, then building the Hilbert space by applying creation and annihilation operators to the field operators. Feynman showed the equivalency of this approach to the path integral formalism, in which an amplitude is computed by exponentiating the action and summing over all possible paths. There are upsides and downsides to each formalism. Recent research in high energy theory has included the development of spinor-helicity formalism to compute scattering amplitudes.

While the quantum field-theoretic notion of a particle is certainly useful, it is not predictive. More predictive power came along in the 1930s when Wigner formulated his namesake theorem, relating the Hilbert space of quantum states to the symmetry transformations they are invariant under. This perspective led to a sharper definition: A particle is an irreducible representation of the Poincaré group. In the rest of this section, we will interrogate this definition, exploring the beautiful relationship between particle physics, group theory, and representation theory along the way. Much of this discussion is inspired by [1] [2].

Representation Theory

To start, we will build up some mathematical machinery, starting with the notion of a group. A *group* G is a set with an associated group action, which we will denote $\times$, that obeys the axioms:

1. There is an identity element $e \in G$ such that $e \times g = g \times e = g$ for each element $g \in G$.

2. For each element $g \in G$, there is an inverse element $g^{-1} \in G$ such that $g \times g^{-1} = g^{-1} \times g = I$.

3. For elements $g_1, g_2, g_3 \in G$, $(g_1 \times g_2) \times g_3 = g_1 \times (g_2 \times g_3)$. In other words, the group action $\times$ is associative.

The group action $\times$ is commonly either addition or multiplication, but can be any other operation that satisfies the axioms. We define G as *abelian* if, for $g_1, g_2 \in G$, $g_1 \times g_2 = g_2 \times g_1$, i.e the group action is commutative on elements of G. The group is *non-abelian* if it is non-commutative.

Given two groups G and H, a map $f : G \to H$ is a *group homomorphism* if the group structure of G is preserved,

$$f(g_1 \times g_2) = f(g_1) \times f(g_2) \ \ \forall \ \ g_1, g_2 \in G. \tag{1.1}$$

The group homormorphism further constitutes a *representation* of G if it maps G onto the *general linear group* $GL(V)$ for a field F,

$$f : G \to GL(V), \tag{1.2}$$

where again the group structure of G is preserved. For an n-dimensional vector space V, $GL(V)$ is a group consisting of the set of all n-by-n invertible matrices plus the group action of matrix multiplication that takes $V \to V$. Representations of groups are important because allow us to study entities which may be abstract — for example, the group of permutation operations on a set of objects — as familiar, linear matrices. A group may admit multiple representations, i.e. there is usually more than one map from G to $GL(V)$ that retains the group structure of G. For finite groups, one must choose a basis for the vector space V; it is common to denote $GL(V)$ by $GL(n, F)$, where n is the matrix dimension and F is the field over which the vector space is defined.

Let's consider a simple example of the group of permutations of the set of numbers

$\{1,2,3\}$. There are $3! = 6$ ways to permute this set:

$$\begin{aligned} &\{1,2,3\} \to \{1,2,3\}, \quad \{1,2,3\} \to \{1,3,2\}, \quad \{1,2,3\} \to \{2,1,3\} \\ &\{1,2,3\} \to \{2,3,1\}, \quad \{1,2,3\} \to \{3,2,1\}, \quad \{1,2,3\} \to \{3,1,2\} \end{aligned} \tag{1.3}$$

The group of permutations consists of the permutations themselves. Adopting cyclic notation[1], there are 5 group elements,

$$p_1 = (1)(2)(3), \quad p_2 = (12)(3), \quad p_3 = (13)(2), \quad p_4 = (1)(23), \quad p_5 = (123), \quad p_6 = (132). \tag{1.4}$$

This permutation group is known as S_3. Note that this indeed forms a group: We've included the identity element, each element has an inverse (the first three elements listed are their own inverses, while the last two are each other's inverses), and composing any two permutations is guaranteed to result in another permutation of the group. Not all p_i, p_j commute, and so the group is non-abelian.

To form a representation of S_3, we need to find a mapping $S_3 \to GL(V)$ that preserves the structure of S_3. For example, we see that the $p_2 p_3 = p_5$, and so the corresponding representation matrices M_2, M_3, M_5 must obey $M_2 M_3 = M_5$. Let's find a representation that maps,

$$f : S_3 \to GL(1, \mathbb{R}). \tag{1.5}$$

The choice of $M_k = 1$ for all k is an easy way to satisfy the group structure. This is an example of a *trivial representation*.

To find more complicated representations, we introduce the idea of the *generating set* of a group G: This is a subset of the group S in which any element of G can be expressed

[1] In this notation, (1 2 3) denotes the permutation that exchanges $1 \to 2$, $2 \to 3$, and $3 \to 1$. The numbers contained in a cycle, denoted by closed parentheses, are exchanged with one another in this way.

as a combination of elements in S or their inverses. For S_3, we can see that all of the patterns of 1.3 can be achieved through successive applications of p_2 and p_4 — these constitute the generating set. Now, let's consider V to be a 3-by-3 dimensional vector space, with basis vectors e_1, e_2, e_3. Explicitly, the map is $f : S_3 \to GL(3, \mathbb{R})$. Then, p_1 should map onto the 3-by-3 identity matrix, p_2 should swap the first and second basis vectors, and so on. We arrive at the representation,

$$\begin{aligned} p_1 &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad p_2 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad p_3 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \\ p_4 &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad p_5 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \quad p_6 = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \end{aligned} \tag{1.6}$$

This is the *defining representation*; we can verify via matrix multiplication that it preserves the S_3 group structure. This representation is an example of a *reducible representation* because it has an invariant subspace: Multiplying any p_i by the vector $v_{sub} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ returns v_{sub}.

Formally, a representation is reducible if it has non-trivial *sub-representation*: Given a representation $f : G \to GL(V)$, if there exists a vector subspace $U \subset V$ such that $f(g)u \in U$ for all $u \in U$, restricting f to the subspace U constitutes a sub-representation. For example, if we have two representations, $f_1 : G \to GL(V_1)$ and $f_2 : G \to GL(V_2)$, over the same field F, we can define a representation $f_1 \oplus f_2 \to GL(V_1 \oplus V_2)$ by the direct sum,

$$(f_1 \oplus f_2)(g) = f_1(g) \oplus f_2(g) = \begin{pmatrix} f_1(g) & \mathbf{0} \\ \mathbf{0} & f_2(g) \end{pmatrix}. \tag{1.7}$$

This new representation is of course reducible, and f_1 and f_2 are sub-representations. If a representation does not contain non-trivial sub-representations, it is *irreducible*.

There is an easy way to see if a representation is reducible or not. we first need to know its *conjugacy classes*. We say that two elements $g_1, g_2 \in G$ are *conjugate* to one another if there exists an element $g_3 \in G$ such that $g_1 = g_3 g_2 g_3^{-1}$, where g_3 does not need to be distinct from g_1, g_2; a conjugacy class is then a set of conjugate group elements.[2] For S_3, one can work out that there are three conjugacy classes: The identity element p_1 constitutes one class, the elements consisting of a single permutation p_2, p_3, and p_4 constitute the second class, and the elements consisting of two permutations p_5 and p_6 constitute the third.

Given the representation f, we can define the *character value* of f by the map $\chi : G \to \mathbb{C}$ such that

$$\chi^{(f)}(g) = \mathrm{Tr}[f(g)]. \tag{1.8}$$

All elements of a group that belong to the same conjugacy class have the same character value. *Equivalent representations* — where two representations $h_1 : G \to GL(V)$ and $h_2 \to GL(V)$ are defined as equivalent if there exists a matrix A such that $h_1(g) = A^{-1} h_2(g) A$ for all $g \in G$ — also have the same character value. We can assemble these character values into a *character vector* for each representation, which lives in a complex vector space with dimension equal to the number of conjugacy classes. [3]

Now, we are ready to state an important theorem for irreducibility: A representation

[2] This is an example of an *equivalence relation*: A binary relation between set elements that is reflexive, symmetric, and transitive. Given this equivalence relation, we can form equivalence classes, i.e. split up our set of group elements into sets that are equivalent, where g_1 and g_2 belong to the same equivalence class if and only if they are equivalent. The equivalence classes will necessarily be disjoint.

[3] The representation theory literature employs the term "character" for both the trace of the representation matrix for a particular conjugacy class and the vector assembling the traces of each conjugacy class; I've separated these definitions here for clarity.

f of a finite group G is irreducible if and only if

$$\sum_{i=1}^{c} n_i |\chi_i^{(f)}|^2 = |G|, \tag{1.9}$$

where c is the number of conjugacy classes of G and n_i is the number of elements in the ith conjugacy class. Let's test this for the defining representation of S_3. We find,

$$\sum_{i=1}^{c} n_i |\chi_i^{(f)}|^2 = (1)(3)^2 + (3)(1)^2 + (2)(0)^2 > |G| = 6, \tag{1.10}$$

and so this representation is indeed reducible. The trivial representation is always irreducible,

$$\sum_{i=1}^{c} n_i |\chi_i^{(f)}|^2 = (1)(1)^2 + (1)(1)^2 + (1)(1)^2 + (1)(1)^2 + (1)(1)^2 + (1)(1)^2 = |G| = 6. \tag{1.11}$$

We can then count the number of irreps. The representations above are unitary and, given two inequivalent unitary irreps, their corresponding matrices must be orthogonal. This follows from Schur's lemma, which states that, for two irreps $h_1 : G \to GL(V_1)$ and $h_2 : G \to GL(V_2)$, if there is a matrix A such that

$$A h_1(g) = h_2(g) A \tag{1.12}$$

for all $g \in G$, then either A is a square invertible matrix and the irreps are equivalent or $A = 0$. Then, for matrix elements of inequivalent irreps $h_1(g)$ and $h_2(g)$,

$$\sum_{g \in G} h_1^\dagger(g)_{ij} h_2(g)_{kl} = 0 \; \forall \; i, j, k, l. \tag{1.13}$$

One can use Schur's second lemma [4] to show more generally that, for any two irreps $h_1(g)$ and $h_2(g)$,

$$\sum_{g \in G} h_1^\dagger(g)_{ij} h_2(g)_{kl} = \frac{|G|}{d} \delta_{il} \delta_{jk} \delta_{h_1 h_2} \; \forall \; i, j, k, l, \tag{1.14}$$

[4] $A = \lambda I$, $\lambda \in \mathbb{C}$ for $h_1 = h_2$, where I is the identity matrix of appropriate dimensionality.

where d is the dimension of the irreps (if equivalent). Letting $i = j$ and $k = l$ and summing over i, k yields,

$$\begin{aligned}\sum_{i,k}\sum_{g\in G} h_1^\dagger(g)_{ii}h_2(g)_{kk} = \sum_{g\in G}\mathrm{Tr}[h_1^\dagger(g)_{ii}]\mathrm{Tr}[h_2(g)_{kk}] &= \sum_{g\in G}\left(\chi^{(h_1)}(g)\right)^*\chi^{(h_2)}(g) \\ &= \sum_{i,k}\frac{|G|}{d}\delta_{ik}\delta_{ik}\delta_{h_1h_2} = |G|\delta_{h_1h_2}.\end{aligned} \tag{1.15}$$

Because all group elements in the same conjugacy class have the same character value, we can equivalently write this result as,

$$\sum_{i=1}^{c} n_i\left(\chi_i^{(h_1)}\right)^*\chi_i^{(h_2)} = |G|\delta_{h_1h_2}, \tag{1.16}$$

which implies that the character vectors of inequivalent irreps are orthogonal in the vector space of conjugacy classes. These irreducible character vectors then form a basis in this vector space, and the character vectors for *reducible* representations can be written as a linear combination of them. We can also think of the vector space of irreducible representations, which has a number of dimensions ρ equal to the number of irreps. One can work out[5],

$$\sum_{\alpha=1}^{\rho}\chi_i^{(\alpha)^*}(g)\chi_j^{(\alpha)}(g) = \frac{|G|}{n_i}\delta_{ij}, \tag{1.17}$$

which tells us that This implies key result: The number of inequivalent irreps of a finite group is equal to the number of conjugacy classes of the group. Finally, we can use the fact that there will always be a trivial representation with a number of elements equal to $|G|$ to find the dimension of each irrep,

$$\sum_{\alpha=1}^{c}\chi_1^{(\alpha)^*}(g)\chi_1^{(\alpha)}(g) = \sum_{\alpha=1}^{c} d_\alpha^2 = |G|. \tag{1.18}$$

For S_3, this means that there are 3 irreps, one of which is the trivial represenation which has dimension 1. This leaves us with 2 irreps to find, of which one should be

[5]The proof of this is rather involved, and we will not reproduce it here. It is, dare I say, an exercise left to the reader.

dimension 1 and the other dimension 2. We will eventually connect this to physics, so we will focus on unitary irreps. Then, we can employ 1.9 to get the constraint,

$$\chi_a^2 + 3\chi_b^2 + 2\chi_c^2 = 6, \tag{1.19}$$

for the character values χ_a, χ_b, and χ_c corresponding to the identity conjugacy class, 2-cycle conjugacy class, and 3-cycle conjugacy class, respectively. We can then use the orthogonality relation of 1.16 and take h_1 to be the trivial representation to obtain,

$$\chi_a + 3\chi_b + 2\chi_c = 0. \tag{1.20}$$

This uniquely fixes the choice $\chi_a = 1$, $\chi_b = -1$, $\chi_c = 1$.[6]

The orthogonality relations can be used to construct a very useful device, the *character table* for a group, which organizes the character values for each representation according to their conjugacy class. The character table for S_3, based on what we know so far, is:

	identity	2-cycles	3-cycles
trivial	1	1	1
sign	1	-1	1
2D	2	a	b

The top row of the table gives the conjugacy class, while the leftmost column gives the irrep. For the 2-dimensional representation, which we have not found yet, we've used the variables a, b to label unknown characters. We do know that the trace of the 2-by-2 identity matrix is 2, and so we automatically know the character of the identity conjugacy class. We can find a and b with a nice trick: The orthogonality relations imply that the

[6]Technically, based on these constraints alone, this choice is fixed up to a sign. The ambiguity is resolved by recalling that this irrep still has to satisfy the properties of being a representation. In this case, it turns out that the opposite sign choice does not result in a valid homomorphism. The valid homorphism is the one that maps even permutations to 1 and odd permutations to -1. This is known as the sign representation.

rows and columns of the character table should be orthogonal. The completed table is then:

	identity	2-cycles	3-cycles
trivial	1	1	1
sign	1	-1	1
2D	2	0	-1

At last, we can use the character table to construct the 2-dimensional representation. We know there is one identity element, three 2-cycles, and two 3-cycles. We know their traces from the character values, we know that their matrix elements satisfy 1.13, and we know that p_2 and p_4 constitute the generating set of S_3. In general, we can use the fact that the representation must maintain the group structure and the representation is unitary to piece together that the 2-dimensional representation consists of,

$$\begin{gathered} p_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad p_2 = \begin{pmatrix} -1/2 & \sqrt{3}/2 \\ \sqrt{3}/2 & 1/2 \end{pmatrix}, \quad p_3 = \begin{pmatrix} -1/2 & -\sqrt{3}/2 \\ -\sqrt{3}/2 & 1/2 \end{pmatrix} \\ p_4 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad p_5 = \begin{pmatrix} -1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & -1/2 \end{pmatrix}, \quad p_6 = \begin{pmatrix} -1/2 & \sqrt{3}/2 \\ -\sqrt{3}/2 & -1/2 \end{pmatrix}. \end{gathered} \tag{1.21}$$

In this section, we've studied the group S_3 and its representations, focusing on permutations of the set of numbers $\{1, 2, 3\}$. A group, however, is an abstract entity, and the "permuted objects" of the set can really be anything we want. For example, we can imagine each element being a corner of an equilateral triangle; the permutations then correspond to rotations and reflections of this triangle. In particular, we can see that the 2-cycle permutations correspond to the three different ways of reflecting the triangle, while the 3-cycle permutations correspond to the two directions in which to rotate it. Because of this direct correspondence to symmetry transformations, the groups S_n are more generally known as symmetric groups.

The S_n groups are quite fundamental in the study of group theory. This is due to *Cayley's theorem*, which states that every group G of order n is isomorphic to a subgroup of S_n. This is a remarkable result, because it means that we can map every finite group $G \to S_n$, and that the group structure can be characterized by bijective mappings $G \times G \to G$. It also allows us to define a *regular representation* for each G based in our knowledge of S_n. This is found as follows. The dimension of the regular representation corresponds to the number of elements in the group, e.g. for S_3, the regular representation consists of six 6-by-6 matrices. Each element is assigned a basis vector: $\{e, e_{(12)(3)}, e_{(13)(2)}, e_{(1)(23)}, e_{(123)}, e_{132}\}$. The matrix corresponding to each element can then be constructed by working out how that particular element acts on the other group elements and permuting the matrix columns accordingly. For example, the matrix for p_2 is,

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}, \tag{1.22}$$

because $p_2 p_1 = p_2$, $p_2 p_2 = p_1$, $p_2 p_3 = p_5$, and so on.

The regular representation is particularly important because it can be decomposed exactly as the direct sum of all irreps of the group,

$$f_{reg} = \bigoplus_{\alpha=1}^{c} d_\alpha f_{(\alpha)}, \tag{1.23}$$

where α labels each irrep and d_α is its corresponding dimension.

Lie Groups and Lie Algebras

We've now built up our study of S_n and seen that groups are a powerful tool for studying symmetries. S_n groups are discrete, and so the rotations and reflects they describe occur about fixed axes. What if instead we are interested in describing, for example, the rotations and reflections of a circle? In this case there are no fixed axes, and an infinite number of rotations and reflections can leave the circle invariant. The group describing these transformations must be continuous.

S_3 is a discrete group, and so the rotations and reflections it describes occur about fixed axes. What if instead we are interested in describing, for example, the rotations and reflections of a circle? In this case there are no fixed axes, and an infinite number of rotations and reflections one could perform that leave the circle invariant. The group describing these transformations must then be continuous, and we can think of it as a continuous space M with dimension corresponding to the number of parameters the group elements depend on. The parameters themselves must also be continuous. In order to associate each point in M with another, in the mapping $M \to M$, we must be able to "walk" along this space without encountering any discontinuities. We must also be able to describe the change from one point to another infinitesimally close point. These requirements restrict M to be a smooth, differentiable manifold. Such groups are known as *Lie groups*. Formally, a Lie group is a group that is also a smooth, differentiable manifold, endowed with group actions of multiplication $\mu : M \times M \to M$ and inversion $\iota : M \to M$ that are smooth maps.

One Lie group we have already encountered is $GL(n, \mathbb{C})$, the group of $n \times n$ invertible matrices. These matrices are described by n^2 parameters which may be complex, and so the dimension of this group is $2n^2$.[7] This Lie group is *closed*, meaning that its underlying

[7]It is useful to emphasize that while the group elements are $n \times n$ matrices, the dimension of the group is given by the number of parameters needed to describe an element, and the dimension of the

manifold is boundaryless and compact[8]. All closed Lie groups are subgroups of $GL(n, \mathbb{C})$, and therefore also matrix groups. Matrix groups we encounter over and over again in the study of particle physics include:

1. $O(n)$: The group of all $n \times n$ orthogonal matrices. Its dimension is $\frac{n(n-1)}{2}$.

2. $U(n)$: The group of all $n \times n$ unitary matrices. Its dimension is n^2.

3. $SO(n)$: The group of all $n \times n$ orthogonal matrices with determinant equal to 1. Its dimension is $\frac{n(n-1)}{2}$.

4. $SU(n)$: The group of all $n \times n$ unitary matrices with determinant equal to 1. Its dimension is $n^2 - 1$.

5. Spin(n): A group that is a double cover[9] of $SO(n)$. It therefore also has dimension $\frac{n(n-1)}{2}$.

Each of these groups can be thought of as groups of transformations on some *target space*, the space of parameters describing the group. For example, the group elements of $SO(2)$ are given by,

$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}, \tag{1.24}$$

for $\theta \in \mathbb{R}$. This is the familiar matrix of rotations in $\mathbb{R}^2$. We can choose any parametrization, though it is useful to choose a set of parameters β_b, with $b = 1, ..., N$, such that the

matrices of matrix Lie group elements are often not equal to the dimension of the group itself.

[8] A topological space is *compact* if its parameters vary over a closed interval.

[9] The *cover* of a group G is another group H that maps onto G in a way that maintains certain structures of G. Formally, H covers G if there is a surjective homomorhism $\phi : H \to G$ such that the kernel of ϕ (the set of group elements that ϕ maps to the identity) is a normal subgroup (a subgroup that is invariant under conjugation of $g \in G$) and the quotient $H/ker(\phi)$ is isomorphic to G. A double cover is a cover in which $|H| = 2|G|$.

group element for $\beta = 0$ is the identity,

$$g(\beta)\big|_{\beta=0} = e, \tag{1.25}$$

where β is to be interpreted as the total set of β_b. The representation f then obeys $f(\beta)|_{\beta=0} = 1$. The Taylor expansion of $f(\beta)$ about $\beta = 0$ is then,

$$f(\beta) = f(\beta)\big|_{\beta=0} + \beta_b \left(\frac{\partial}{\partial\beta} f(\beta_b)\right)\big|_{\beta=0} + ... = 1 + \beta_b \left(\frac{\partial}{\partial\beta_b} f(\beta)\right)\Big|_{\beta=0} + ... \tag{1.26}$$

where we sum over b. This is more commonly written as,

$$f(\beta) = 1 + i\beta_b X_b + ... \tag{1.27}$$

where we define

$$X_b \equiv -i \left(\frac{\partial}{\partial\beta_b} f(\beta)\right)\Big|_{\beta=0} \tag{1.28}$$

to be the *group generators*. We will see that these generators indeed generate the group. Consider a set of infinitesimal parameters β_b/k, where k is large. An infinitesimal group element near the identity is then given by,

$$f(d\beta) = 1 + i\left(\frac{\beta_b}{k}\right) X_b. \tag{1.29}$$

We can then recover any group element by raising the infinitesimal one to the power of k,

$$\left(1 + i\left(\frac{\beta_b}{k}\right) X_b\right)^k = 1 + i\beta_b X_b - \frac{\beta_b^2(k-1)}{2k} X_b^2 - \frac{i\beta_b^3(k-2)(k-1)}{6k^2} X_b^3 + ... \tag{1.30}$$

In the limit $k \to \infty$, this becomes

$$\lim_{k\to\infty} \left(1 + i\left(\frac{\beta_b}{k}\right) X_b\right)^k = 1 + i\beta_b X_b - \frac{1}{2}\beta_b^2 X_b^2 - \frac{i}{3!}\beta_b^3 X_b^3 + ... = e^{i\beta_b X_b}. \tag{1.31}$$

This means that we can write the group elements directly in terms of the generators.

Now, let's say we have two elements of the same group $g_1, g_2 \in G$ in the representation f. According to this result, we can write the elements as $f(g_1) = e^{i\beta_a X_a}$ and $f(g_2) = e^{i\beta_b X_b}$, where g_1 and g_2 may be generated by different generators (or combinations thereof). While they are parameterized in the same way, they will not in general share the same parameter values. Multiplying these elements yields,

$$f(g_1)f(g_2) = e^{i(\beta_a X_a + \beta_b X_b)}. \tag{1.32}$$

We know that $f(g_1)f(g_2)$ should yield some $f(g_3)$, and so there must be some parameter choices β_3 and combination of generators X_3 such that,

$$e^{i(\beta_a X_a + \beta_b X_b)} = e^{i\beta_c X_c}. \tag{1.33}$$

This implies that the generators X_b must form an algebra, which we call the *Lie algebra* of the group. We can see why this is by finding the commutator of the group generators. We can write,

$$\ln\left(1 + e^{i\beta_a X_a} e^{i\beta_b X_b} - 1\right) = i\beta_c X_c. \tag{1.34}$$

Taylor expanding, we find,

$$\begin{aligned}
&\left(e^{i\beta_a X_a} e^{i\beta_b X_b} - 1\right) - \frac{1}{2}\left(e^{i\beta_a X_a} e^{i\beta_b X_b} - 1\right)^2 + ... \\
&= \left(i\beta_a X_a + i\beta_b X_b - \beta_a X_a \beta_b X_b - \frac{1}{2}\beta_a^2 X_a^2 - \frac{1}{2}\beta_b^2 X_b^2\right) - \frac{1}{2}\left(i\beta_a X_a + i\beta_b X_b - ...\right)^2 + ... \\
&\qquad = \left(i\beta_a X_a + i\beta_b X_b - \beta_a X_a \beta_b X_b - \frac{1}{2}\beta_a^2 X_a^2 - \frac{1}{2}\beta_b^2 X_b^2\right) \\
&\qquad - \frac{1}{2}\left(-\beta_a^2 X_a^2 - \beta_a X_a \beta_b X_b - \beta_b^2 X_b^2 - \beta_b X_b \beta_a X_a + ...\right) + ... \\
&\qquad\qquad = i\beta_a X_a + i\beta_b X_b - \frac{1}{2}\beta_a X_a \beta_b X_b + \frac{1}{2}\beta_b X_b \beta_a X_a + ... = i\beta_c X_c.
\end{aligned} \tag{1.35}$$

This yields the commutator,

$$[\beta_a X_a, \beta_b X_b] = 2i\left(\beta_a X_a + \beta_b X_b - \beta_c X_c\right) + ... \tag{1.36}$$

where the ellipses denote further terms in all expansions. The result, leaving the details to [2] is,

$$[X_a, X_b] = if_{abc}X_c \tag{1.37}$$

for generators X_a, X_b, and X_c, and constants f_{abc} known as structure constants. These can be computed for any nontrivial representation and define the algebra for each group.

The Lie algebra is typically easier to work with than the Lie group, because the algebra is a vector space. Formally, the Lie algebra of a group G is the vector space $\mathfrak{g}$ over a field F, endowed with a binary operation called the Lie bracket that maps $[\ ,\ ] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$. The bracket satisfies the properties:

1. Bilinearity: [ax + by, cz] = [ax, cz] + [by, cz] for all $a, b \in F$ and $x, y, z \in \mathfrak{g}$.

2. Asymmetry: [x,y] = -[y,x] for all $x, y \in \mathfrak{g}$.

3. The Jacobi identity: [x, [y,z]] + [z, [x,y]] + [y, [x,z]] = 0 for all $x, y, z \in \mathfrak{g}$.

To recap, we've introduced the related concepts of both Lie groups and Lie algebras. A Lie group is a group that is also a smooth differentiable manifold, and we can apply all of the tools of representation theory to them. The group elements can be written in terms of infinitesimal group generators, which are defined such that there is one generator for each continuous parameter of the group. The set of all linear combinations of the generators constitutes a vector space which is the Lie algebra. Given the Lie algebra, one can find the group elements by exponentiating the generators according to 1.31. Notationally, a Lie group is written as e.g. $SU(n)$, whereas its corresponding Lie algebra is written as $\mathfrak{su}(\mathfrak{n})$. Each Lie group has a corresponding Lie algebra, and multiple Lie groups may share the same Lie algebra. We can construct representations of the Lie algebra in analogous way to Lie groups as the mappings $\mathfrak{f} : \mathfrak{g} \to \mathfrak{gl}(V)$ such that the

algebra homomorphism is preserved. The homomorphism for a Lie algebra is,

$$\mathfrak{f}([x,y]) = \mathfrak{f}(x)\mathfrak{f}(y) - \mathfrak{f}(y)\mathfrak{f}(x). \tag{1.38}$$

An especially nice property of a Lie algebra is its *Casimir invariant*, which is an element that commutes with all elements of the Lie algebra. The most commonly used Casimir invariant is the quadratic Casimir, defined as

$$C = \sum_{i,j} \kappa^{ij} X_i X_j, \tag{1.39}$$

for $X \in \mathfrak{g}$. κ^{ij} is the inverse of the Killing form[10]. Casimirs are important for labeling the irreps of a Lie algebra according to their eigenvalues.

Given a Lie group, we would like to be able to find its irreps. Because there is a corresponding Lie algebra for every Lie group, and the Lie algebras are linear, we can study the representations of Lie groups by studying the representations of their Lie algebras. Below, we will introduce representations that are particularly useful to our central question of how to understand particles as irreps of the Poincaré group.

The adjoint representation. The adjoint representation is comprised of the generators themselves, and is way to represent the elements of the group as linear transformations of its Lie algebra. Denoting the matrices of the adjoint representation by T_a, we have,

$$[T_a]_{bc} \equiv -f_{abc}. \tag{1.40}$$

These matrices by definition obey,

$$[T_a, T_b] = i f_{abc} T_c. \tag{1.41}$$

The dimension of the adjoint representation is then given by N, the number of generators of the group (also the number of parameters specifying each group element).

[10]This is a bilinear form $\kappa : \mathfrak{g} \times \mathfrak{g} \to \mathbb{R}$ or $\mathbb{C}$ such that $\kappa(X_i, X_j) = \mathrm{Tr}(T_i T_j)$, where T gives $X \in \mathfrak{g}$ in the adjoint representation, described below.

The fundamental representation. For Lie groups, the fundamental representation is equivalent to the *defining* representation, which is given by the matrices of the Lie group itself. For example, the fundamental representation of $SO(n)$ is given by the $n \times n$ simple orthogonal matrices.

Tensor representations. The fundamental representations of $SO(n)$ associate each element of $SO(n)$ to a rotation matrix that specifies how an n-vector transforms in $\mathbb{R}^n$,

$$v'_a = \sum_{i=1}^{n} R_{ai} v_i. \tag{1.42}$$

This particular representation is sometimes known as the *vector representation* because of this property. We can also consider how higher-dimensional objects transform under the rotation matrices like matrices,

$$M'_{ab} = \sum_{i=1}^{n} R_{ai} R_{bj} M_{ij} \tag{1.43}$$

and tensors,

$$T'_{abc} = \sum_{i=1}^{n} R_{ai} R_{bj} R_{ck} T_{ijk}. \tag{1.44}$$

Letting $n = 3$ for illustrative purposes, we can see that a 3-vector will transform according to 1.42, a 3×3 matrix will transform according to 1.43, and a $3 \times 3 \times 3$ tensor will transform according to 1.44. While the **3** representation[11] tells us how a 3-component vector transforms under $SO(3)$, we can similarly arrange the 9 components of M'_{ab} in a 9-component vector, with each element of $SO(3)$ represented by a 9×9 matrix with components $R_{ai}R_{bj}$. This is the **9** representation. We notice that this representation is built from the tensor product of the **3** representation,

$$\mathbf{9} = \mathbf{3} \otimes \mathbf{3}. \tag{1.45}$$

[11] The notation $\mathbf{n}$ denotes a representation of a Lie group in which each element is associated with an $n \times n$-dimensional matrix. For complex representations, we distinguish between the $\mathbf{n}$ representation and the complex conjugate $\overline{\mathbf{n}}$ representation.

Similarly, we can construct a representation,

$$\mathbf{27} = \mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3}, \tag{1.46}$$

in which we arrange the 27 components of T'_{abc} into a 27-component vector, with each element of $SO(3)$ represented by a 27×27 dimensional matrix with components $R_{ai}R_{bj}R_{ck}$. We can build higher-rank tensor representations in the same way.

Recall that we can always decompose a tensor into its symmetric and anti-symmetric parts. For example,

$$M_{ij} = S_{ij} + A_{ij}, \tag{1.47}$$

with

$$S_{ij} = \frac{1}{2}(M_{ij} + M_{ji}), \quad A_{ij} = \frac{1}{2}(M_{ij} - M_{ji}). \tag{1.48}$$

One can show that S and A are invariant under the action of 1.43. Because S has 6 independent components and A has 3 independent components, we can decompose the **9** representation as,

$$\mathbf{9} = \mathbf{6} \oplus \mathbf{3}. \tag{1.49}$$

The **6** representation can be further decomposed into $\mathbf{6} = \mathbf{5} \oplus \mathbf{1}$, because the trace (which is symmetric) is invariant under 1.43. In full, the decomposition into invariant subspaces is,

$$\mathbf{9} = \mathbf{5} \oplus \mathbf{3} \oplus \mathbf{1}. \tag{1.50}$$

For $SO(n)$, a rank-m tensor transforms in the tensor product representation,

$$\mathbf{n} \otimes \mathbf{n} \otimes ... \otimes \mathbf{n}, \tag{1.51}$$

where we take the direct product $m-1$ times. Studying how these general tensors decompose into irreps leads to the familiar Clebsch-Gordan coefficients.

Let's now consider tensor representations of $SU(n)$. The matrices are now unitary instead of orthogonal, and can be complex-valued. We then construct tensor representations with $\mathbf{n}$ as well as $\overline{\mathbf{n}}$. Objects with upper indices transform under products of $U \in SU(n)$ while objects with lower indices transform under products of $U^\dagger \in SU(n)$. For example, the tensor T^{ij}_k then transforms under products of unitary matrices $U \in SU(n)$ as,

$$(T')^{ij}_k = \sum_{i,j,k=1}^{n} U^i_a U^j_b (U^\dagger)^c_k T^{ab}_c, \tag{1.52}$$

and defines the tensor product representation $\mathbf{n} \otimes \mathbf{n} \otimes \overline{\mathbf{n}}$.

Spin representations. A spin representation is a type of *projective representation* [12]. We've already been introduced to the spin group $Spin(n)$ as the double cover of $SO(n)$, meaning that there is a group homomorphism $Spin(n) \to SO(n)$ with kernel $\mathbb{Z}_2$. The spin *representations* of $SO(n)$ are defined to be the representations of $Spin(n)$ that cannot be mapped onto ordinary representations of $SO(n)$.

We will construct explicit examples of the spin representations for the Lie algebra $\mathfrak{su}(2)$, which will be useful to later discussions. The derivation is familiar in the context of introductory quantum mechanics, but we will be explicit in our treatment of representation theory. Let L_1, L_2, L_3 form a basis for $\mathfrak{su}(2)$. Its bracket is then,

$$[L_i, L_j] = i \sum_{i,j=1}^{3} \epsilon_{ijk} L_k. \tag{1.53}$$

Representations of the algebra — here given by the matrices $\Gamma_1, \Gamma_2, \Gamma_3$ — will also follow

[12] A projective representation is given by matrices $T(g)$, $g \in G$ such that $T(g_1)T(g_2) = c(g_1, g_2)T(g_1, g_2)$, and is defined only for $SO(n)$. In other words, the group structure is maintained only up to a constant $c(g_1, g_2)$.

this structure,

$$[\Gamma_i, \Gamma_j] = i \sum_{i,j=1}^{3} \epsilon_{ijk} \Gamma_k. \tag{1.54}$$

Constructing the representation up to a similarity transformation, we can choose Γ_3 to be diagonal. Its eigenstates can be labeled by integer m,

$$\Gamma_3 |m\rangle = m|m\rangle. \tag{1.55}$$

We then label the *highest weight state* by j, where j is the largest eigenvalue of Γ_3 in a representation. Note that our representation is finite-dimensional. We define raising and lowering operators,

$$\Gamma_\pm = \frac{1}{\sqrt{2}}(\Gamma_1 \pm i\Gamma_2). \tag{1.56}$$

The states are normalized,

$$\Gamma_+ |m\rangle = N_m |m+1\rangle, \tag{1.57}$$

where $N_m = \frac{1}{2}j(j+1) - \frac{1}{2}m(m+1)$. This can be computed by fixing the normalization $\Gamma_- |m\rangle = |m-1\rangle$ and noting that $\Gamma_+ |j\rangle = 0$. We can then recursively construct the $2j+1$ states, $|-j\rangle, ...|0\rangle, ..., |j\rangle$. The $(2j+1)$-dimensional representation is known as the spin-j representation, with states labeled by $|j, m\rangle$. The representation must have positive integer dimension, and so j must be a multiple of $1/2$. These are irreps of $\mathfrak{su}(2)$, and there are an infinite number of them.

We can then find the spin representation explicitly for a given j. The states of the spin-$1/2$ representation are $|1/2, 1/2\rangle$ and $|1/2, -1/2\rangle$. Thinking of the states as basis vectors of a two-dimensional vector space,

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} = |1/2, 1/2\rangle, \quad \begin{pmatrix} 0 \\ 1 \end{pmatrix} = |1/2, -1/2\rangle, \tag{1.58}$$

we can solve for $\Gamma_1, \Gamma_2, \Gamma_3$. We find that they are the Pauli matrices,

$$\Gamma_1 = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \Gamma_2 = -\frac{i}{2}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \Gamma_3 = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{1.59}$$

The quadratic Casimir for $\mathfrak{su}(2)$ is $\Gamma^2 \equiv \Gamma_1^2 + \Gamma_2^2 + \Gamma_3^2$. One can find,

$$\Gamma^2|j,m\rangle = j(j+1)|j,m\rangle, \tag{1.60}$$

which corresponds to the total angular momentum, as seen in introductory quantum mechanics.

Lie groups and Lie algebras are a rich topic, and we could say many more interesting things about them. To make progress with our central question, we will now focus on one particular Lie group, the Poincaré group.

Defining a Particle

The *Poincaré group* is the group of all spacetime symmetries in Minkowski spacetime. It acts on spacetime coordinates x^μ as,

$$x^\mu \to (x')^\mu = L^\mu_\nu x^\nu + a^\mu, \tag{1.61}$$

where we sum over spacetime indices as usual, L^μ_ν are components of a Lorentz transformation, and a^μ denotes a spatial translation. The Lorentz group is the group $SO(3,1)$, where the time and spatial dimensions are denoted separately, and it is the group of rotations in four dimensions with respect to the Minkowski metric. Because it shares many properties with $SO(4)$, the group of rotations in four dimensions with respect to the Euclidean metric, we will start our discussion there.

SO(4). The fundamental (defining) representation of $SO(4)$ is given by the set of special orthogonal 4×4 matrices. Its double cover is $Spin(4)$. We can further show that there is an isomorphism $Spin(4) \simeq SU(2) \times SU(2) \implies SO(4) \simeq SU(2) \times SU(2) \implies \mathfrak{so}(4) \simeq \mathfrak{su}(2) \oplus \mathfrak{su}(2)$. Building on the finite irreps we found for $\mathfrak{su}(2)$, the double cover relationship tells us that the irreducible projective representations of $\mathfrak{so}(4)$ are labeled by two half-integers j_1, j_2 and have dimension $(2j_1+1)(2j_2+1)$. The $\mathfrak{so}(4)$ Lie algebra is then given by two copies of the Lie algebra of $\mathfrak{su}(2)$, which we label by generators $L_\pm$,

$$\begin{aligned}[L_{+i}, L_{+j}] &= i \sum_{i,j=1}^{3} \epsilon_{ijk} L_{+k} \\ [L_{-i}, L_{-j}] &= i \sum_{i,j=1}^{3} \epsilon_{ijk} L_{-k}.\end{aligned} \tag{1.62}$$

The L_+, L_- should also commute,

$$[L_{+i}, L_{-j}] = 0, \tag{1.63}$$

for $i, j = 1, 2, 3$. The $\mathfrak{so}(4)$ Lie algebra is more commonly written as,

$$\begin{aligned}[J_i, J_j] &= i \sum_{i,j=1}^{3} \epsilon_{ijk} J_k \\ [J_i, K_j] &= i \sum_{i,j=1}^{3} \epsilon_{ijk} K_k \\ [K_i, K_j] &= i \sum_{i,j=1}^{3} \epsilon_{ijk} K_k,\end{aligned} \tag{1.64}$$

which is the same as that given by $L_\pm$ with the basis change $L_{\pm i} = \frac{1}{2}(J_i + K_i)$.

The Lorentz group. The Lie algebra of the Lorentz group is given by,

$$\begin{aligned}[J_i, J_j] &= i\sum_{i,j=1}^{3}\epsilon_{ijk}J_k \\ [J_i, K_j] &= i\sum_{i,j=1}^{3}\epsilon_{ijk}K_k \\ [K_i, K_j] &= -i\sum_{i,j=1}^{3}\epsilon_{ijk}K_k,\end{aligned} \tag{1.65}$$

which is the same as the Lie algebra for $\mathfrak{so}(4)$, except we've picked up a minus sign on the third commutator, i.e. $K_i \to iK_i$. Going through the same arguments as in the case of $\mathfrak{so}(4)$, we find that we can study the Lorentz group as the Lie algebra isomorphism,

$$\mathfrak{so}(3,1)_{\mathbb{C}} \simeq \mathfrak{su}(2)_{\mathbb{C}} \oplus \mathfrak{su}(2)_{\mathbb{C}}, \tag{1.66}$$

where the subscript denotes the basis change in the third commutator[13]. We can identify the generators J as generating rotations. Satisfying the commutators leads to the boost matrices for K, and we identify these as the generators of the boosts. We see that the Lie algebra corresponding to the Lorentz group is 6-dimensional.

The Poincaré group. Finally, the Lie algebra of the Poincaré group is given by,

$$\begin{aligned}[P_\mu, P_\nu] &= 0 \\ [M_{\mu,nu}, P_\rho] &= -i(\eta_{\mu\rho}P_\nu - \eta_{\nu\rho}P_\mu) \\ [M_{\mu\nu}, M_{\rho,\sigma}] &= -i(\eta_{\mu\rho}M_{\nu\sigma} - \eta_{\mu\sigma}M_{\nu\rho} - \eta_{\nu\rho}M_{\mu\sigma} + \eta_{\nu\sigma}M_{\mu\rho}),\end{aligned} \tag{1.67}$$

where the rotations are given by $J_i = \frac{1}{2}\epsilon_{imn}M^{mn}$ and the boosts are given by $K_i = M_{i0}$, the P_μ are the generators of the translations, and $\eta_{\mu\nu}$ is the Minkowski metric.

There are two Casimir invariants for the Poincaré group. The first is $P^2 = P_\mu P^\mu$, and the second is $W^2 = W_\mu W^\mu$, with $W_\mu \equiv -\frac{1}{2}\epsilon_{\mu\nu\rho\sigma}M^{\nu\rho}P^\sigma$, which is known as the

[13]This is known as *complexification*. We can get back to real Lie algebras by noting that there is an isomorphism between $\mathfrak{su}(2)_{\mathbb{C}}$ and $\mathfrak{sl}_2(\mathbb{C})$, where the latter is the special linear group. One eventually finds that $\mathfrak{so}(3,1)_{\mathbb{C}} \simeq \mathfrak{sl}_2(\mathbb{C})$, and from there we can restrict to the reals.

Pauli-Lubanski pseudovector. Because the eigenvalue of P^2 on state $|p\rangle$ is $-m^2$, the representations of the Poincaré group are labeled by m^2. We can then study the irreps of the Poincaré group based on the mass and momentum of eigenvector $|p\rangle$:

1. Zero mass and zero momentum. The only finite-dimensional unitary representation in this case is the trivial representation. This is the *vacuum*, and it is invariant under all symmetries of the Poincaré group.

2. $m > 0$. For positive mass in the rest frame, $p_\mu = (m, 0, 0, 0)$. We then find the *little group*: The subgroup of the Poincaré group that leaves p_μ invariant. We find that it is $SO(3)$, whose Lie algebra is isomorphic to $\mathfrak{su}(2)$. $SO(3)$ then has an infinite number of irreps, labeled by half-integer j with dimensions $2j+1$. We've seen that j corresponds to the spin angular momentum.

3. Zero mass and nonzero momentum. Choosing $p_\mu = (p, 0, 0, p)$, the little group is the *special Euclidean group* $SE(2)$. It has one generator of rotations J, and two generators of translations P_1, P_2. The eigenvector under translations is $P_i|k\rangle = k|k\rangle$ and, for $k = 0$, the little group is $SO(2)$. Its irreps are one-dimensional and indexed by *helicity* h. Because there are spin representations, h should also be a half-integer. For $k \neq 0$, the little group is trivial.

We are now ready to state *Wigner's theorem*, which classifies unitary representations of the Poincaré group. Following the above analysis, the irreps of the Poincaré group are indexed by the mass m of the particle, and either its spin j (if it is massive) or its helicity h (if it is massless). A *particle* is then an irreducible representation of the Poincaré group, and it can be either massive or massless. Massive particles are indexed by their spin, while massless particles are indexed by their helicity.

We can go a step further and begin to classify these particles. Consider again irreps for which $m > 0$, whose little group is $SO(3)$. Further, recall that spin representations are

defined for $SO(n)$ and are projective representations, which preserve the group structure up to a constant. *Bargmann's theorem* states that every projective unitary representation[14] of a group G can be lifted[15] to an ordinary representation of the universal cover of G. The universal cover of $SO(3)$ is $SU(2)$, and so all projective unitary representations of $SO(3)$ come from $SU(2)$. We can then find the projective irreps of $SO(3)$ by exponentiating the irreps of $\mathfrak{su}(2)$. And because $Spin(3)$ is the double cover of $SO(3)$, these projective representations will be spin representations.

We found that the irreps of $\mathfrak{su}(2)$ can be indexed by a half integer j and are $(2j+1)$-dimensional. Let's consider both $j = 1/2$ and $j = 1$. For $j = 1/2$, we found that the generators of the Lie algebra were the Pauli matrices. We can then choose elements of the Lie algebra $\theta_1\Gamma_3, \theta_2\Gamma_3$ where $\theta_1, \theta_2 \in \mathbb{R}$ to find two elements of $SO(3)$ via exponentiation,

$$R(\theta_1) = e^{i\theta_1\Gamma_3} = \begin{pmatrix} e^{i\theta_1/2} & 0 \\ 0 & e^{-i\theta_1/2} \end{pmatrix}, \quad R(\theta_2) = e^{i\theta_1\Gamma_3} = \begin{pmatrix} e^{i\theta_2/2} & 0 \\ 0 & e^{-i\theta_2/2} \end{pmatrix}. \tag{1.68}$$

Multiplying these elements together should also yield an element of the group,

$$R(\theta_1)R(\theta_2) = \begin{pmatrix} e^{i(\theta_1+\theta_2)/2} & 0 \\ 0 & e^{-i(\theta_1+\theta_2)/2} \end{pmatrix}. \tag{1.69}$$

For $\theta_1 + \theta_2 = 2\pi$, this becomes,

$$R(\theta_1)R(\theta_2) = -\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{1.70}$$

In rotating by 2π, we have picked up a minus sign, rather than returning to the identity. This is the definition of a projective representation; the group structure has been preserved up to a constant. This representation is therefore a spin representation.

[14] Technically, this theorem only applies to groups whose Lie algebra cohomology satisfy certain properties; we will not detail this further except to say that all groups we will consider satisfy such properties.

[15] A projective representation is *lifted* if it is possible to choose a projective representation $T(g) \in \tilde{T}(g)$, where $\tilde{T}(g)$ is a representation of the projective linear group $PGL(V)$, such that the group structure is preserved exactly.

For $j = 1$, we can again solve for $\Gamma_1, \Gamma_2, \Gamma_3$ to find,

$$\Gamma_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \Gamma_2 = -\frac{i}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}, \quad \Gamma_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}. \tag{1.71}$$

We can do the same exercise to find,

$$R(\theta_1)R(\theta_2) = e^{i\theta_1\Gamma_3}e^{i\theta_2\Gamma_3} = \begin{pmatrix} e^{i(\theta_1+\theta_2)} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & e^{-i(\theta_1+\theta_2)} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ for } \theta_1 + \theta_2 = 2\pi. \tag{1.72}$$

Rotating by 2π has yielded the identity matrix. We can repeat this process for any half integer j, and the punchline is: For integer $j = 1, 2, ...$, the elements of $SO(3)$ transform as ordinary representations, while for half integer $j = 1/2, 3/2, ...$, the elements of $SO(3)$ transform as spin representations. Objects that transform under the ordinary representations are vectors, while the objects that transform under the spin representations are *spinors*. This is a nice way to see that bosonic particles, with integer spin, transform as vectors while fermionic particles, with half integer spin, transform as spinors.

Spinors are equivalently, and more commonly, defined as elements of a vector space linearly represented by the *Clifford algebra*. This is a Lie algebra with a bracket

$$\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu} \tag{1.73}$$

where γ^μ are the *gamma matrices*, defined as,

$$\gamma^0 = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}, \quad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}, \tag{1.74}$$

where I is the 2×2 identity matrix, and σ^i are the Pauli matrices. The γ^5 defined by $\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3$. A spinor field is "left-handed chiral" or "right-handed chiral" depending

on whether it transforms as,

$$\psi_L = P_L\psi \;\text{ or }\; \psi_R = P_R\psi, \tag{1.75}$$

respectively. P_L and P_R are the projection matrices,

$$P_L = \frac{1-\gamma^5}{2}, \quad P_L = \frac{1+\gamma^5}{2}. \tag{1.76}$$

As defined, ψ_L and ψ_R are *Weyl spinors*, while $\psi = \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix}$ is the *Dirac spinor*.

Particles and fields. In quantum mechanics, our system is built from states occupying some Hilbert space and operators that act upon the states. A transformation on the system can be captured by a unitary operator U acting on the state,

$$|\psi\rangle \to |\psi'\rangle = U|\psi\rangle. \tag{1.77}$$

Ensuring that this operator is unitary means that the inner product of states is unchanged, and so probabilities are conserved. The system should also evolve in time as,

$$|\psi'(t)\rangle = e^{-iH(t-t_0)}U|\psi'(t_0)\rangle, \tag{1.78}$$

which implies that U commutes the Hamiltonian H. Each eigenvalue of the Hamiltonian can be associated with an irrep of its symmetry group, and the eigenfunctions give the basis for the space of states. This state space — the Hilbert space — decomposes into a direct sum of subspaces, each of which transform irreps of the symmetry group.

In QFT, fields are operator-valued and act upon an infinite-dimensional Hilbert space. We can then think of our symmetries as acting on the operators,

$$\mathcal{O} \to \mathcal{O}' = U^\dagger(g)\mathcal{O}U(g), \tag{1.79}$$

where U is again unitary. The symmetry then acts linearly on the fields,

$$\phi_i(x) \to \phi_i'(x) = U^\dagger(g)\phi_i(x)U(g) = M_i^j(g)\phi_j(x), \tag{1.80}$$

where $M_i^j(g)$ is a representation of the symmetry group.

There are many layers of mathematical structure we've built up to reach this point. First, we've developed the notion of a symmetry group, which is an abstract object. Then, we built up representations of the group on a vector space V over a field F. Finally, the group can be represented by matrices which act on the fields. To finish out the connection, we can think of the irreps of the Poincaré group — the particles — as forming a basis for the Hilbert space of the QFT, and the fields are operator-valued and act on these states.

1.1.2 Symmetries and their Breaking

As we've now seen, symmetry plays in indispensable role in QFT. Thus far we've studied Poincaré invariance, a *spacetime symmetry* and example of a *local* symmetry due to its dependence on a spacetime coordinate. Local symmetries are important due to their connection to charges and conservation laws; they are in contrast to *global* symmetries which are the same at every point in spacetime. In addition to spacetime symmetries, particles may obey *internal symmetries*, which act on the internal space generated by the fields. The *Coleman-Mandula theorem* tells us that, in spacetime dimensions greater than $1+1$, the symmetry group of an interacting QFT will factorize as the Poincaré symmetry $\times$ internal symmetries [16].

A QFT may be invariant under any number of symmetries, whether *accidental* or *imposed*, and they lead to a rich variety of phenomena. In this section, we will review aspects of both continuous and discrete symmetries, for both global and local transformations.

[16] There are two important exceptions to this theorem, conformal invariance and supersymmetry, which we will not discuss here.

Continuous Symmetries

Emmy Noether, in her namesake theorem, proved the correspondence between continuous symmetries of the action and conserved currents. To illustrate, let's consider a Lagrangian density $\mathcal{L}$ that is invariant under a translation symmetry $x^\mu \to x'^\mu = x^\mu + a^\mu$. The variation of the action is given, in general, by

$$\delta S = \int \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \delta(\partial_\mu \phi) + \frac{\partial \mathcal{L}}{\partial \phi} \delta\phi \right) d^4x. \tag{1.81}$$

Integrating by parts and applying the Euler-Lagrange equations yields,

$$\delta S = \int \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \delta\phi \right) d^4x. \tag{1.82}$$

We can Taylor expand to find the change in the field under the infinitesimal translation: $\delta\phi = -a^\nu \partial_\nu \phi$. The variation of the action is then,

$$\delta S = -a^\nu \int \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \partial_\nu \phi \right) d^4x. \tag{1.83}$$

Which can be rewritten in terms of the energy-momentum tensor,

$$\delta S = a^\nu \int \partial_\mu T^{\mu\nu} \, d^4x, \tag{1.84}$$

where

$$T^{\mu\nu} = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \partial_\nu \phi - \eta^{\mu\nu} \mathcal{L}. \tag{1.85}$$

$\eta^{\mu\nu}$ is the usual Minkowski metric. For the action to be invariant under the translation, $\delta S = 0$. This implies,

$$\partial_\mu T^{\mu\nu} = 0. \tag{1.86}$$

This implies that, if the action is invariant under translations, both energy T^{00} and momentum T^{0i} are conserved.

Global Symmetries

Noether's theorem motivates the concept of charge. If we consider two complex scalars, ϕ_1 and ϕ_2, we can impose a symmetry under a phase rotation,

$$\phi_1 \to e^{i\alpha_1\theta}\phi_1, \quad \phi_2 \to e^{i\alpha_2\theta}\phi_2. \tag{1.87}$$

The Lagrangian for a model with these fields can be built by requiring each term to be invariant under both $U(1)$ symmetries,

$$\mathcal{L} = \partial^\mu\phi_1^\dagger\partial_\mu\phi_1 + \partial^\mu\phi_2^\dagger\partial_\mu\phi_2 - V(\phi_1,\phi_2), \tag{1.88}$$

where $V(\phi_1,\phi_2)$ is a general, q_1- and q_2-conserving potential. The *Noether currents* are,

$$j_1^\mu = i(\phi_1^\dagger\partial^\mu\phi_1 - \phi_1\partial^\mu\phi_1^\dagger), \quad j_2^\mu = i(\phi_2^\dagger\partial^\mu\phi_2 - \phi_2\partial^\mu\phi_2^\dagger) \tag{1.89}$$

By Noether's theorem, these are conserved quantities. We can integrate the first component of the current, corresponding to the charge density, to find the *Noether charge*,

$$Q_1 = \int d^3x\, j_1^0. \tag{1.90}$$

While Q_1 gives the total amount of charge for the system, q_1 corresponds to the charge quantum number, and is what we typically think of as the "charge" of a particle. Because we've discussed only global continuous symmetries, with no spacetime dependence, this scenario corresponds to classical electrodynamics. We can further identify q_1 and q_2 as the generators of the corresponding $U(1)$ symmetries.

In the case of fermions, consider two spinor fields, ψ_1 and ψ_2. A $U(1)$ phase rotation corresponds to,

$$\psi_L \to e^{iq_L\theta}\psi_L, \quad \psi_R \to e^{iq_R\theta}\psi_R. \tag{1.91}$$

If $q_L \neq q_R$, this symmetry is known as *chiral*. If not, it is a *vectorial* symmetry.

Local Symmetries

Considering now just one scalar field ϕ, we can see what happens when make θ a continuous parameter. The transformation is,

$$\phi(x) \to e^{iq\theta(x)}\phi(x), \tag{1.92}$$

with the complex conjugate transforming as usual,

$$\phi(x)^\dagger \to e^{-iq\theta(x)}\phi(x)^\dagger. \tag{1.93}$$

The kinetic term transforms under this symmetry as,

$$\partial_\mu\phi^\dagger(x)\partial^\mu\phi(x) \to \left[\partial_\mu\phi^\dagger(x) - iq\left(\partial_\mu\theta(x)\right)\phi^\dagger(x)\right]\left[\partial^\mu\phi(x) + iq\left(\partial^\mu\theta(x)\right)\phi(x)\right] \neq \partial_\mu\phi^\dagger(x)\partial^\mu\phi(x). \tag{1.94}$$

The kinetic term then apparently violates this symmetry, and we find the same in the case of fermion fields. To make the Lagrangian invariant under this $U(1)$ symmetry, we have to "subtract" the extra terms that arise. To this end, we introduce the *covariant derivative*,

$$D^\mu = \partial^\mu + igqA^\mu, \tag{1.95}$$

where g is the gauge coupling and A^μ is a vector field that transforms as,

$$A_\mu \to A_\mu - \frac{1}{g}\partial_\mu\theta. \tag{1.96}$$

With this new definition, the kinetic term remains invariant. Because we are adding new vector fields, we introduce a kinetic term for them, which is the familiar *field strength*,

$$\mathcal{L} \supset -\frac{1}{4}F^{\mu\nu}F_{\mu\nu}, \tag{1.97}$$

where the field strength is defined by,

$$[D^\mu, D^\nu] = igqF^{\mu\nu}, \tag{1.98}$$

and can be written in terms of the vector fields,

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu. \tag{1.99}$$

These vector fields are known as *gauge fields*, and a Lagrangian that does not change under such transformations is *gauge invariant*. Interestingly, the gauge fields themselves must be massless because an explicit mass term $\frac{1}{2}m^2 A^\mu A_\mu$ is not gauge invariant. The strength of the interaction of the gauge field is set by gq, and so depends directly on the charge. Finally, the transformation law for A^μ shows that the gauge field transforms in the adjoint representation.

Gauge invariance is not a symmetry in the same sense that spacetime symmetries and global symmetries are symmetries. It is rather a "symmetry of description", because a gauge-invariant Lagrangian may be rewritten in several different ways that all lead to the same physics.

Non-Abelian Symmetries

Consider a field ϕ in a representation R with N components. The transformation law for a general local symmetry is,

$$\phi_i \to \left(e^{iT_a(x)\theta_a(x)}\right)_{ij} \phi_j, \tag{1.100}$$

where the T_a are the generators of the symmetry. For abelian symmetries, the generator is simply the charge q_i. In the case of non-abelian symmetries, we simply need to know the generators. Let's consider the case of ϕ transforming under $SU(n)$. The T_a are then the $SU(n)$ generators with the algebra,

$$[T_a, T_b] = i f_{abc} T_c, \tag{1.101}$$

while ϕ is represented by $N \times N$ matrices. Following the same procedure as in the abelian case, we find that we must introduce a covariant derivative in order to ensure the

invariance of the kinetic term,

$$D^\mu = \partial^\mu + igT_a G_a^\mu, \tag{1.102}$$

where g is again a coupling constant and we introduce a vector field G_a^μ that transforms as,

$$G_\mu^a \to G_\mu^a - f^{abc}\theta^b G_\mu^c - \frac{1}{g}\partial_\mu \theta^a. \tag{1.103}$$

The field strength is now defined by the commutator,

$$[D_\mu, D_\nu] = igT^a G_{\mu\nu}^a \tag{1.104}$$

and can be written in terms of G_a^μ as,

$$G_{\mu\nu}^a = \partial_\mu G_\nu^a - \partial_\nu G_\mu^a - g f^{abc} G_\mu^b G_\nu^c. \tag{1.105}$$

Again, we find that we cannot write down explicit mass terms for the gauge fields that are gauge invariant.

There are a rich variety of phenomena that arise from gauge theories, which we will not attempt to detail here, but can be found in excellent resources such as [3].

Discrete Symmetries

In QFT, the set of C, P, and T discrete symmetries — charge conjugation, parity, and time reversal, respectively — are especially important. All of these are spacetime symmetries and their combination, CPT, must be conserved for any local theory that obeys Lorentz invariance. Indeed, no experimental violations of CPT have ever been observed.

C and P are especially interesting in theories involving fermions. While C symmetry exchanges the sign of charges, P exchanges the chirality of space, i.e. $\mathbf{x} \to -\mathbf{x}$. This has

the effect of changing the chirality of fermions; under P, $\psi_L \leftrightarrow \psi_R$. The combined CP symmetry performs both operations at once, and so makes the exchanges,

$$\phi \to \phi^\dagger, \quad \psi_{Li} \to \psi_{Li}, \quad \psi_{Ri} \to \psi_{Ri}. \tag{1.106}$$

If there is a basis of the Lagrangian in which all parameters are real, CP is guaranteed to be conserved.

Broken Symmetries

If the Lagrangian contains terms do not obey a symmetry, but whose coefficients are small parameters, this symmetry is said to be *explicitly broken*. A symmetry may also be *spontaneously broken* if the Lagrangian of the theory is invariant under the symmetry but its vacuum state is not. To illustrate this, consider the theory of two scalar fields ϕ_R and ϕ_I related by $\phi \equiv \phi_R + i\phi_I$. The Lagrangian for this theory, after imposing an $SO(2)$ rotational symmetry, is

$$\mathcal{L} = \frac{1}{2}\partial_\mu\phi_R\partial^\mu\phi_R + \frac{1}{2}\partial_\mu\phi_I\partial^\mu\phi_I - \frac{\mu^2}{2}\left(\phi_R^2 + \phi_I^2\right) - \frac{\lambda}{4}\left(\phi_R^2 + \phi_I^2\right)^2, \tag{1.107}$$

where μ^2 and λ are real parameters. Assuming $\mu^2 < 0$, and defining $v^2 = -\mu^2/\lambda$, we can rewrite the potential as,

$$V = \lambda\left(\phi^\dagger\phi - \frac{v^2}{2}\right)^2. \tag{1.108}$$

The *vacuum expectation value* (VEV) $2\langle\phi^\dagger\phi\rangle = \langle\phi_R^2 + \phi_I^2\rangle$ is then equal to $-\mu^2/2\lambda = v^2$. This means that the minimum of the potential V lies on a circle of radius v. We can choose the configuration $\langle\phi_R\rangle = v$, $\langle\phi_I\rangle = 0$ and redefine our fields such that the VEV vanishes. This occurs for $h = \phi_R - v$, $\xi = \phi_I$. The Lagrangian can be rewritten in terms of these fields as,

$$\mathcal{L} = \frac{1}{2}(\partial_\mu h)(\partial^\mu h) + \frac{1}{2}(\partial_\mu\xi)(\partial^\mu\xi) - \lambda v^2h^2 - \lambda vh(h^2 + \xi^2) - \frac{\lambda}{4}(h^2 + \xi^2)^2. \tag{1.109}$$

This Lagrangian is no longer invariant under $SO(2)$, and we can see that we went from a Lagrangian with a nonzero VEV to one with a zero VEV. In other words, the theory has been spontaneously broken.

We can notice a few properties. First, our rewritten Lagrangian contains a scalar field h whose mass is $m^2 = \lambda v^2$. Second, the original Lagrangian was invariant under $SO(2)$, which has one generator, and this is no longer the case in the resulting Lagrangian. The generator is said to be broken; in exchange, we've picked up a massless scalar ξ. *Goldstone's theorem* states as much: For a continuous spontaneously broken symmetry, the number of broken generators correspond to the number of new, massless scalars. These scalars are the *Goldstone bosons*.

1.1.3 The SM Gauge Group

In addition to invariance under the Poincaré group, the particles of the SM are symmetric under a set of internal symmetries,

$$SU(3) \times SU(2)_L \times U(1)_Y, \tag{1.110}$$

which make up the SM *gauge group*. Each representation is denoted $(p, q)_Y$, where p is the representation of $SU(3)$, q the representation of $SU(2)_L$, and Y the hypercharge. The subscript L denotes that the particles that transform under $SU(2)$ all have left-handed chirality, while the hypercharge is the $U(1)$ component of the electroweak symmetry. The

SM representations are,

Particle	**Name**	**Representation**
Spin 1		
B	Z boson	$(1,1)_0$
W	W boson	$(1,3)_0$
G	gluon	$(8,1)_0$
Spin $\frac{1}{2}$		
q_L	left-handed quark	$(3,2)_{\frac{1}{3}}$
u_L^c	left-handed antiquark (up)	$(\bar{3},1)_{-\frac{4}{3}}$
d_L^c	left-handed antiquark (down)	$(\bar{3},1)_{\frac{2}{3}}$
ℓ_L	left-handed lepton	$(1,2)_1$
ℓ_L^c	left-handed antilepton	$(1,1)_2$
Spin 0		
H	Higgs boson	$(1,2)_1$

(1.111)

The SM is further endowed with the electroweak symmetry breaking structure $SU(2)_L \times U(1)_Y \to U(1)_{EM}$, where EM denotes the electromagnetic gauge group. All of the gauge bosons, as illustrated in the previous section, transform under the adjoint representation, while all fermions transform as spinors.

1.2 Effective Field Theory

Effective field theories (EFTs) are surprisingly simple in their core idea: We do not need to know every single degree of freedom in the universe in order to understand a given system. These degrees of freedom may be separated by orders of magnitude in scale, and decouple from one another in certain regimes. Practically, our observational instruments can only make measurements to a finite degree of precision, and so we need

only consider the effects that contribute in a non-negligible way to the observable at hand. This is the reason why are are able to focus on Newtonian mechanics to understand the physics of billiard balls, rather than starting with general relativity. While the latter approach would return the same answer as the former, it is far more laborious and rather unnecessary. The easier approach is to construct a theory that contains the pertinent degrees of freedom that make up a system at a given energy scale.

This is rather convenient for us, as there are a number of phenomena that (as far as we can tell) do not require an extremely coarse-grained understanding of the world in order to make meaningful predictions. Across scientific disciplines, "effective theories" are built with the goal of capturing the essential components of a particular system. In quantum field theory, the formalized organization of an EFT allows us to parameterize our ignorance of physics in the far UV in order to understand the workings of a theory in the IR.

In this section, we will review the details of such parameterizations, focusing on a Lagrangian-based approach to EFT. We begin our discussion with a review of some baseline concepts in constructing QFTs. From there, we illustrate some scale-dependent properties of QFTs, discuss the operator redundancies that are rife in Lagrangian-based descriptions of QFTs, and end by introducing the procedure for building an EFT from the bottom-up.

1.2.1 Lagrangian Formalism

A QFT can be described via an action S. We are familiar with the concept of the action from classical Lagrangian-based mechanics, which is built upon the principle of least action: Given a classical system and its generalized coordinates, we can compute the system's trajectory between two points as the one that minimizes the energy required to

take this trajectory. In quantum mechanics, we no longer have a single, classical trajectory, and instead must sum over all possible paths in order to compute the propagator. For a quantum field theory, this becomes a sum over all possible field configurations. Feynman famously showed that the propagator can be written as,

$$\langle \phi(x_0, x) | \phi(x_0', x') \rangle \sim \int \mathcal{D}\phi e^{i\mathcal{S}}, \tag{1.112}$$

where $\mathcal{D}\phi$ denotes the sum over field configurations and we omit normalization factors. For simplicity, we have written the propagator of a scalar field, though we can of course consider more complicated objects.

The action itself is often written in terms of its Lagrangian density, $S = \int (d^D x)\mathcal{L}$, with D the number of spacetime dimensions. The Lagrangian density $\mathcal{L}$ is most often referred to simply as the Lagrangian. This is a functional that depends on the fields of the theory, and encodes the symmetries it obeys and its interactions. Because the propagator is related to a sum of the action over all possible field configurations, the Lagrangian must include all possible interactions of the fields in a given theory. It is constructed such that several physical principles are maintained, including:

1) Locality. This is the idea that an object can only influence objects in its immediate surroundings. The speed of propagation is bounded by the speed of light.

Mathematically, we require that the Lagrangian is a function only of fields and their derivatives, each of which depend on a single point in spacetime. All spacetime dependence of the Lagrangian is contained in the fields.

2) Causality. Objects that are not contained in each other's past light cone cannot influence one another. This means that fields at two points in spacetime that are

spacelike-separated should commute,

$$[\phi(x), \phi(y)] = 0. \tag{1.113}$$

The principles of locality and causality are addressed by ensuring that the Lagrangian obeys Poincaré symmetry. This is the full symmetry of special relativity, which includes invariance under spatial translations and rotations as well as Lorentz boosts. Considering a scalar field $\phi(x)$, a Lorentz transformation transforms this field as

$$\phi_a(x) \to \phi_a(x')' = [M(\Lambda)]_a^b \phi_b(\Lambda^{-1}\phi'), \tag{1.114}$$

where Λ defines the Lorentz symmetry. The matrix $M(\Lambda)$ forms a representation for ϕ.

More generally, the fields which construct the Lagrangian are irreducible representations of the Poincaré group.

3) Unitarity. This is the principle that ensures that the sum of probabilities in scattering processes between sets of particles remains equal to 1 as the system evolves in time. Consider a general multi-particle state denoted by Ψ_α, where α is a general index describing the state.

For a scattering process, we generically have some "in" states Ψ_α^{in} that interact and result in some "out" states Ψ_α^{out}. This is formalized by the $\mathcal{S}$-matrix, which is defined as

$$\mathcal{S} = \langle \Psi_\alpha^{out} | \Psi_\alpha^{in} \rangle, \tag{1.115}$$

where we use the notation $\mathcal{S}$ to be distinct from the action S. In order for the probabilities to sum to 1, we must have,

$$\mathcal{S}^\dagger \mathcal{S} = 1, \tag{1.116}$$

the requirement of unitarity of the $\mathcal{S}$-matrix. We can further split up the $\mathcal{S}$-matrix into a noninteracting, "free field" component and an interacting component,

$$\mathcal{S} = 1 + i\mathcal{T}, \tag{1.117}$$

where we must also have

$$\mathcal{T}^\dagger \mathcal{T}. \tag{1.118}$$

Given the input of the local fields and the symmetries they obey, the above properties get us most of the way to a sensible Lagrangian. For our theory to be dynamical, it must contain a kinetic term. Massive theories may contain an explicit mass term quadratic in the fields. As a general principle, all interactions that are not explicitly forbidden by a symmetry are included. From a model-building perspective, one may wish to add additional ingredients to the Lagrangian, such as an imposed symmetry.

In an ideal world, field theorists would simply be able to write down an action S for a theory of their choice and evaluate 1.112. In practice, the path integral cannot be computed exactly except in the case of very simple theories. There are then several steps to get from the Lagrangian of a theory to answers for observable quantities, and there are several challenges inherent in this formalism. Some of them are:

1) Perturbative expansion in the computation of the path integral. It is not possible to compute the path integral exactly for a general Lagrangian,

$$\mathcal{L} = \frac{1}{2}\partial_\mu \phi \partial^\mu \phi - \frac{1}{2}m^2\phi^2 - V(\phi), \tag{1.119}$$

where we consider a scalar field ϕ. One can, however, exactly compute the path integral for the free field action, with $V(\phi) = 0$. Carrying out this computation returns the well-known free field propagator. Let's say we want to add an interaction term to this theory,

$$\mathcal{L} = \frac{1}{2}\partial_\mu \phi \partial^\mu \phi - \frac{1}{2}m^2\phi^2 - \frac{\lambda}{4!}\phi^4. \tag{1.120}$$

The path integral we wish to compute is then,

$$\int \mathcal{D}\phi e^{i\int d^4x\left(\frac{1}{2}\partial_\mu\phi\partial^\mu\phi-\frac{1}{2}m^2\phi^2-\frac{\lambda}{4!}\phi^4+J\phi\right)}, \tag{1.121}$$

where $J\phi$ is a source term. To proceed, can separate the free field and interacting parts of the integral and Taylor-expand the interacting part order by order in λ. This yields,

$$\int \mathcal{D}\phi\left[e^{i\int d^4x\left(\frac{1}{2}\partial_\mu\phi\partial^\mu\phi-\frac{1}{2}m^2\phi^2+J\phi\right)}\left(\int d^4x\left[1-\frac{i}{4!}\lambda\phi^4+\frac{1}{2}\left(\frac{i}{4!}\right)^2\lambda^2\phi^8+...\right]\right)\right]. \tag{1.122}$$

which can be rewritten as,

$$\int d^4w\left(1-\frac{i}{4!}\lambda\left(\frac{\delta}{\delta J(w)}\right)^4+\frac{1}{2}\left(\frac{i}{4!}\right)^2\lambda^2\left(\frac{\delta}{\delta J(w)}\right)^8+...\right)\int \mathcal{D}\phi e^{i\int d^4x\left(\frac{1}{2}\partial_\mu\phi\partial^\mu\phi-\frac{1}{2}m^2\phi^2+J\phi\right)} \tag{1.123}$$

$$=\sqrt{\frac{2\pi}{m^2}}\int d^4w\left(1-\frac{i}{4!}\lambda\left(\frac{\delta}{\delta J(w)}\right)^4+\frac{1}{2}\left(\frac{i}{4!}\right)^2\lambda^2\left(\frac{\delta}{\delta J(w)}\right)^8+...\right)e^{-i\int\int d^4xd^4yJ(x)D(x-y)J(y)}, \tag{1.124}$$

where $D(x-y)$ is the free field propagator. For details, see the excellent discussion in chapter 1 of [1]. We can further expand the exponential on the right-hand side to obtain a total expansion in terms of both λ and J. This expansion maps on to the familiar Feynman diagrams we've come to associate with perturbative QFT, with the order of J corresponding to the number of particles involves in each diagram and λ corresponding to the number of times they intersect.

In order to obtain this result, we have considered small λ: the weak-coupling regime. This ensures that we can safely truncate our expansion of the interaction term, as terms with higher powers of λ will negligibly contribute to our overall result for the path integral. In a strong-coupling regime, we cannot safely make this assumption and must consider other methods.

There are several nuances of this discussion that are worth emphasizing. In particular, this perturbative approximation is not valid for every QFT. In practice, there are only a small class of QFTs for which we can utilize this machinery and be ensured that we are sufficiently capturing the phenomena of the theory. The perturbative approximation of QCD, for example, famously does not capture important non-perturbative effects like instantons. QFTs themselves are not mathematically well-defined outside of the perturbative regime, and it remains unknown whether or not every QFT even admits such a Lagrangian description. Finally, although we've considered a path integral approach here, one can similarly consider these ideas from the approach of canonical quantization.

2) Divergences in computations of loop diagrams. The perturbative expansion of the path integral is one way to see how Feynman diagrams simplify our calculation. Each term can be associated with a Feynman diagram, and in turning the crank to compute the contribution of each term to the path integral, one can notice patterns that comprise the Feynman rules. Knowing the Feynman rules for a particular theory allows one to draw each diagram and compute its individual amplitude and sum these amplitudes to arrive at the total.

Feynman diagrams for tree-level processes are relatively straightforward. Complications arise when we consider loops, which require performing an integral over an internal loop of momenta. When considering the scattering process $\phi\phi \to \phi\phi$, we have the loop diagram:

To evaluate it, we must compute the integral,

$$\frac{1}{2}(-i\lambda)^2 \int \frac{d^4k}{(2\pi)^4} \left(\frac{i}{k^2 - m^2 + i\epsilon}\right) \left(\frac{i}{(k_1 + k_2 - k)^2 + m^2 + i\epsilon}\right). \tag{1.125}$$

When k is large, the integral becomes,

$$-\frac{1}{2}(-i\lambda)^2 \int \frac{d^4k}{(2\pi)^4} \left(\frac{1}{k^4}\right) = \frac{1}{24(2\pi)^4}(-i\lambda)^2 \log(k), \tag{1.126}$$

which is divergent. The resolution to this apparent problem lies in recognizing that our integral involved summing over modes of arbitrarily high momentum. Because we have only verified our knowledge of physics up to some energy scale, and indeed we know that there is some UV physics that we do not know, it does not necessarily make sense to include these high energy modes in the sum. We therefore introduce a physical parameter Λ that tells us up to what momenta we can reasonably compute our integral. This *cutoff* scale *regularizes* the integral, which is now formally,

$$\frac{1}{2}(-i\lambda)^2 \int_0^{\Lambda} \frac{d^4k}{(2\pi)^4} \left(\frac{i}{k^2 - m^2 + i\epsilon}\right) \left(\frac{i}{(k_1 + k_2 - k)^2 + m^2 + i\epsilon}\right), \tag{1.127}$$

and the amplitude itself now depends on the cutoff. We can evaluate this integral using Feynman's trick, $\int_0^1 dx(xA + (1-x)B)^{-2} = (AB)^{-1}$. The computation can be found in standard QFT textbooks [4] [5]. At one-loop, after summing all one-loop diagrams, the result is,

$$\mathcal{M}(\phi\phi \to \phi\phi) = -i\lambda + iC\lambda^2 \left[\log\left(\frac{\Lambda^2}{s}\right) + \log\left(\frac{\Lambda^2}{t}\right) + \log\left(\frac{\Lambda^2}{u}\right)\right] + \mathcal{O}(\lambda^3), \tag{1.128}$$

where s, t, and u are the Mandelstam variables and C is a constant. We have successfully removed the momentum divergence, as the amplitude now scales logarithmically with the cutoff. However, the scattering-amplitude is a measurable quantity, and should not depend on the specific choice of Λ. Further, an experiment can only measure λ at a given value of s, t, and u, which we will label s_0, t_0, and u_0. The λ that shows up in the Lagrangian is not necessarily a physical quantity.

The idea of *renormalization* is simple, albeit confusing in practice, and shows up in many branches of physics. The logic is this: The parameters of a theory will change when we add in interactions or compute an observable to a higher order in perturbation theory. As we've seen, we can regulate the theory to only depend on distance scales for which the physics is known, but this introduces a cutoff parameter. By renormalizing

the theory to depend only on physical, measurable quantities, we can remove this cutoff dependence and arrive at a sensible theory from which we can make predictions. In our example, instead of the coupling λ, we wish to instead write our answer in terms of some physically-measurable λ_p. This λ_p should be a function of Λ such that the amplitude remains invariant as we scale Λ up or down. In other words, we should not be able to choose an arbitrary value of Λ that changes the physics. Instead, we can absorb this dependence into the coupling parameters, which defines the physical couplings for us, rather than relying on the unknown parameter λ we wrote in our Lagrangian.

We then must solve for λ_0, the *bare parameter*[17], in terms of λ_p. One choice is to resum the higher-order corrections into the definition of the measured coupling, giving a renormalization condition,

$$-i\lambda_p = -i\lambda_0 + iC\lambda_0^2\left[\log\left(\frac{\Lambda^2}{s_0}\right) + \log\left(\frac{\Lambda^2}{t_0}\right) + \log\left(\frac{\Lambda^2}{u_0}\right)\right] + \mathcal{O}(\lambda_0^3), \tag{1.129}$$

for some chosen values of the Mandelstam variables at which we perform our experiment[18]. We can then solve for the bare parameter,

$$i\lambda_0 = i\lambda_p + iC\lambda_0^2\left[\log\left(\frac{\Lambda^2}{s_0}\right) + \log\left(\frac{\Lambda^2}{t_0}\right) + \log\left(\frac{\Lambda^2}{u_0}\right)\right] + \mathcal{O}(\lambda_0^3). \tag{1.130}$$

Plugging this into our result for $\mathcal{M}$, we find,

$$\begin{aligned}\mathcal{M}(\phi\phi\to\phi\phi) &= -i\lambda_p - iC\lambda_0^2\left[\log\left(\frac{\Lambda^2}{s_0}\right) + \log\left(\frac{\Lambda^2}{t_0}\right) + \log\left(\frac{\Lambda^2}{u_0}\right)\right] + \mathcal{O}(\lambda_p^3)\\ &\quad + iC\lambda_0^2\left[\log\left(\frac{\Lambda^2}{s}\right) + \log\left(\frac{\Lambda^2}{t}\right) + \log\left(\frac{\Lambda^2}{u}\right)\right] + \mathcal{O}(\lambda_0^3)\\ &= -i\lambda_p + iC\lambda_0^2\left[\log\left(\frac{s}{s_0}\right) + \log\left(\frac{t}{t_0}\right) + \log\left(\frac{u}{u_0}\right)\right] + \mathcal{O}(\lambda_0^3).\end{aligned} \tag{1.131}$$

[17] These are the parameters that appear in the regularized Lagrangian, as well as in the path integral.

[18] Crucially, this is a choice that is dependent on the *renormalization scheme*. Here, we illustrate the process in the on-shell renormalization scheme, but there are other schemes which define the physical coupling differently. As long as we are consistent about what we measure, and ensure that it corresponds to the chosen definition, this should not be a problem. While the numerical value of λ_p changes from scheme to scheme, physical measurements do not.

We can make the substitution $\lambda_0^2 \to \lambda_p^2$, which will only change the expression at higher order in λ_p. Finally, we have,

$$\mathcal{M}(\phi\phi \to \phi\phi) = -i\lambda_p + iC\lambda_p^2 \left[\log\left(\frac{s}{s_0}\right) + \log\left(\frac{t}{t_0}\right) + \log\left(\frac{u}{u_0}\right)\right] + \mathcal{O}(\lambda_p^3), \quad (1.132)$$

and our expression for the scattering amplitude no longer depends on a cutoff Λ, only the measured coupling λ_p and the chosen values of the Mandelstam variables at which our experiment is undertaken.

Suppose we instead would like to write down our Lagrangian from the beginning in terms of only physical couplings. Consider again the Lagrangian 1.120. Technically, this should be written,

$$\mathcal{L} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}m_0^2\phi^2 - \frac{\lambda_0}{4!}\phi^4, \quad (1.133)$$

where the subscripts denote the fact that we are using the bare parameters. We can instead express this only in terms of the physical parameters. To this end, we write,

$$\mathcal{L} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}m_p(\Lambda)^2\phi^2 - \frac{\lambda_p(\Lambda)}{4!}\phi^4 + A\partial_\mu\phi\partial^\mu\phi + B\phi^2 + C\phi^4, \quad (1.134)$$

where we must choose A, B, and C such that this Lagrangian results in the same physics as the one before. If we assume that the kinetic term is canonically normalized at the scale Λ, we can take $A = 0$. Then, we consider which computations lead us to actual observables of the theory. In this example, we can find B through the two-point function by computing the contribution of the counterterm Lagrangian to the propagator. C can be found in a similar manner through computation of the four-point function. We will not perform the calculation here, but the details can be found in several introductory textbooks [5] [4].

The model considered here is an example of a *renormalizable theory*, because we can successfully compute observables in terms of physical parameters. This means that, in

practice, as long as we can perform the necessary measurements, we can make meaningful predictions based on the values of these parameters. In introducing counterterms, we can see that we need to perform three renormalizations in order to fully express the Lagrangian in terms of physical quantities. Not all theories share this luxury, and are *nonrenormalizable* in the sense that an infinite number of counterterms would need to be introduced. We will expand upon this concept, and the conditions that lead to it, in 1.2.2.

3) Inclusion of higher order terms in the action. We can decompose the Lagrangian into a free field part plus an interacting part,

$$\mathcal{L} = \mathcal{L}_0 + \mathcal{L}_{int}. \tag{1.135}$$

The question then arises: How many terms should we include in $\mathcal{L}_{int}$? In principle, the answer is as many as we like. However, we need to make sure that, like the higher order terms in the perturbative expansion of the path integral, these additional terms do not lead to divergences in the path integral. Additionally, computations of loop corrections in higher order terms will in general lead to divergences in computations of scattering amplitudes. Depending on the particular term, these divergences may be severe.

To contend with these issues, the technique of *power counting* introduces explicit dependence on the cutoff Λ. This organizes the Lagrangian in terms of Λ in order to keep track of how each successive term scales. It is easiest to see how this is done by considering an in-depth example.

1.2.2 An Illustrative Example

Let's now consider a simple theory involving a scalar field ϕ in D spatial dimensions with a $\mathbb{Z}_2$ symmetry. Its action is,

$$S \supset \int d^D x \left(\frac{1}{2} \partial_\mu \phi \partial^\mu \phi - \frac{1}{2} m^2 \phi^2 - \frac{\lambda}{4!} \phi^4 - \frac{\zeta}{6!} \phi^6 + ... \right) \tag{1.136}$$

The action itself, in natural units, must have a mass dimension of zero, and so each term in the Lagrangian density must have mass dimension equal to D. Each derivative $\frac{\partial}{\partial x^\mu}$ has units of inverse length, giving it a mass dimension $[\frac{\partial}{\partial x^\mu}] = 1$. We then conclude that ϕ has mass dimension $[\phi] = \frac{D}{2} - 1$.

We immediately notice that, in knowing the mass dimension of the scalar field, each of the parameters in the above Lagrangian are dimensionful. It is not surprising to see that m^2 has mass dimension $[m^2] = 2$. We can also see that $[\lambda] = 4 - D$, $[\zeta] = 6 - 2D$, and so on.

We can understand the long-distance behavior of this theory by rescaling x^μ by a dimensionless parameter s,

$$x^\mu \to s x^{\mu'}. \tag{1.137}$$

In order to ensure that the kinetic term remains canonically normalized, we then make the field redefinition,

$$\phi \to s^{(1-D/2)} \phi' \tag{1.138}$$

The action is then,

$$S \supset \int d^D x' \left(\frac{1}{2} \partial_\mu \phi' \partial^\mu \phi' - \frac{1}{2} s^2 m^2 \phi'^2 - \frac{\lambda}{4!} s^{4-D} \phi'^4 - \frac{\zeta}{6!} s^{6-2D} \phi'^6 + ... \right) \tag{1.139}$$

The powers of s in each term are exactly the mass dimensions of each of their corresponding parameters. To see the long-distance behavior, we can now keep $x^{\mu'}$ fixed and

take the limit $s \to \infty$. The behavior of each term clearly depends on D; to be explicit, let's take the phenomenologist's favorite, $D = 4$. The action is,

$$S \supset \int d^4x' \left(\frac{1}{2}\partial_\mu \phi' \partial^\mu \phi' - \frac{1}{2} s^2 m^2 \phi'^2 - \frac{\lambda}{4!} s^0 \phi'^4 - \frac{\zeta}{6!} s^{-2} \phi'^6 + ... \right) \tag{1.140}$$

Because the kinetic term is canonically normalized, its behavior does not change in the $s \to \infty$ limit. The m^2 term will grow, the λ term will not change, and the ζ operator will tend to zero.

The reparameterization in terms of s is a nice way to explicitly categorize the behavior of each term. However, we could have skipped this exercise altogether and just looked at the mass dimension of the parameter corresponding to each term, or equivalently to the mass dimension of the operators themselves. To complete the categorization, we define each operator $\mathcal{O}$ as

1. **relevant** if $[\mathcal{O}] > 0$,
2. **marginal** if $[\mathcal{O}] = 0$, or
3. **irrelevant** if $[\mathcal{O}] < 0$.

In 1.140, the quadratic operator is relevant, the quartic operator is marginal, and the last operator written is irrelevant. Despite the terminology, it should be emphasized that irrelevant operators are not necessarily unimportant. To see this, consider again the action 1.139. We've seen that we can write the Lagrangian in terms of the physical coupling constants, which are functions of the cutoff scale Λ. We can explicitly make these parameters functions of this scale by rewriting them as, for example,

$$\zeta = \frac{c_1}{\Lambda^2}, \tag{1.141}$$

where c_1 is a dimensionless constant and we must have two inverse powers of Λ so that the whole term has a mass dimension of 4. The Lagrangian can then be written as an

expansion in powers of Λ,

$$\mathcal{L} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}m^2\phi^2 - \frac{\lambda}{4!}\phi^4 + \sum_{i=1}\frac{c_i}{\Lambda^{2i}}\phi^{2i+4}, \tag{1.142}$$

where we've absorbed further constants into the definitions of the c_i. This is the power counting technique. We can see that, if the energy scale E at which our interactions proceed is much less than Λ, irrelevant operators can be heavily suppressed. At larger Λ, these contributions will become more and more important.

The categorization of relevant, marginal, and irrelevant operators is directly related to their renormalization properties: All operators with mass dimension ≤ 4 are renormalizable, while operators with higher mass dimension are not. Power counting is an easy way to see this. Recall that a theory is renormalizable if we are able to rescale the couplings by introducing counterterms that completely absorb divergences without introducing new ones. Further, recall that our ϕ^4 operator led to a logarithmic divergence in the scattering amplitude at one-loop, stemming from the momentum dependence $1/k^4$, which can be fully absorbed by three counterterms. If we were to go to two-loop order, we would encounter an integral,

$$\mathcal{M}(\phi\phi \to \phi\phi)_{two\ loop} = \int \frac{d^4k_1}{(2\pi)^4}\frac{d^4k_2}{(2\pi)^4}\frac{1}{k_1^2k_2^2(k_1+k_2)^2(k_1+p_1+p_2)^2}, \tag{1.143}$$

where k_1 and k_2 are the momenta flowing in each loop, and p_1 and p_2 are the external momenta. We've omitted the dependence on m and $i\epsilon$ for brevity. This integral can be computed via dimensional regularization[19], in which we shift the number of dimensions $d = 4 - \epsilon$ for some small parameter ϵ which can be later taken to zero after the integral is performed. This has the advantage of converting divergences into simple poles of the

[19]This is a common renormalization scheme that, instead of computing integrals in d dimensions, introduces a small parameter ϵ and computes the integral in $d - \epsilon$ dimensions. This has the effect of shifting the integral away from the d dimensions in which the divergence occurs. At the end of the computation, we can take the limit $\epsilon \to 0$. This scheme is particularly useful because it preserves both Lorentz invariance and gauge invariance. Other schemes may not obey these symmetries, and the results must be treated with extra caution.

form $1/\epsilon$. We will not go through the details of the computation; the result is,

$$\mathcal{M}(\phi\phi \to \phi\phi)_{two\ loop} \sim \frac{1}{(4\pi)^4}\left(\frac{1}{\epsilon} + \text{finite terms}\right), \tag{1.144}$$

which is another logarithmic divergence. This means that the existing ϕ^4 counterterm can additionally absorb this divergence. This is true for all loop orders, and so we only end up needing three counterterms to render the amplitude finite.

If we now consider ϕ^6, the scattering process is $\phi\phi\phi \to \phi\phi\phi$. We will again spare the details of the computations, and simply state that the scattering amplitude at one-loop is,

$$\mathcal{M}(\phi\phi\phi \to \phi\phi\phi)_{one\ loop} \sim \frac{1}{(4\pi)^4}\left(\frac{1}{\epsilon} + \text{finite terms}\right) \tag{1.145}$$

and two-loop is,

$$\mathcal{M}(\phi\phi\phi \to \phi\phi\phi)_{two\ loop} \sim \frac{1}{(4\pi)^4}\left(\frac{1}{\epsilon^2} + \frac{1}{\epsilon} + \text{finite terms}\right). \tag{1.146}$$

In contrast to the phi^4 case, our two-loop amplitude is now quadratically divergent. This means that, to absorb this divergence, we must introduce a counterterm at the level of ϕ^8. In turn, the ϕ^8 counterterm will introduce a more severe divergence, requiring a counterterm at the order of ϕ^{10}. This eventually leads to an infinite number of counterterms, making ϕ^6 nonrenormalizable.

We can still study non-renormalizable theories up to some cutoff scale Λ — as long as the energy scale E is much less than Λ — they will just have much less predictive power than a completely renormalizable theory. There are many scenarios in which we wish to explore the theory at high order, or models of physics beyond the SM which include operators with high mass dimension.

1.2.3 Wilson's Approach

The approach to renormalization developed by Ken Wilson revolves around a low energy effective action,

$$S_{eff} = S_0 + S_{int}, \tag{1.147}$$

where S_{int} is the interacting part. This effective action is a function of both the fields and the cutoff scale Λ, so that the path integral is,

$$\int \mathcal{D}\phi_{|k|<\Lambda} e^{-S_{eff}[\phi]}. \tag{1.148}$$

We now only include modes with momentum $|k| < \Lambda$; modes above this scale are said to have been *integrated out*.

We treat the interacting part as an expansion in some power counting parameter E/Λ, with $E \ll \Lambda$. As we've seen, this is done by explicitly writing in the powers of Λ required to maintain a mass dimension equal to the spacetime dimension — such that the mass dimension of the action is zero — for each term in the Lagrangian. Higher-order terms then depend on dimensionless constants c_i, which are known as *Wilson coefficients*. These encode the effects of the high-energy physics on the low-energy observables, and are fixed by data through the renormalization process. We've seen that the physical couplings are functions of the cutoff scale, e.g. $\lambda_p(\Lambda)$. These are known as *running couplings*, because depend on a measurement made at some renormalization scale — which we've characterized by the Mandelstam variables s_0, t_0, and u_0. The Λ dependence on these couplings can be understood through the *beta function*,

$$\beta(g) \equiv \frac{\partial g}{\partial \ln \Lambda} = \Lambda \frac{\partial g}{\partial \Lambda}. \tag{1.149}$$

As we lower the scale Λ, we integrate out more high-energy modes in the effective action; the physics of these modes is now encoded in the couplings themselves. This process of

lowering Λ from the UV to the IR is known as *renormalization group (RG) flow*. Because we are successively integrating out modes, we are coarse-graining our theory. This flow is therefore one-way: While we can flow from the UV to the IR, this decreases the the available information content in our theory — we cannot reverse course from the IR to the UV. In fact, several points in the UV may flow to the same point in the IR, demonstrating how different microscopic details may exhibit the same behavior on large-scales. There is much more one could say about RG flow, which is outside the scope of this thesis, but a highly recommended excursion.

1.2.4 Operator Redundancies

Formulating the action in the Lagrangian formalism is convenient and lends itself well to the familiar Feynman diagram expansion, allowing one to find the Feynman rules for a theory and easily compute its leading diagrams. The prescription for calculating scattering amplitudes in this way is well-trodden ground in introductory QFT textbooks. These amplitudes, however, are the observable quantity at hand, not the fields themselves. It is well-known that the Lagrangian for a theory is not unique; this is why we need to agree upon a particular basis in which to perform calculations of the quantities that experiments will measure. We of course need to be able to agree that a measured value corresponds to a particular, well-defined parameter, so that our experimental conclusions are reproducible and universal.

There are several ways that one can transform a Lagrangian to obtain an equivalent Lagrangian. Namely,

1. **Integration by parts (IBP)**. Consider a scalar field ϕ. Its kinetic term, up to normalization factors, is $\partial_\mu\phi\partial^\mu\phi$. We can integrate by parts in 4 spacetime

dimensions to obtain,

$$\int d^4x \left(\partial_\mu \phi \partial^\mu \phi\right) = \partial_\mu \left(\phi \partial^\mu \phi\right) - \int d^4x \left(\phi \partial^2 \phi\right). \tag{1.150}$$

The total derivative term can be rewritten as a boundary term via Stokes' theorem, which then vanishes because it is assumed that ϕ vanishes at spacetime infinity. We are then free to make the replacement $\partial_\mu \phi \partial^\mu \phi \leftrightarrow -\phi \partial^2 \phi$. A Lagrangian that contains both of these terms would be said to contain an operator redundancy.

2. **Equations of motion (EOM)**. We can always obtain the EOMs for a field in a given theory via the Euler-Lagrange equations. For a simple scalar theory with Lagrangian

$$\mathcal{L} = \frac{1}{2}\partial_\mu \phi \partial^\mu \phi - \frac{1}{2} m^2 \phi^2, \tag{1.151}$$

one can easily find the EOM for ϕ to be

$$\ddot{\phi} = -m^2 \phi. \tag{1.152}$$

We are then free to make replacements using this expression.

3. **Field redefinitions**. We are free to redefine our fields as we see fit as long as the scattering amplitudes (the observables) themselves are unchanged. Consider a field redefinition on the ϕ field above,

$$\phi \to \phi + f(\phi). \tag{1.153}$$

The Lagrangian transforms as,

$$\mathcal{L} \to \frac{1}{2}\partial_\mu \phi \partial^\mu \phi + \partial_\mu \phi \partial^\mu \phi f'(\phi) + \frac{1}{2}\partial_\mu \phi \partial^\mu \phi (f'(\phi))^2 - \frac{1}{2} m^2 \phi^2 - m^2 \phi f(\phi) - \frac{1}{2} m^2 f(\phi)^2. \tag{1.154}$$

How can we be sure that that this Lagrangian leads to the same phyiscs as the previous one? Consider the generating functional,

$$Z[J] = \int \mathcal{D}\phi \, e^{iS[\phi]+i\int d^4x \, J(x)\phi(x)}. \tag{1.155}$$

Applying the field redefinition, we obtain,

$$Z'[J] = \int \mathcal{D}(\phi + f) \, e^{iS[\phi+f]+i\int d^4x \, J(x)(\phi(x)+f(x))}. \tag{1.156}$$

Because f is a fixed function, it does not affect the integration measure. We have, however, picked up a shift in the source term, $J(x)(\phi + f) = J(x)\phi(x) + J(x)f(x)$. We can write $Z'[J]$ in terms of $Z[J]$ as,

$$Z'[J] = e^{i\int d^4x \, J(x)f(x)} Z[J]. \tag{1.157}$$

To compute the correlation functions, we then take functional derivatives of the generating functional with respect to $J(x)$. The two-point function, for example, can be found via,

$$\langle 0|T\{\phi(x)\phi(y)\}|0\rangle = \frac{1}{Z[0]} \frac{\delta^2 Z[J]}{\delta J(x)\delta J(y)}\bigg|_{J=0}. \tag{1.158}$$

We find a shifted correlation function,

$$\langle 0|T\{\phi(x)\phi(y)\}|0\rangle + \langle 0|T\{\phi(x)f(\phi(y))\}|0\rangle + \langle 0|T\{f(\phi(x))\phi(y)\}|0\rangle + \langle 0|T\{f(\phi(x))f(\phi(y))\}|0\rangle. \tag{1.159}$$

The correlation functions are related to the the S-matrix elements — which are the observables of the system — via the LSZ reduction formula,

$$\langle f|S|i\rangle = \lim_{p_i^2 \to m^2} \prod_{i=1}^{n} \left(\int d^4x_i \, e^{ip_i\cdot x_i} (\Box_{x_i} + m^2) \right) \prod_{j=1}^{m} \left(\int d^4y_j \, e^{-iq_j\cdot y_j} (\Box_{y_j} + m^2) \right) \times \langle 0|T\{\phi(x_1)\ldots\phi(x_n)\phi(y_1)\ldots\phi(y_m)\}|0\rangle. \tag{1.160}$$

To compute the S-matrix elements using the LSZ formula, we need to extract the physical on-shell poles in the correlation function. We've picked up three new terms in the two-point function due to our field redefinition. The cross-terms involve correlations between the original field and the redefined part. However, because $f(\phi)$ is a function of the original field, it will not introduce new physical poles in the propagator. These cross terms do not contribute new poles in the on-shell limit, and so they do not change the physical S-matrix. Further, Similarly, the last term involves the function $f(\phi)$, but since $f(\phi)$ is typically chosen to be smooth and invertible, it does not alter the pole structure of the theory. $f(\phi)$ may modify the correlation functions off-shell, but it does not introduce new poles or alter the residues of the original poles. Hence, this term also does not affect the physical S-matrix.

We can then safely perform field redefinitions without worrying that they are changing the underlying physics. It should be noted that

It should be noted that EOMs are just a special case of field redefinitions. While field redefinitons may change the field in a non-trivial way, EOMs contain at most infinitesimal changes in the fields, because the variation of the action is required to vanish. We've seen above that we can make redefinitions on fields without changing the correlation functions; utilizing EOMs guarantees this as well.

When writing down an EFT, there are several input parameters: The fields, the symmetries these fields obey, and the cutoff scale and Wilson coefficients. Our Lagrangian may also include terms up to arbitrarily high mass dimension, and we must take care to ensure that we are not including redundant operators in our theory. At low mass dimension, it is straightforward enough to write down all possible operators that are consistent with the symmetries of the theory and check by hand that we've not included opera-

tors which may be transformed into one another via IBP, EOMs, or field redefinitions. However, at high mass dimension, this operation quickly becomes difficult.

This redundancy of description is intimately related to gauge invariance. Gauge invariance introduces redundancies in the description of the fields because gauge-equivalent field configurations lead to the same physical situation. Similarly, in EFT, operator redundancies arise because different ways of writing the operators in the Lagrangian can correspond to the same physical interaction. The physical observables, namely the scattering amplitudes, do not depend on how the operators are written as long as the physical content remains unchanged. Both gauge invariance and operator redundancies reflect the idea that not all degrees of freedom in a field theory correspond to physically meaningful quantities, and some are merely mathematical conveniences that can be transformed or removed without altering the physical content.

One might wonder if there is a way to formulate QFT without all of these redundancies. The answer is, so far, yes: There is a fruitful program of research into formulating on-shell scattering amplitudes in the spinor-helicity formalism. Redundancies may still show up in intermediate steps in these calculations, but the Ward identity ensures that the final answer will be gauge invariant. In some cases, this formalism drastically reduces messy algebra and condenses calculations that would take hundreds of pages in the Lagrangian formalism down to a few lines. These structures are proving to be invaluable computational devices, and are certainly are pushing forward our understanding of QFT itself. However, the beauty of EFTs is not to be disregarded; they remain powerful tools whose utility lies in their ability to and understand the low-energy consequences of higher-energy physics. In the quest to uncover new physics, and study theories that might constitute new physics, EFTs prove indispensable yet.

1.3 ΛCDM Cosmology

The early 1990s saw the onset of the "golden age of cosmology" in which large-scale cosmological experiments began to gather enormous quantities of high quality data on an array of astrophysical and cosmological objects. The launch of the COBE satellite in 1992 enabled the first measurements of the cosmic microwave background (CMB), providing a new window into the physics of the universe at the time of recombination. It discovered anisotropies in the CMB, a key prediction of inflation and component of models of large-scale structure formation. The angular power spectrum of the CMB was precisely mapped by WMAP in the 2000s, revealing the acoustic peaks that allow for measurements of the universe's curvature as well as the cosmological densities of both baryonic and dark matter. Observations of the bullet cluster and gravitational lensing experiments provided further evidence for dark matter, and distance measurements using type Ia supernovae suggested the accelerated expansion of the universe. The first observation of gravitational waves by LIGO in 2015 confirmed a key prediction of general relativity and allowed for detailed studies into the dynamics of black hole and neutron star inspirals. A host of upcoming cosmological experiments are poised to continue this age of discovery, with ever more precise gravitational wave observatories and infrared telescopes aimed at measuring the 21-cm hydrogen transition line.

Modeling the formation and evolution of the universe itself is a notably different endeavor than modeling the universe's particle content. We were able to verify the SM experimentally by accelerating particles up to a fraction of the speed of light, colliding them, and analyzing the subsequent map of energy deposition on the collider's calorimeter[20]. This is a repeatable procedure, and we can build up statistics of these

[20]This is, of course, an extremely coarse-grained view of how this process has played out. On a finer level, our knowledge of the SM is built up from a variety of tabletop experiments, various accelerator components aimed at distinguishing between different types of particles, and an enormous amount of sophisticated data analysis.

particle collisions over many runs of the experiment. In contrast, we can only observe the universe we are living in, and our measurements are inherently indirect. Rather than measuring direct products of collisions, we can only measure the byproducts of the universe's expansion and evolution: Galaxy structures and distributions, the cosmic microwave background (CMB), the astrophysical objects whose light is visible to our telescopes, and low-frequency gravitational waves (thus far). The universe was only formed once, and so precise observations of these byproducts and theory-building efforts to predict them from underlying principles are crucial in formulating our understanding. While terrestrial particle physics experiments are comprised of continuous testing and iterating, cosmological observation is more like piecing together evidence from the scene of a crime.

Rapid development on the observational side has enjoyed great agreement with the ΛCDM model of cosmology, the current paradigm describing the energy content and dynamics of the universe on large scales. It is the second "standard model" we encounter in this thesis, and we will review its key components in this section.

1.3.1 Spacetime Expansion

The first significant piece of evidence in formulating a thermal history of the universe lies in its expansion. Galaxy redshift data places the age of the universe at approximately 13.8 billion years. We now know, via measurements of redshifts from very early galaxies, that the universe is expanding, as light that is systematically shifted toward the red end of the spectrum points to an outward expansion. Observations of Cepheid variable stars — which radially pulsate at constant intervals and act as *standard candles* for distance measuring because they have the same intrinsic brightness — give further evidence to an accelerated nature of this expansion. An expanding universe that also obeys the

cosmological principle[21] is described by the the Friedmann-Robertson-Walker (FRW) metric,

$$ds^2 = -dt^2 + a^2(t)\gamma_{ij}dx^i dx^j, \tag{1.161}$$

where γ_{ij} is the metric of a maximally-symmetric 3-dimensional spacetime and $a(t)$ is the scale factor, which parametrizes the expansion rate. Written in spherical coordinates, this can be further parametrized in terms of a curvature parameter κ,

$$ds^2 = -dt^2 + a(t)^2\left[\frac{dr^2}{1-kr^2} + r^2\left(d\theta^2 + \sin^2\theta\, d\phi^2\right)\right], \tag{1.162}$$

where κ may lie anywhere between -1 and 1. $\kappa = 0$ corresponds to a flat space, $\kappa = 1$ corresponds to a maximally positively curved space, and $\kappa = -1$ corresponds to a maximally negatively curved space. We can solve Einstein's equations with this metric. Assuming that we can model the universe as a perfect fluid, the energy-momentum tensor is,

$$T_{\mu\nu} = (\rho + p)u_\mu u_\nu + p g_{\mu\nu}, \tag{1.163}$$

where ρ is the energy density of the fluid, p its pressure, and u_μ its four-velocity. Solving the temporal component[22],

$$R_{00} - \frac{1}{2}Rg_{00} + \Lambda g_{00} = \frac{8\pi G}{c^4}T_{00}, \tag{1.164}$$

we find,

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho + \frac{\Lambda}{3} - \frac{\kappa}{a^2}. \tag{1.165}$$

[21] This states that, on large scales, the universe is spatially homogeneous and isotropic. This is often taken as an assumption in cosmology, mostly in the context of it simplifying computations. The validity of the cosmologcial principle is an ongoing area of research.

[22] In this section, we will explicitly write factors of G and c.

The parameter $H \equiv (\dot{a}/a)$ is known as the *Hubble parameter*. Applying the same procedure to the spatial components,

$$R_{ii} - \frac{1}{2}Rg_{ii} + \Lambda g_{ii} = 8\pi T_{ii}, \tag{1.166}$$

yields,

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3p) + \frac{\Lambda}{3}. \tag{1.167}$$

Equations 1.165 and 1.167 are known as the *Friedmann equations*; with them, we can see how the dynamics of the universe depend directly on its contents. We can combine the Friedmann equations to obtain the *continuity equation*,

$$\dot{\rho} = -3H(\rho + p), \tag{1.168}$$

and classify how the energy density scales with $a(t)$. Non-relativistic matter can be assumed to be pressure-less and evidently scales as $\rho_m \propto a(t)^{-3}$, while (relativistic) radiation has a pressure $p = \frac{1}{3}\rho$ and scales as $\rho_r \propto a(t)^{-4}$. This implies that, by understanding how the species of matter and radiation in the universe evolved over time, we can correspondingly understand its expansion history. We can study each species according to its number density, energy density, and pressure,

$$\begin{aligned} n_a &= g_a \int \frac{d^3k}{(2\pi)^3} f_a(\mathbf{x}, \mathbf{k}) \\ \rho_a &= g_a \int \frac{d^3k}{(2\pi)^3} f_a(\mathbf{x}, \mathbf{k}) E(k) \\ P_a &= g_a \int \frac{d^3k}{(2\pi)^3} f_a(\mathbf{x}, \mathbf{k}) \frac{k^2}{3E(k)}, \end{aligned} \tag{1.169}$$

where a labels each species, $f_a(\mathbf{x}, \mathbf{k})$ gives its distribution in phase space, and g_a is the number of degrees of freedom of each species. Early on in the expansion history, we can approximate the species as being in equilibrium and sharing a common phase space distribution function,

$$f_a(E) = \frac{1}{e^{(E_a - \mu_a)/T} \pm 1}, \tag{1.170}$$

where T is the equilibrium temperature, μ_a is the chemical potential, and the $\pm$ factor becomes a $+$ for fermions and a $-$ for bosons. Due to the Boltzmann suppression factor, we can see that when the temperature T drops below the mass m_a of a given species, it will no longer significantly contribute to the distribution function. This implies that the energy density is radiation-dominated. We can solve for the radiation from 1.169, finding,

$$\rho_r = \frac{\pi^2}{30} g_* T^4, \text{ with } g_* \equiv \sum_{a=b} g_a + \frac{7}{8} \sum_{a=f} g_a, \tag{1.171}$$

where b denotes a sum over bosonic species and f a sum over fermionic species. Our equilibrium scenario is disrupted for interacting matter, which is what eventually transitions the universe out of a radiation-dominated era. There are two competing factors that determine the dominant energy density: The expansion rate and the interaction rate between species. For an expansion rate greater than the interaction rate, interacting species will eventually *freeze out*.

As the particles of the SM with the weakest interactions, neutrinos were the first known species to freeze out. After protons and neutrons froze out, light elements could be synthesized in a process known as *Big Bang Nucleosynthesis* (BBN). Both of these freeze outs occurred shortly after the Big Bang. Due to the high ambient temperature, these species existed in a dense, ionized plasma until about 380,000 years after the Big Bang, when the universe had cooled enough to form the first neutral hydrogen atoms in a process known as *recombination*. Prior to recombination, the abundance of free electrons led to a high rate of Thomson scattering, and so the universe was opaque. As neutral hydrogen atoms formed, the rate of Thomson scattering decreased enough for photons to decouple. As the universe expanded, photons eventually underwent a *last scattering* off an electron; these are the photons that make up the surface of last scattering, visible to us at the CMB.

1.3.2 The Cosmic Microwave Background

A hydrogen atom has a minimum ionization energy of 13.6 eV. Prior to the epoch of recombination, the universe was hot enough such that the energy of each photon exceeded this margin, and hence all hydrogen was fully ionized. To study recombination more concretely, we introduce the parameter χ_e describing the fraction of electrons in the universe that are free. We can then write this in terms of the number densities of free electrons, protons, and hydrogen,

$$\chi_e = \frac{n_e}{n_H + n_p} \tag{1.172}$$

During recombination, we assume a reaction,

$$p + e^- \longleftrightarrow H + \gamma \tag{1.173}$$

We make the assumption that this process occurs fast enough to ensure equilibrium. To estimate the relative abundances of protons, electrons, and hydrogen under this assumption, we can employ the *Saha equation*,

$$\frac{n_p n_e}{n_H} = \Big(\frac{m_e k_B T}{2\pi\hbar^2}\Big)^{3/2} e^{-E_I/k_B T} \tag{1.174}$$

Where m_e is the mass of the electron, k_B is Boltzmann's constant, T is the temperature, and E_I is the hydrogen ionization energy. We can rewrite this equation in terms of χ_e,

$$\frac{\chi_e^2}{1-\chi_e} = (n_H - n_p)^{-1}\Big(\frac{m_e k_B T}{2\pi\hbar^2}\Big)^{3/2} e^{-E_I/k_B T}, \tag{1.175}$$

and express quantities on the right-hand side of this equation in terms of redshift z. The temperature T scales as $T = 2.725(1+z)$, where T is units of Kelvin and

2.725 gives the current temperature of the CMB. The total hydrogen density scales as $n_H + n_p = 1.6(1+z)^{-3}$ in units of inverse cubic meters. With these relationships, we can estimate the redshift at which recombination occurred, and arriving a value $z \approx 1380$[23]. Comparing the size of the universe at this redshift to its present size, using the scale factor relationship $\rho_r \propto a^{-4}$, we find that the universe was roughly one-thousandth of its current size.

The CMB is almost perfectly described by the spectrum of a blackbody[24]. The energy density of this black-body radiation is given by,

$$\epsilon(\nu)d\nu = \frac{8\pi h}{c^3}\frac{\nu^3 df}{e^{h\nu/k_B T} - 1}, \tag{1.176}$$

where we write energy per unit volume of black-body radiation at a given temperature T. As the universe expands, the form of the blackbody is preserved, albeit at a lower temperature. This follows from the proportionality,

$$T_f \propto T_i \frac{a_i}{a_f}, \tag{1.177}$$

where the subscripts i and f denote the initial and final temperatures and scale factors.

The majority of radiation density in the universe is made up of CMB photons. Since these photons originated from the last scattering surface, and since we have seen that this surface was formed when the universe was roughly a thousandth of its current size, we observe them as coming from a large sphere centered on our galaxy.

Anisotropies

While the CMB is largely described by a blackbody radiation spectrum, it also contains tiny anisotropies on the order of one part in 10^5. In the image above, we see these

[23] For reference, recall that we take the redshift at the current epoch to be 0 by definition.

[24] This can be understood by the fact that this radiation was in a highly-interacting thermal state at the time of recombination

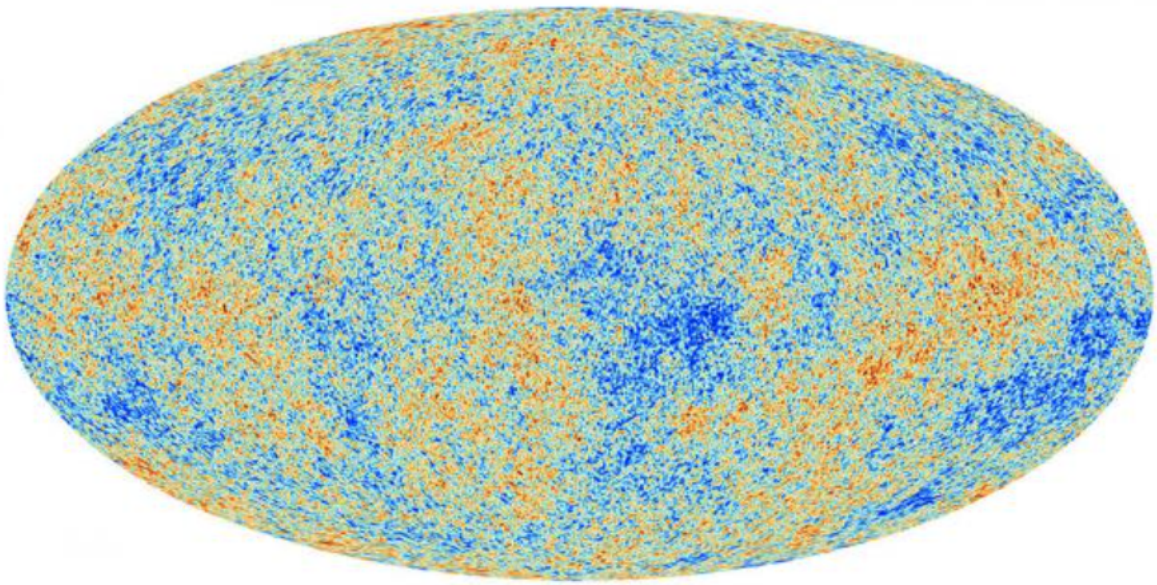

Figure 1.1: The CMB as seen by the Planck satellite. [6]

anisotropies as points with a slightly higher temperature, shown in red, and points with a slightly lower temperature, shown in blue, with respect to its average blackbody temperature. These scalar features of the CMB can be categorized primary and secondary anisotropies, depending on their source. The former were created during last scattering while the latter were created during the photon's path to our observational line of sight. There is a wide range of effects that create secondary anisotropies as the photon travels from the surface of last scattering, including reionization, an encounter with a gravitational potential well, or effects of gravitational lensing.

To analyze the primary anisotropies, we consider two main mechanisms by which they are created: Gravitational effects and adiabatic effects, which we will analyze in further detail. For general primary anisotropies Θ, we have,

$$\Theta = \Theta_{grav} + \Theta_{ad}. \tag{1.178}$$

A gravitational potential $\delta\Phi$ can affect CMB radiation via gravitational redshift as well as an added effect of time dilation, whereby we observe a slightly younger universe in overdense regions. Gravitational redshift and time dilation add anisotropies as, respectively,

$$\Theta \approx \delta\Phi, \ -\frac{2\delta\Phi}{3}. \tag{1.179}$$

The total primary anisotropy due to gravitational effects is then,

$$\Theta_{grav} \approx \frac{\delta\Phi}{3}. \tag{1.180}$$

This gravitational effect is more commonly known as the *Sachs-Wolfe effect*. Looking now at adiabatic perturbations, or perturbations that can be expressed in terms of some time shift, we can treat the dense plasma present in the universe prior to recombination as a perfect fluid of photon-baryons. This is because baryons were coupled to photons via Coulomb interactions, and Thomson scattering processes coupled photons to electrons. With radiation pressure acting as a restoring force, this means we can apply typical fluid dynamic analyses to the adiabatic perturbations. We can relate the local gravitational potential wells, $\delta\Phi$, to the matter density perturbations via Poisson's equation. It can then be shown using the continuity equation that the adiabatic perturbations are related to the matter density ρ by,

$$\Theta_{ad} = \frac{\delta\rho}{3\rho}. \tag{1.181}$$

Considering a model in which perturbations are affected only by changes in radiation pressure, ignoring gravity and effects of baryon mass-energy, we can use the fact that photon number is conserved and take into account the universe's expansion to arrive at a key equation,

$$\ddot{\Theta} + \frac{1}{3}k^2\Theta, \tag{1.182}$$

where k is a wave number in Fourier space. We see that this equation closely resembles that of a harmonic oscillator. Given this simplified model, this means that the radiation pressure acts as a restoring force for any initial perturbations. The factor of 1/3 in our equation can be related to the adiabatic speed of sound, and we further conclude that these oscillations will propagate at the speed of sound in our simplified fluid. This means that when matter and radiation decouple during recombination, any temperature fluctuations are effectively "frozen" into the photons, and these oscillations can be seen based on the correlations of the temperature anisotropies in the CMB.

Power Spectrum

The angular *power spectrum* of the CMB can be obtained by mapping the correlations between the anisotropies of the temperature map. We can first express the temperature map $T(\theta, \phi)$ as an expansion,

$$T(\theta, \phi) = \sum_{\ell m} a_{\ell m} Y_{\ell m}(\theta, \phi), \tag{1.183}$$

where $a_{\ell m}$ are coefficients and $Y_{\ell m}$ are the spherical harmonics,

$$Y_{\ell m}(\theta, \phi) = \sqrt{\frac{2\ell + 1}{4\pi} \frac{(\ell - m)!}{(\ell + m)!}} P_{\ell m}(\cos\theta) e^{im\phi}. \tag{1.184}$$

As usual, ℓ labels the multipole number and m the azimuthal number. The angle $\cos\theta$ depends on the relative orientation between two correlations. The power spectrum for a particular value of ℓ is defined as a sum over m[25],

$$C_\ell = \frac{1}{2\ell + 1} \sum_{m=-\ell}^{\ell} |a_{\ell m}|^2. \tag{1.185}$$

[25]Because the CMB is approximately homogeneous and isotropic, the power spectrum should not depend on the aximuthal angle.

The value $\ell = 0$ corresponds to the temperature averaged over the entire sky, $\ell = 1$ corresponds to the angular dipole, and so on. Because these fluctuations are assumed to be chosen randomly and distributed uniformly over the sky, we can define a Gaussian average to characterize the entire power spectrum,

$$\overline{C}_\ell = \left\langle \frac{1}{2\ell+1} \sum_{m=-\ell}^{\ell} |a_{\ell m}|^2 \right\rangle . \tag{1.186}$$

When we measure the power spectrum, our instruments take individual measurements over various angular scales corresponding to the value of ℓ[26]. At low values of ℓ, there are only a small number of measurements to average. This means that, even with zero measurement error, there may be random fluctuations that give values higher or lower than expected; this is known as *cosmic variance*. As we increase ℓ, we increase the number of measurements, and so any error due to cosmic variance is gradually decreased.

The power spectrum contains a succession of peaks known as *acoustic peaks*, and we can glean a lot of information from them. The first peak corresponds to the largest scale on which the photon-baryon fluid was able to undergo one complete oscillation by the time of recombination. The angular size of this *sound horizon* depends on the geometry of the universe: For $\kappa = 0$, the angular size corresponds directly to the ratio of the physical size of the sound horizon to the distance to the surface of last scattering. In a closed universe, $\kappa = 1$, the angular size of objects appears larger than in a flat universe because light travels along curved paths, making distant objects appear closer. In an open universe, $\kappa = -1$, the angular size of objects appears smaller because light is spread out over a larger distance due to the curvature. Measurements of the CMB show that the first angular peak occurs at $\ell \approx 220$, corresponding to an angular size of about 1°. This corresponds to a universe that is very close to flat.

[26]The angular scale of each ℓ is $360/(\ell+1)$. For example $\ell = 1$ corresponds to an angular scale of 180 degrees, $\ell = 2$ corresponds to an angular scale of 120 degrees, and so on.

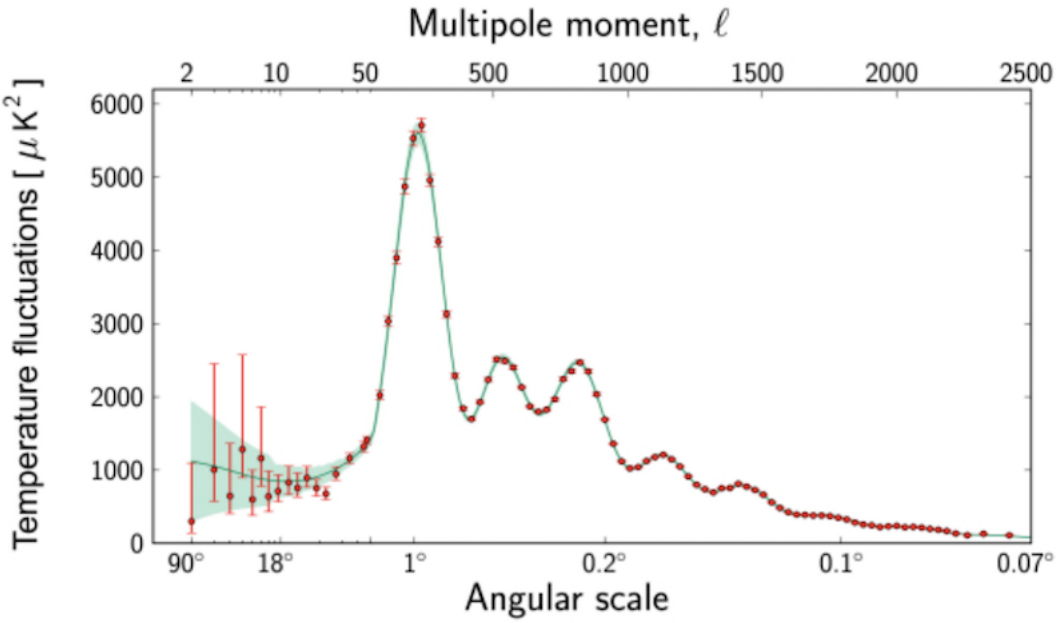

Figure 1.2: The CMB power spectrum. [7]

The more baryons there are, the more tightly the photon-baryon plasma is coupled in the early universe, which affects the amplitude of the acoustic oscillations. If the baryon density is higher, the gravitational wells are deeper during the compression phase, making it harder for photon pressure to push the plasma outward in the rarefaction phase. This results in a lower second peak. The baryon density can then be determined by the relative heights of the first and second peaks.

Transfer Functions

In analyzing the underlying physics of the power spectrum, our goal is to track the effects of fluctuations during recombination directly to the CMB anisotropies. Consider perturbations in the metric, using Newtonian gauge,

$$ds^2 = a^2(\tau)\left[-(1+2\Phi)d\tau^2 + (1-2\Psi)\delta_{ij}dx^i dx^j\right], \tag{1.187}$$

where $\Phi \approx \Psi$ act as gravitational potentials. One can show by integrating the geodesic equation that the photon momentum p is described by,

$$\frac{d}{d\tau}\ln(ap) = -\frac{d\Phi}{d\tau} + \nabla(\Phi + \Psi). \tag{1.188}$$

p is related to the temperature anisotropy as,

$$ap \propto a\bar{T}\left(1 + \frac{\delta T}{\bar{T}}\right). \tag{1.189}$$

Integrating 1.188 from the time of emission to the time of observation, and rewriting in terms of temperature, we arrive at,

$$\frac{\delta T}{\bar{T}}(\hat{n}) = \left(\frac{1}{4}\delta_\gamma + \Phi + \hat{n}\cdot v_e\right)_{\tau=\tau_*} + \int_{\tau_*}^{\tau_0} d\tau\,(\Phi + \dot{\Psi}). \tag{1.190}$$

The term $\hat{n}\cdot v_e$ accounts for the Doppler shift due to motion of electrons at the last scattering surface — $\hat{n}$ is the direction from which the photon enters a detector, while v_e is the electron's velocity. $\frac{1}{4}\delta_\gamma + \Phi$ describes the gravitational redshift of the photons from moving out of a small gravitational well. These are the dominant contributions to the temperature anisotropies. Substituting this result into the angular power spectrum gives us an expression,

$$C_\ell = \frac{4\pi}{(2\ell+1)^2}\int d\ln k\, T_\ell^2(k) P_\zeta(k), \tag{1.191}$$

where the power spectrum of each Fourier mode ζ_k is given by $P_\zeta(k) \equiv \frac{k^3}{2\pi^2}\langle\zeta_k\zeta_{-k}\rangle_0$. $T_\ell(k)$ is the *transfer function*, which describes the projection of the fluctuations onto the surface of last scattering as well as their evolution from recombination. It is given by,

$$T_\ell(k) = T_{\rm SW}(k) j_\ell(kr_*) + T_{\rm D}(k) j'_\ell(kr_*), \tag{1.192}$$

where $T_{\rm SW}(k) \equiv \left(\frac{1}{4}\delta_\gamma + \Phi\right)_* \zeta(k)$ and $T_{\rm D}(k) \equiv -\left(v_e\right)_* \zeta(k)$. The subscript denotes evaluation at the surface of last scattering.

1.3.3 The Inflationary Paradigm

The inflationary paradigm posits that the universe underwent a brief period of extremely rapid, exponential expansion in its earliest moments, within a fraction of a second after the Big Bang. While not a component of traditional Big Bang cosmology, the inflationary paradigm has gained much traction in recent decades due to its ability to elegantly solve several Big Bang puzzles and provide a compelling framework to understand structure formation. Here, we will review these puzzles, and review the formalism and consequences of inflation.

Motivations

The scale of the CMB's temperature anisotropies is on the order of 10^{-5}, while the CMB overall of a relatively uniform temperature of about 2.7 K. This is a significant result, because it means that places in the sky which are billions of light years away from each other somehow have the same temperature. As these regions are not causally connected, this poses a conundrum.

To be more rigorous, consider the size of the universe. As the universe evolves and expands, the particle horizon[27] changes by a factor of a_0/a_i, where a_i denotes the initial scale factor and a_0 denotes the scale factor at present time. Therefore, initially, this horizon scale was smaller by a factor of a_i/a_0. This tells us that the size of the universe at this time was roughly

$$l_i \approx ct_0 \frac{a_i}{a_0}. \tag{1.193}$$

We can then compare this estimate to the size of an initial casual region, the size of portions of space in causal contact with one another. This is given by $l_c \approx ct_i$. We can

[27] This is given by the distance that light has traveled since the beginning of the universe, integrated over time, considering the expansion of the universe

then take the ratio ,

$$\frac{l_i}{l_c} \approx \frac{t_0}{t_i}\frac{a_i}{a_0}. \tag{1.194}$$

We can find a numerical estimate from this expression by noting the relationship between the scale factor and temperature, $a_i/a_0 = T_0/T_i$. Estimating that primordial radiation dominates at the Planck time, we can plug in the Planck temperature to get,

$$\frac{l_i}{l_c} \approx 10^{28}. \tag{1.195}$$

This tells us that the initial size of the universe, at the Planck time, is greater than the size of a causal region by 10^{84}, or that there are approximately 10^{48} of these causally disconnected regions. This means that the matter distribution, which we observe as largely homogeneous, is unnaturally fine-tuned. This is often known as the *horizon problem*.

We can further consider the initial curvature via the Friedmann equation,

$$\Omega(t) - 1 = \frac{\kappa}{(Ha)^2}. \tag{1.196}$$

where $\Omega(t)$ is the cosmological density parameter, defined for a particular component as $\Omega_X = \rho_X/\rho_c$, with ρ_c the critical density[28]. The curvature parameter κ measured today is very close to zero, and so $\Omega(t)$ must be very close to 1. We can further find conditions on Ω at various epochs in the universe's evolution,

$$\begin{aligned} |\Omega(a_{\rm BBN}) - 1| &\leq \mathcal{O}(10^{-16}) \\ |\Omega(a_{\rm GUT}) - 1| &\leq \mathcal{O}(10^{-55}) \\ |\Omega(a_{\rm pl}) - 1| &\leq \mathcal{O}(10^{-61}). \end{aligned} \tag{1.197}$$

This is known as the *flatness problem* because this means that the early universe must have been very flat in order to satisfy cosmological constraints.

[28]The critical density, $\rho_c = \frac{3H^2}{8\pi G}$, is the energy density of a spatially flat universe

The punchline of this analysis[29] is that there are several aspects of Big Bang cosmology which become a story about extreme fine-tuning. While it is not necessarily forbidden that the universe was formed with a certain set of specific parameters, it makes the story less plausible and lessens the overall predictive power of Big Bang cosmology. Are these parameters due to an anthropic principle, or is there a physical way to explain why the universe is so homogeneous and isotropic? Can we explain these specific, fine-tuned initial conditions through a physical mechanism?

The answer is yes. The basic idea of *inflation* is that there was a period of exponential expansion early on in the evolution of the universe. By allowing several e-folds, or periods of time where the universe doubles in size due to exponential expansion, we can essentially "smooth out" the initial conditions. The inflationary paradigm typically includes an *inflaton* field responsible for inflation, whose potential energy vastly dominated its kinetic, leading to an equation of state $p \approx -\rho$ that drives the expansion. As inflation ends, the inflaton field decays to other particles in a process known as *reheating*.

Formalism

To understand exactly how inflation resolves the problems of initial conditions, let us formally consider inflation within the context of general relativity. We have seen in the previous discussion of initial conditions that the ratio $\frac{\dot{a}_0}{\dot{a}_i}$, which we can also write as $(aH)^{-1}$, seems to play a special role in both the horizon and flatness problems. In fact, these problems arise because of the behavior of this quantity; since the comoving Hubble radius is increasing with time, there are more causally disconnected regions and the flatness continually decreases. This leads to an idea: Invert this ratio early on in the universe.

[29] There is a third commonly-cited problem with Big Bang cosmology known as the monopole problem. In short, it asks why there are no magnetic monopoles, which are an abundant feature of many grand unified theories, in our universe.

First developed by Alan Guth in 1979 [8], the main idea behind inflation is that the universe underwent a period of adiabatic exponential expansion between 10^{-36} seconds after the big bang to 10^{-33} seconds after the big bang. We can use the Friedmann equations to see why the solution to the problems outlined above rests with exponential expansion.

By observational constraints, we can set the curvature $\kappa = 0$. The mechanism of inflation, accelerated expansion, requires that $\ddot{a} > 0$, which corresponds to $P < -\rho/3$. We can assume that inflation was not driven by radiation or ordinary, so a reasonable guess, based off of the behavior of cosmological constant-dominance, would give us $P \approx -\rho$. This yields a scale factor,

$$a(t) = a_i e^{H_i(t-t_i)} \tag{1.198}$$

This scale factor solves the horizon problem by allowing the comoving horizon to decrease with time. To be consistent with observations, it must be the case that all anisotropies today were in causal contact with each other in the far past, i.e. they were inside the same horizon. This is the condition,

$$\frac{1}{a_0 H_0} < \frac{1}{a_i H_i}. \tag{1.199}$$

The comoving particle horizon is given by,

$$\chi_p(t) = \int_0^t \frac{dt'}{a(t)'} = \int_0^a \frac{da}{aH^2} = \int_0^a d\ln(\frac{1}{aH}). \tag{1.200}$$

We can then solve for the number of *e-folds* of inflation: The number of times the universe expands by a factor of e. This is dependent on the details of reheating, but most models place the e-fold number somewhere in the range of $50 - 70$.

We also see that an exponentially increasing scale factor allows $\Omega(a)$ to tend toward 1, driving the universe toward flatness and allowing for a resolution to the flatness problem. All in all, the paradigm allows for a wider range of initial conditions for the early

universe. These initial conditions are then "inflated away" as the exponential expansion drives causal patches in the universe far away from each other and pushes the curvature parameter toward zero.

Scalar Perturbations

Through inflation, quantum fluctuations can be stretched to cosmic scales that seed large-scale structure. The Hubble radius H^{-1} — approximately the radius of the observable universe — is roughly constant during inflation, since the Hubble parameter is given by $\dot{a}/a$ and the scale factor is an exponential. Quantum fluctuations in the matter and scalar fields during this time are being constantly generated and are stretched as space expands, eventually stretching beyond the Hubble radius and exiting the horizon. Scalar fluctuations in particular can be characterized by the power spectrum,

$$\langle \mathcal{R}_k \mathcal{R}_{k'} \rangle = (2\pi)^3 \delta(k+k') P_{\mathcal{R}}(k), \quad \Delta_s^2 \equiv \Delta_{\mathcal{R}}^2 = \frac{k^3}{2\pi^2} P_{\mathcal{R}}(k), \tag{1.201}$$

where the second quantity defines the *dimensionless power spectrum*. The scale dependence of the power spectrum is parameterized by the *spectral tilt* n_s,

$$n_s - 1 \equiv \frac{d \ln \Delta_s^2}{d \ln k}. \tag{1.202}$$

The momentum dependence is parameterized by the *running of the spectral index*,

$$\alpha_s \equiv \frac{dn_s}{d \ln k}. \tag{1.203}$$

This allows us to parameterize the dimensionless power spectrum as an approximate power law,

$$\Delta_s^2(k) = A_s(k_*) \left(\frac{k}{k_*} \right)^{n_s(k_*) - 1 + \frac{1}{2}\alpha_s(k_*) \ln\left(\frac{k}{k_*}\right)}, \tag{1.204}$$

given in terms of a reference scale k_* known as the *pivot scale*.

Now, let's consider the generation of these scalar perturbations during inflation and compute the resultant power spectra. We define the comoving curvature perturbation, a gauge-invariant quantity, as

$$\mathcal{R} \equiv \Psi - \frac{H}{\rho + p}\delta q \tag{1.205}$$

where Ψ is a spatially flat hypersurface and δq is the scalar part of the 3-momentum density. Cosmological inhomogeneities are characterized by $\mathcal{R}$, which tells us about the intrinsic spatial curvature of co-moving or constant hypersurfaces. This quantity is conserved outside the horizon, and so we quantize scalar fluctuations during inflation by writing the equation of motion for $\mathcal{R}$ in the form of a simple harmonic oscillator. We start with a single-field slow roll model[30], in which we have one scalar inflaton field ϕ, defined by the action

$$S = \frac{1}{2}\int d^4x\sqrt{-g}[R - (\nabla\phi)^2 - 2V(\phi)]. \tag{1.206}$$

Here, we choose units where $M_{pl}^{-2} = 1$. In order to fix our reparametrizations, we choose a gauge. In this case, it is convenient choose the gauge where the inflaton field is unperturbed and scalar degrees of freedom are characterized by metric fluctuations,

$$\delta\phi = 0,\ g_{ij} = a^2[(1 - 2\mathcal{R})\delta_{ij} + h_{ij}],\ \partial_i h_{ij} = 0. \tag{1.207}$$

We can then expand the action to second order in $\mathcal{R}$ to get,

$$S = \frac{1}{2}\int d^4x a^3 \frac{\dot{\phi}^2}{H^2}[\dot{\mathcal{R}}^2 + a^{-2}(\partial_i\mathcal{R}^2]. \tag{1.208}$$

We then define the *Mukhanov variable* $v \equiv z\mathcal{R}$ where $z \equiv a^2\frac{\dot{\phi}^2}{H^2}$. Rewriting the action in conformal time yields

$$S = \frac{1}{2}\int d\tau dx^3 \left[(v'^2) + (\partial_i v)^2 + \frac{z''}{z}v^2\right]. \tag{1.209}$$

[30] We will consider alternative models of inflation in the next section.

Using a Fourier expansion of the field v, we arrive at the *Mukhanov equation.*

$$v_k'' + \Big(k^2 + \frac{z''}{z}\Big)v_k = 0 \tag{1.210}$$

Since the function z depends on background dynamics, this equation is difficult to entirely solve. We can arrive at boundary conditions characterizing the solutions by first quantizing the field v,

$$\hat{v}_k = v_k(\tau)\hat{a}_k + v_{-k}^*(\tau)\hat{a}_{-k}^\dagger, \tag{1.211}$$

with the creation and annihilation operators satisfying the usual commutation relations if the mode functions are normalized as,

$$\langle v_k, v_k\rangle \equiv \frac{i}{\hbar}\left(v_k^* v_k' - v_k^{*'} v_k\right) = 1. \tag{1.212}$$

We can choose the vacuum state for the fluctuations to be the Minkowski vacuum in the far past, $\tau \to -\infty$. The mode equation in this limit becomes that of a simple harmonic oscillator, which motivates the initial condition,

$$\lim_{\tau\to-\infty} v_k = \frac{e^{-ik\tau}}{\sqrt{2k}}. \tag{1.213}$$

In the de Sitter limit, perhaps the closest approximation to inflationary expansion, $z''/z = 2/\tau^2$. Plugging this, and our ansatz for v_k, back into the Mukhanov equation yields,

$$v_k = \frac{e^{-ikr}}{\sqrt{2k}}\Big(1 - \frac{i}{\kappa\tau}\Big), \tag{1.214}$$

the *Bunch-Davies* mode function. With this result, we can now compute the power spectrum; consider the field defined by $\hat{\psi}_{\mathbf{k}} = \hat{v}_{\mathbf{k}}/a$,

$$\langle\hat{\psi}_k(\tau)\hat{\psi}_{k'}(\tau)\rangle = (2\pi)^3\delta(k+k')\frac{H^2}{2k^3}\left(1 + k^2\tau^2\right). \tag{1.215}$$

On superhorizon scales, for which $k\tau \ll 1$, this becomes a constant,

$$\langle\hat{\psi}_k(\tau)\hat{\psi}_{k'}(\tau)\rangle \to (2\pi)^3\delta(k+k')\frac{H^2}{2k^3}. \tag{1.216}$$

We can relate the curvature perturbation in a spatially-flat gauge directly to the perturbations of the inflation field ϕ driving inflation: $\mathcal{R} = H\psi/\dot{\phi}$. The power spectrum at horizon crossing $k = a(t_*)H_*$ is,

$$\langle \mathcal{R}_k(t)\mathcal{R}_{k'}(t)\rangle = (2\pi)^3\delta(k+k')\frac{H^2}{2k^3}\frac{H^2}{\dot{\phi}^2}. \tag{1.217}$$

The dimensionless power spectrum is then,

$$\Delta^2_{\mathcal{R}}(k) = \frac{H_*^2}{(2\pi)^2}\frac{H_*^2}{\dot{\phi}_*^2}. \tag{1.218}$$

We can see that the spectrum at horizon crossing, because $\mathcal{R}$ approaches a constant, determines the spectrum until the modes re-enter the horizon.

We can also the compute the power spectrum of the vacuum fluctuations, given by the operator,

$$\hat{v}(\tau,\mathbf{x})\int\frac{d^3k}{(2\pi)^3}\hat{v}_k e^{i\mathbf{k}\cdot\mathbf{x}}. \tag{1.219}$$

While the expectation value of $\hat{v}(\tau,\mathbf{x})$ vanishes, it has a nonzero variance,

$$\langle|\hat{v}(\tau,\mathbf{x})|^2\rangle = \int\frac{d^3k}{(2\pi)^3}|v_k(\tau)|^2 = \int d\ln k\frac{k^3}{(2\pi)^3}|v_k(\tau)|^2. \tag{1.220}$$

The dimensionless power spectrum is then,

$$\Delta^2_v(k,\tau) = |v_k(\tau)|^2 = a^2\left(\frac{H}{2\pi}\right)^2(1+(k\tau)). \tag{1.221}$$

As $k\tau\to 0$, the superhorizon limit, the power spectrum is the same for all momenta, and is therefore scale-invariant.

Tensor Perturbations

The production of a stochastic gravitational wave background is a key prediction of inflation. In a similar way that electromagnetic waves are derived as propagating solutions

of Maxwell's equations, and come in two distinct linear polarizations, gravitational waves are derived as propagating solutions to Einstein's equations. It can be shown that these equations imply two distinct polarizations for gravitational waves as well, which we will denote + and ×. Gravitational waves are the transverse, traceless components of the metric perturbation. This can be described as a perturbation to the FRW metric

$$g_{ij} = a^2(\delta_{ij} + 2h_{ij}) \tag{1.222}$$

We can consider these perturbations via the Einstein-Hilbert action. Expanding this action to quadratic order in h_{ij} yields

$$S = \frac{1}{4}\int dt \int d\mathbf{x}^3 a^3 \left[(1/2)(\dot{h_{ij}})^2 - \frac{1}{2a^2}(\partial_k h_{ij})^2\right] \tag{1.223}$$

where $\mathbf{x}$ is a comoving coordinate. We can reformulate this equation in terms of Fourier components and the two gravitational wave polarization amplitudes h_+ and $h_\times$.

$$S = \sum_{p=+,\times}\sum_{k}\int dt a^3 \left[\frac{1}{2}|\dot{v_{p,\mathbf{k}}}^2| - \frac{1}{2}(k/a)^2|v_{p,\mathbf{k}}^2|\right] \tag{1.224}$$

where $v_p = h_p/2$ and $\mathbf{k}$ is the usual wave vector. The structure of this resulting equation tells us that the gravitational wave Fourier modes and polarization state amplitudes behave as simple harmonic oscillators. We can then define the gravitational-wave power spectrum by noticing that each perturbation is a realization of a random field that is randomly drawn from a Gaussian distribution. The variance of v, utilizing the usual result for a simple harmonic oscillator, is then $\langle v^2\rangle = \hbar/2\omega$. The variance of the distribution of tensor fluctuations will then be,

$$\langle|h_{p,\mathbf{k}}|^2\rangle = \frac{H_*^2}{4k^3}. \tag{1.225}$$

The dimensionless power spectrum is then given by

$$\Delta_t^2 = 2\frac{k^3}{2\pi^2}\langle|h_{p,\mathbf{k}}|^2\rangle = \frac{2H_*^2}{\pi^2}, \tag{1.226}$$

where we multiply by a factor of 2 in order to account for the two different polarization states. We can again see that this spectrum is scale-invariant.

The power spectra of both the scalar and tensor perturbation depend, as expected, on the Hubble scale at horizon exit, and therefore on the energy scale of inflation. This also implies that can probe this energy scale through the *tensor-to-scalar ratio*,

$$r \equiv \frac{\Delta_t^2(k)}{\Delta_s^2(k)}, \tag{1.227}$$

where $\Delta_s^2(k)$ denotes the dimensionless power spectrum of the scalar fluctuations. The inflationary potential V — which we will detail in the next section — is proportional to H^2. V can be related to the tensor-to-scalar ratio as,

$$V^{1/4} \sim \left(\frac{r}{0.01}\right)^{1/4} 10^{16}\,\text{GeV}, \tag{1.228}$$

and thus constitutes a direct measurement of Hubble during inflation. Upper limits on the tensor-to-scalar ratio by BICEP and Planck [9] give $r \lesssim 0.06$. This translates into a bound on $H \lesssim 10^{13}$ GeV.

1.3.4 Cold Dark Matter

The first hints of the existence of large quantities of extra, unseen matter in the universe came about via the observation of galaxy rotation curves in the 1930s. The rotation curves of spiral galaxies – showing the orbital speeds of visible matter in the galaxy as a function of their distance from the galactic center – appeared flat, rather than displaying the $\propto r^{-1/2}$ dependence expected from Newtonian gravity. If we assume that galaxy rotation follows Newtonian gravity, as other orbital systems do, the galactic mass would have to be up to five times larger than what could be accounted for by visible matter. The rotation curves then imply that this extra mass would scale as $M(r) \propto r$ in order to match the observed flatness; its density would scale as $\propto M(r)/r^3 \sim r^{-2}$. This

suggests that the distribution of this matter in galaxies is more like a halo, in contrast with the disk-like structure observed from visible matter.

At this point, there were several possibilities for how to reconcile this observation. Perhaps this signaled a breakdown of the predominant gravitational theory, or that observational instruments could not capture enough light necessary to reveal the presence of a preponderance of dim astrophysical objects. In the subsequent decades, these questions were sharpened with the development of ΛCDM model of cosmology, which constitutes the leading modern paradigm. As every piece of observational evidence strengthened the predictions of General Relativity (GR) and it became clear that known astrophysical objects alone could not account for enough matter to form the necessary massive halo structures, the possibility of a new type of matter altogether emerged.

Further observations have also pointed to the presence of a large amount of hidden matter. Gravitational lensing measurements can map out this matter indirectly by observing the gravitationally-assisted bending of light around galaxies and other massive objects. Lensing measurements of the Bullet Cluster, consisting of a pair of colliding galaxy clusters, reveal a center of mass inconsistent with that predicted by the cluster's visible matter. Measurements of the mass-energy density of the universe via the power spectrum of the CMB by Planck have led to the following breakdown: Only about 5 percent of this mass-energy, assuming a ΛCDM model of cosmology, is made up of baryonic matter. Seventy percent consists of dark energy. The remaining twenty-five percent is other, non-visible matter; this is what we mean by dark matter (DM).

This breakdown comes from measurements of the power spectrum. We've seen that the first and second peaks are sensitive to the amount of baryonic matter in the universe at the time of recombination. The balance between radiation pressure and gravitational pull due to DM is also reflected in the positions and heights of the power spectrum peaks. DM creates deeper wells than baryons alone would, which leads to stronger compressions

during the odd-numbered peaks. This means that more dark matter makes the third peak more pronounced relative to the second peak. The position of the first peak, when combined with the observed matter density and the flatness of the universe, implies the existence of a component that makes up the remaining energy density to achieve $\Omega_{tot} \approx 1$. This missing component is attributed to the cosmological constant.

The DM of ΛCDM has specific properties. While remaining agnostic to the actual microscopic model, the DM in this paradigm is:

1. **Cold**. In particular, the DM is non-relativistic. This is essential for the growth of structure formation. Low velocities of dark matter allow efficient clumping, providing gravitational wells in which baryonic matter can also collapse to form stars and galaxies. If dark matter were hot and relativistic, its high velocities would prevent it from clustering on small scales, as it would have a tendency to smooth out small-scale density fluctuations due to its free-streaming behavior[31]. This would hinder the formation of small structures and delay large-scale structure formation.

2. **Non-dissipative**. One of the primary reasons dark matter must be non-dissipative in the ΛCDM model is that it is electromagnetically neutral. It then cannot emit or absorb light, so it cannot cool through radiaton. The non-dissipative nature of DM is crucial for explaining why it forms diffuse halos around galaxies, while baryonic matter collapses into the dense, luminous regions we observe as galaxies.

3. **Collisionless**. Along the same lines, DM must be collisionless, interacting with each other primarily through gravity. The Bullet Cluster provides direct evidence that DM is non-dissipative and collisionless: While the baryonic gas was slowed

[31]The free-streaming length is the distance a particle can travel before it is significantly slowed by gravity. Cold dark matter has a very small free-streaming length, which means that it clusters on small scales.

down by collisions, the DM — inferred through gravitational lensing — passed through the collision without interacting, indicating that DM does not dissipate energy.

While ΛCDM posits the existence of cold DM, it does not posit a microscopic model for it. In the next section, we will provide an overview of several models of DM, some of which deviate from the above guidelines in a small enough way to maintain the successes of ΛCDM and remain consistent with observation.

1.4 Hints of New Physics

The SM and the ΛCDM model both constitute our current best understanding of the particle physics and cosmology of our universe. It is also clear that this cannot be the end of the story. In this section, we will dive into the long list of reasons why there must be physics beyond the SM. We will start with the naturalness problems, then consider models to explain the nature of DM, introduce microscopic models of inflation, and end with an overview of many other rich topics that we will not fully detail here.

1.4.1 The Naturalness Problems

In our study of EFTs, we saw that physical phenomena that are characterized by different scales decouple from one another and can be understood in a largely independent fashion. A theory violates the *naturalness principle* if its parameters depend on scales that are significantly separated from one another. For example, in computing the self-energy of the electron with a cutoff scale Λ, it receives a classical contribution of the form $\Delta m_e \sim e^2\Lambda$, signifying a divergence. The electron self-energy must then be *fine-tuned* in order to fit observational data.

Such talk of divergences recalls renormalization: We've seen that the standard procedure for curing divergences in computations of QFT observables is to regularize the integral by introducing a cutoff scale Λ, then redefine our parameters to be anchored to a physical measurement. By parameterizing our ignorance of the high momentum modes, renormalization allows us to make predictions of the physics at lower momentum. However, our ultimate goal is to systematically discover the new physics beyond the cutoff. Indeed, the divergence in the electron self-energy was eventually cured by the inclusion of the positron, whose divergent contribution to the self-energy cancelled that of the electron. Further, this modification was motivated by the inclusion of a (then) new symmetry: CPT. In a similar fashion, divergences between the self-energy of neutral kaons led to the successful prediction of a (then) new degree of freedom: The charm quark [10]. Historically, finding solutions to problems that display fine-tuning has guided the discovery of new physics.

There are a few ways that we can more rigorously define naturalness as a principle. The main idea is that parameters in a more complete theory of physics should not require fine-tuning, and instead admit dynamical explanations for their values. Parameters are said to be *technically natural* if their size in the UV is not significantly influenced by intermediate-scale physics, as in the example of the electron self-energy. More strongly, they are *'t Hooft natural* if, when the parameter is set to zero, a symmetry is restored.

There are three major outstanding problems that consist of a finely-tuned parameter that is natural under either definition. Collectively, they make up the *naturalness problems*. We will go through each of them below, and focus our extended discussion on the strong CP problem, which will be important for later chapters of this thesis.

The Hierarchy Problem

The *hierarchy problem* is the question of why the Higgs mass is 125 GeV, in light of the fact that it can receive large quantum corrections that should drive it up to much higher values. Quantum corrections to the Higgs mass take the form,

$$\delta m_H^2 \sim \frac{\Lambda^2}{16\pi^2}, \tag{1.229}$$

where Λ is the usual cutoff scale, which in this case may be as high as the Planck mass. Without fine-tuning, the Higgs mass should be much larger than its observed value, unless a mechanism exists to cancel or regulate these large contributions.

Supersymmetry (SUSY) has long been a popular solution to the hierarchy problem. By introducing an R-parity symmetry — which exchanges bosons for their fermionic superpartner and fermions for their bosonic superpartner — the divergent quantum corrections to the Higgs mass now receive a contribution of the opposite sign from the superpartner of each particle. The corrections then take the form,

$$\delta m_H^2 \sim \frac{\Lambda^2}{16\pi^2} - \frac{\Lambda^2}{16\pi^2}, \tag{1.230}$$

and the Higgs mass is rendered stable without the need for fine-tuning. Although SUSY has since fallen out of favor due to the null LHC results of the past decade, it may still address the hierarchy problem if the mass difference between the particles and their superpartners is not too large. In other words, SUSY must be softly broken to remain a viable solution.

The Cosmological Constant Problem

The cosmological constant problem refers to the large discrepancy between the theoretically predicted value of the vacuum energy density — the cosmological constant Λ — from quantum field theory and the observed value from cosmological measurements.

The vacuum energy density $\rho_{\rm vac}$ in QFT is expected to receive contributions from various sources, including the zero-point energies of quantum fields. These contributions are expected to be on the order of the Planck scale,

$$\rho_{\rm vac} \sim \frac{M_{\rm Pl}^4}{16\pi^2}. \tag{1.231}$$

However, the observed value of the cosmological constant, inferred from the accelerated expansion of the universe, is extremely small,

$$\rho_{\rm obs} \sim 10^{-120} M_{\rm Pl}^4. \tag{1.232}$$

representing a discrepancy of 120 orders of magnitude between the theory prediction and observational result. The value in natural units of the cosmological constant is,

$$\Lambda \approx 1.1 \times 10^{-52}\,{\rm m}^{-2}, \tag{1.233}$$

an incredibly tiny value that renders the cosmological constant problem the most severe of the naturalness problems. It is also the problem for which there exist the fewest satisfactory explanations, and constitutes one of the most significant unresolved problems in high energy physics.

The Strong CP Problem

CP, describes the invariance of the behavior of a system under the simultaneous transformations of charge conjugation and parity. The weak force has been observed to violate CP, while the strong force has thus far produced no such observed occurrence.

However, using our model-building techniques of writing down all terms in the Lagrangian of a theory that obey all discrete and continuous symmetries, there should be a CP-violating term in the QCD Lagrangian. Namely,

$$\mathcal{L}_{QCD} \supset \theta \frac{g^2}{32\pi^2} {\rm Tr}(G_{\mu\nu}\tilde{G}^{\mu\nu}), \tag{1.234}$$

where G is the gluon field strength, g its gauge coupling, and θ a phase. This term is a total derivative, and arises from non-perturbative effects. CP violation also occurs due to the Yukawa terms, and so the total amount of CP violation in the strong sector is characterized by,

$$\bar{\theta} = \theta + \text{Arg}(\text{Det}M_q), \tag{1.235}$$

where M_q is the quark mass matrix. Bounds on $\bar{\theta}$ come from measurements of the neutron eDM, which gives the constraint,

$$\bar{\theta} \lesssim 10^{-10}. \tag{1.236}$$

The strong CP problem is then the question of why $\bar{\theta}$ is so small. This is a fine-tuning problem of 1 part in 10^{10}, which is why it is categorized as a naturalness problem. A solution would be able to explain the smallness of $\bar{\theta}$, rather than relying on this parameter to be an input.

A massless up quark. In principle, the strong CP problem may be solved in a straightforward way — if the up quark were massless. When $m_u = 0$, there is a chiral symmetry in the quark sector of the QCD Lagrangian. In particular, we can perform a chiral rotation,

$$u \to e^{i\alpha\gamma_5} u, \tag{1.237}$$

which allows the θ term to be rotated away through an appropriate choice of the phase α. However, this solution is disfavored. The experimentally measured pion masses and decay constants are consistent with the up quark having a small but nonzero mass, and the mass differences between the proton and neutron would be difficult to explain if the up quark were massless. Lattice QCD calculations also suggest the up quark mass is small but nonzero.

The Nelson-Barr mechanism. The idea of the Nelson-Barr mechanism is to introduce spontaneous CP violation and an extended particle content in such a way that θ is automatically zero at tree-level. The mechanism assumes that CP is a fundamental symmetry of the theory, but it is spontaneously broken by a scalar field acquiring a VEV, which avoids the introduction of explicit CP-violating terms in the Lagrangian. We introduce new vector-like quarks such that the Yukawa terms are,

$$\mathcal{L}_{\text{NB}} = -(y_u)_{ij}\bar{Q}_{L,i}HU_{R,j} - (y_d)_{ij}\bar{Q}_{L,i}HD_{R,j} - (y_q)_{ij}\bar{q}_{L,i}\phi q_{R,j} - (y_{\text{mix}})_{ij}\bar{Q}_{L,i}Hq_{R,j} + \text{h.c.}, \tag{1.238}$$

where Q denote the SM quarks, q the new vector-like quarks, and y the Yukawa matrices. We further introduce a new scalar ϕ responsible for spontaneous CP-breaking, with $\langle\phi\rangle = v_\phi e^{i\alpha}$. The full quark mass matrix is given by,

$$\mathcal{M} = \begin{pmatrix} M_Q & M_{\text{mix}}^\dagger \\ M_{\text{mix}} & M_q \end{pmatrix}, \tag{1.239}$$

After spontaneous CP-breaking, $M_q = v_\phi e^{i\alpha} y_q$ and $M_{\text{mix}} = v y_{\text{mix}}$. The determinant of the mass matrix is,

$$\det(\mathcal{M}) = \det(M_q)\det(M_Q - M_{\text{mix}}M_q^{-1}M_{\text{mix}}^\dagger), \tag{1.240}$$

where the off-diagonal terms cancel the phase contribution in the full determinant, ensuring that $\det(\mathcal{M})$ is real and its argument is zero. Due to the CP symmetry prior to its breaking, θ is initially zero.

While solving the strong CP problem, this solution introduces new sources of fine-tuning to ensure that both the θ angle vanishes and the determinant of the mass matrix remains real. Solutions based on CP are therefore quite fragile, and have become disfavored [11].

The axion. Perhaps the most currently popular proposed solution to the strong CP problem is the axion, a a dynamical scalar particle a that is the pseudo-Nambu-Goldstone boson of a spontaneously broken $U(1)_{P,Q}$ symmetry, known as Peccei-Quinn (PQ) symmetry. Adding this into the QCD Lagrangian yields,

$$\mathcal{L} \supset \frac{\alpha_s}{4\pi}\left(\bar{\theta} - \frac{a}{f_{PQ}} G_{\mu\nu}\tilde{G}^{\mu\nu}\right), \tag{1.241}$$

where α_s is the strong coupling constant and f_{PQ} the scale of the PQ symmetry breaking. The idea is then that a may be rotated into $\bar{\theta}$ by the PQ symmetry, such that it finds a minimum at $\langle a \rangle = \bar{\theta}$, giving an overall contribution of zero above f_{PQ}.

The axion is appealing for many reasons, the first of which is that it is a fairly simple ingredient to add into the model. $U(1)$ symmetries show up frequently in renormalizable models of BSM physics, and it is not a huge stretch to posit the existence of an additional $U(1)$ at higher energies. They are also appealing for an entirely different reason: They are able to make up the dark matter relic abundance, simultaneously solving two of the major outstanding problems of particle physics.

There is a crucial snag, however, that takes the name of the *axion quality problem*. In order to be a robust solution to the strong CP problem, $U(1)_{PQ}$ must be a *high quality* global symmetry. There may be other interactions that change the shape of the axion potential, and it is particularly sensitive to Planck-suppressed operators, which could throw off the careful cancellation that occurs in 1.241. There also may be further sources of CP violation from as-yet unknown BSM physics, such as CP violation that explains the baryon-anti-baryon asymmetry.

Anthropics

While the naturalness principle has certainly led to successful predictions and postdictions of new particles, its logic is inductive. On one hand, there is a logical precedent to

flesh out the full range of solutions to the naturalness puzzles that retain the naturalness principle. On the other hand, perhaps induction can only take us so far, and naturalness will at some point fail, if it hasn't already done so.

In some sense, the naturalness problems are not problems per se. The values for the parameters of the SM, as calculated by Weinberg, fall just in the range that leads to the conditions for life [12]. For example, if the cosmological constant significantly larger than its observed value, the universe would have expanded too quickly for galaxies to form. Conversely, if it were negative and too large in magnitude, the universe would collapse too early. This leads to an anthropic bound on the cosmological constant,

$$\Lambda \lesssim \frac{3}{8\pi G \rho_{\text{matter}}}, \tag{1.242}$$

which translates to an upper bound that lies quite close to the observed value. Perhaps, then, we observe a universe that has these values because this is one of the few that we could even exist in.

The anthropic principle, however, does not make testable predictions, and relies on the possible existence of other universes in which the fundamental parameters are different. Further, while it can in some sense explain why we observe small values for certain constants, it does not provide a dynamical mechanism for why these values are chosen, effectively sidestepping the problem.

1.4.2 What's the (Dark) Matter?

While ΛCDM cosmology posits the existence of cold DM, it remains very much an open question as to what this DM actually is. In this section, we will discuss the prevailing question of what could constitute DM and give a brief survey of the current landscape of DM candidates. Our discussion is inspired by material presented in the lecture notes [13] [14].

Making a DM Candidate

There have been numerous attempts to explain what DM is, utilizing a variety of creative approaches. First, what if DM actually is baryonic matter, but simply too dim to be properly accounted for? A number of astrophysical objects, such as black holes, neutron stars, and brown dwarfs, are difficult to detect and could easily fit this description. Unfortunately, this perhaps simplest explanation has been ruled out by gravitational lensing measurements. Importantly, *primordial* black holes, which are thought to form through large density perturbations in the early universe, remain a compelling DM candidate. This is because black holes formed through gravitational stellar collapse are restricted to certain ranges of masses based on the Schwarzschild radius, while primordial black holes may originate prior to stellar formation via primordial density perturbations and are therefore not subject to the same constraints.

Since the discrepancy in the measurement of visible matter versus its predicted distribution assumes General Relativity (GR), one might ask if GR is insufficient to describe the dynamics of galaxies. Indeed, modified theories of gravity have been constructed that can fit well with observed galaxy rotation curves, and we know that GR cannot be an all-encompassing theory of gravity. However, these theories often fail to mesh with the larger cosmological picture. GR has consistently passed every experimental test we've thrown at it, including precision measurements of the gravitational constant, observations of weak-lensing in the solar system, and the direct detection of gravitational waves. Thus far, modified theories have not been able to simultaneously predict flat galaxy rotation curves while maintaining the overwhelming success of GR. Further, modified theories are more difficult to justify in light of (no pun intended) indirect observations of DM via lensing.

If DM is not comprised of hidden baryonic matter and not a product of corrections

to GR, it is natural to consider explanations rooted in the hypothesis of new wave-like or particle-like matter. After all, we have a Standard Model of baryonic matter that has enjoyed excellent agreement with data. There is no a priori reason why DM could not also adhere to similar particle properties. What, then, makes for a good microscopic DM candidate? There are a few criteria to consider:

1) Its interactions must be predominantly gravitational. Observational evidence strongly suggests that DM interactions are primarily gravitational. DM is also often assumed to be collisionless. However, some models propose weak self-interacting dark matter (SIDM), where dark matter particles interact with each other through a new, weak force. In other models, DM could carry a small, nonzero electric charge, making it millicharged and allowing for very weak electromagnetic interactions.

2) Its interactions must be non-dissipative. However, there are dissipative models in which DM particles can interact with each other via forces beyond gravity, leading to energy loss. A key component of these models is the existence of a dark sector with new particles and forces. These interactions could be mediated by dark photons or other light bosons that act similarly to photons in the SM. DM particles could emit dark photons, carrying away energy allowing dark matter to cool and collapse into compact structures.

3) It must result in the necessary relic abundance. There are two primary mechanisms that may lead to this: *freeze-out* and *freeze-in*.

The freeze-out mechanism is the most commonly considered scenario for DM production. It assumes that DM particles were in thermal equilibrium with the SM particles in the early universe, with interactions like $\chi\bar{\chi} \leftrightarrow SM$ with χ a generic SM particle. At early times, the number density follows the Boltzmann distribution. As the universe expands and cools, the interaction rate $\Gamma \sim n_\chi \langle \sigma v \rangle$, where n_χ is the number density

and $\langle \sigma v \rangle$ the thermally-averaged cross-section, falls below the expansion rate. Then, DM particles can no longer annihilate or be produced efficiently. They decouple and their number density becomes fixed, leading to the observed relic abundance. The evolution of the number density is governed by the Boltzmann equation,

$$\frac{dn_\chi}{dt} + 3Hn_\chi = -\langle \sigma v \rangle \left(n_\chi^2 - n_{\chi,\text{eq}}^2 \right), \tag{1.243}$$

where $n_{\chi,\text{eq}}$ is the number density of DM during equilibrium. The relic abundance is inversely proportional to the annihilation cross section,

$$\Omega_\chi h^2 \propto \frac{1}{\langle \sigma v \rangle}, \tag{1.244}$$

where h is the dimensionless Hubble constant. The annihilation cross-section must lie within a certain range to match the observed relic abundance. If the interactions are too weak, the DM would overclose the universe[32], and if the interactions are too strong, there would be too little DM today.

In the freeze-in mechanism, DM particles are never in thermal equilibrium with SM particles. Instead, they are produced through very weak interactions, like decays or scatterings off heavy particles, $SM + SM \to SM + \chi$. As the temperature decreases, production continues until the universe cools sufficiently that further production becomes negligible and their number density freezes in, leading to the relic abundance. This number density is described by the Boltzmann equation,

$$\frac{dn_\chi}{dt} + 3Hn_\chi = \langle \sigma v \rangle n_{SM}^2, \tag{1.245}$$

where n_{SM} is the number density of SM particles. The relic abundance is determined by the strength of the interactions with the thermal bath, and it scales with the interaction

[32]This is a scenario in which the total density of matter and energy in the universe exceeds the critical density, leading to a closed geometry that will eventually stop expanding and begin to contract.

strength as,

$$\Omega_\chi h^2 \propto \langle \sigma v \rangle. \tag{1.246}$$

Freeze-in requires precise tuning of the interaction cross-section and production mechanism, so that the interaction rate is weak enough that DM never thermalizes, but strong enough to produce the correct relic abundance. This abundance is,

$$\Omega_\chi h^2 \approx 0.12. \tag{1.247}$$

4) It must lead to a consistent cosmology. DM was originally introduced to explain galactic rotation curves, and through lensing experiments and cluster observations, is thought to form a halo-like structure surrounding a galaxy. This has interesting implications for the formation of large-scale structure in the universe. DM must account for the gravitational pull necessary to explain the formation of large-scale structures over cosmic time, while having a free-streaming length small enough to allow for the formation of small structures such as galaxies. Models with slightly higher free-streaming lengths can still be viable but must reproduce the correct galaxy-scale structure. Too much free streaming, as in the case of hot DM, would wash out small-scale structures.

A DM model must also fit the anisotropies in the CMB, in particular leading to the right amount to account for the power spectrum peaks, as well as fit the baryon acoustic oscillations[33] observed in surveys of large-scale structure.

Finally, DM should not significantly alter the predictions of BBN, which is highly sensitive to the energy content of the universe during its first few minutes. DM must not

[33]BAO are regular, periodic fluctuations in the density of visible matter (baryons) in the universe, imprinted in the large-scale structure as a result of sound waves that propagated through the early universe. These sound waves were generated during recombination in the primordial plasma composed of photons, baryons, and DM. They appear as a slight enhancement in the probability of finding galaxies separated by a characteristic scale corresponding to corresponds to the maximum distance the sound waves could have traveled before the universe cooled enough for photons and baryons to decouple. This feature can be seen in the 2-point correlation function or the power spectrum of galaxy surveys.

contribute too much energy during this time or interact strongly with these processes.

5) It must have thus far evaded detection. Like any new particle added to the SM, there must be properties of DM that explain why it has not yet been detected by current or previous experiments, while adhering to bounds set by direct and indirect detection experiments.

The Current Landscape

Current bounds leave open large swaths of parameter space for possible DM candidates, which physicists have slowly been chipping away at in the past decades. There are a number of creative proposals, as well as experiments aimed at various possible production channels [14]. To get a sense for this landscape, we will focus on a few well-considered DM candidates in this section, in order to illustrate the main ideas in performing a DM calculation.

Weakly-Interacting Massive Particles (WIMPs). WIMPS are natural candidates for DM obeying a freeze-out mechanism. They are predicted to have a relic density consistent with the observed dark matter abundance if their annihilation cross-section is approximately the weak interaction scale, which leads to the so-called"WIMP miracle". WIMP masses typically range from a few GeV to a few TeV, which is a natural range in many BSM theories such supersymmetry.

While WIMPs are theoretically compelling, direct and indirect detection experiments have yet to observe their signatures. The absence of detection has pushed the bounds on the WIMP-nucleon interaction cross-section to extremely small values, making it harder to discover WIMPs unless their interactions are weaker or their mass is different from the initially expected range.

Axions. We've introduced axions already in the context of solving the strong CP problem. They also constitute perhaps the current most popular proposal for a particle DM candidate, following the realization that they are naturally light and weakly interacting. Axions are predicted to have a very small mass, in the range of 10^{-5} eV to 10^{-2} eV. This makes them much lighter than WIMPs, but because of their production mechanism, they are non-relativistic at the time of structure formation. They also interact very weakly with other particles, primarily through couplings to photons, fermions, and gluons, which are suppressed by the scale f_{PQ}. In order for axions to have evaded detection thus far, this scale lies around 10^9 GeV to 10^{12} GeV.

Axions may be non-thermally produced in the early universe through different mechanisms. In the vacuum misalignment mechanism, the axion field is initially displaced from its minimum in the potential after PQ symmetry breaking. As the universe expands and cools, the axion field begins to oscillate around the minimum of the potential. These oscillations correspond to a population of non-relativistic axions, making them a natural candidate for cold DM. Further, the axion number density remains constant after the oscillations start, which is analogous to the freeze-out process for WIMPs but without requiring thermal equilibrium. The relic density of axions from this mechanism is,

$$\Omega_a h^2 \sim \left(\frac{10^{12}\,\text{GeV}}{f_a}\right)^{7/6}, \tag{1.248}$$

which becomes close to the observed relic abundance for $f_{PQ} \sim 10^{12}$ GeV.

Another commonly-considered production mechanism involves the topological defects formed from PQ symmetry breaking in the early universe, such as axion strings and domain walls. Axion strings can form after breaking, and produce a population of axions when they decay. Domain walls are formed in an analogous way, and also contribute to axion density upon decay.

The list goes on... Aside from axions and WIMPs, several alternative dark matter candidates are being actively explored. Sterile neutrinos are a leading candidate for warm dark matter (WDM), produced through non-thermal mechanisms and detectable via potential X-ray emissions from their decay. Self-interacting dark matter (SIDM) posits that dark matter particles can interact with each other, potentially solving small-scale structure issues in galaxies, such as the core-cusp problem. Fuzzy dark matter, consisting of ultralight bosons with masses around 10^{-22} eV, exhibits wave-like behavior on galactic scales and could help explain galactic core structures. Primordial black holes (PBHs), which form in the early universe, offer a non-particle-based dark matter candidate, potentially detectable through gravitational lensing or mergers observed by gravitational wave observatories. Finally, light dark matter (LDM) and dark photons offer possibilities in the sub-GeV mass range, with searches focused on accelerator-based experiments and novel direct detection techniques. These diverse candidates each address different theoretical challenges and remain key targets in the ongoing search for dark matter. There are certainly no shortage of additional DM models, with the development of many more an ongoing process.

1.4.3 The Inflationary Mechanism

The inflationary paradigm, though not verified, is quite successful at resolving the traditional problems with Big Bang cosmology. Actually implementing inflation, however, requires new physics, because do not know what the inflating mechanism is. There are several key predictions of inflation theory that would serve as evidence toward its paradigm, and the tests of the paradigm are dependent on which specific type of inflation is being studied. There is no current consensus among cosmologists as to what the triggering mechanism of inflation was, and so several variations on the inflation idea exist,

each of which center around a slightly different formulation of the inflaton potential and subsequent field.

It is typically considered to be inflaton field, a scalar ϕ, that has a high potential energy prior to inflation that drives the expansion itself. We will detail this model, and outline a few others, in this section.

Single-Field Slow Roll Model

The simplest model is single-field slow roll inflation, in which a single scalar inflaton field ϕ has a high potential energy prior to inflation that drives the expansion itself. In this model, the total energy density of the inflaton is,

$$\rho = \frac{1}{2}\dot{\phi}^2 + \frac{1}{2}\frac{(\nabla\phi)^2}{a^2} + V(\phi) \tag{1.249}$$

If we make the assumption that the scalar field is spatially homogeneous, the gradient term goes away and we are left with,

$$\rho = \frac{1}{2}\dot{\phi}^2 + V(\phi). \tag{1.250}$$

The corresponding pressure of the inflaton field is,

$$p = \frac{1}{2}\dot{\phi}^2 - V(\phi). \tag{1.251}$$

The magic of the slow-rolling field then comes in since we notice that, if the inflaton field only changes slowly, we obtain the condition,

$$\dot{\phi}^2 \ll V(\phi). \tag{1.252}$$

In this case, we see that

$$\rho \approx -p \approx V(\phi), \tag{1.253}$$

and the inflaton field acts similarly to a cosmological constant. This means that an exponential expansion can be driven via an inflaton field if the field is slow-rolling through

some period of time and if its potential energy dominates the energy density of the universe, analogous to the behavior of a cosmological constant. The slow roll is characterized by a parameter,

$$\varepsilon = -\frac{\dot{H}}{H^2}. \tag{1.254}$$

When $\varepsilon < 1$, accelerated expansion occurs. The expansion must also be sustained over a long-enough period of time, giving the condition,

$$|\ddot{\phi}| < |V_{,\phi}|. \tag{1.255}$$

where the subscript on V indicates that we take the derivative of the potential with respect to ϕ. This leads to a second slow roll parameter,

$$\eta = -\frac{\ddot{\phi}}{H\dot{\phi}}, \tag{1.256}$$

whose absolute value must also be less than 1. We can write the slow roll parameters in terms of the inflaton potential,

$$v(\phi) \equiv \frac{M_{\rm pl}^2}{2}\left(\frac{V_{,\phi}}{V}\right)^2, \quad \eta_v(\phi) \equiv M_{\rm pl}^2\frac{V_{,\phi\phi}}{V}. \tag{1.257}$$

We've seen that, to formulate single-field slow roll inflation, we can make use of an action with a standard kinetic term minimally coupled to gravity,

$$S = \int d^4x\sqrt{-g}\Big[\frac{1}{2}M_{pl}^2 R - \frac{1}{2}g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - V(\phi)\Big], \tag{1.258}$$

where R is the Ricci scalar. During inflation, spacetime is approximately de-Sitter, $a(t) \sim e^{Ht}$. Varying the action with this metric, we obtain the equations of motion for the inflaton field,

$$\ddot{\phi} + \partial_\phi V = 0. \tag{1.259}$$

Models with Modified Potentials

There are no shortage of further inflationary models to discuss that build off of the simplest single-field slow roll model. Perhaps the most immediate aspect of inflationary models that can be tinkered with is the potential. What features should this potential have? We have seen above that we can start with general slow-roll conditions, although models are not necessarily constrained to this requirement. One feature that each potential must have, however, is some mechanism to exit inflation. Specifically, a "graceful exit" is a necessity: One that does not interfere with the structure of the universe immediately after inflation. This is because we want to preserve the homogeneity and isotropy, which were the motivations for introducing an inflationary paradigm in the first place. Each of the following potentials can be characterized by its state prior to inflation as well as its mechanism for a graceful exit. In Figure 1.4.3 we can see three commonly used

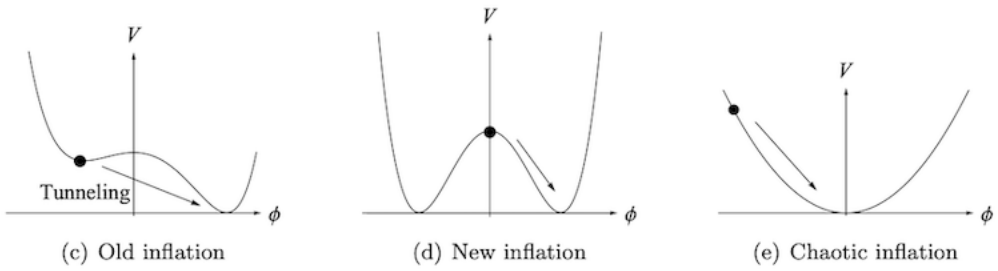

Figure 1.3: A sampling of commonly-used inflationary potentials. [15]

inflationary potentials. The first one, old inflation, assumes that our scalar field starts at the local minimum of the potential, driving inflation as it tunnels to the true vacuum. The second potential, which corresponds to a Coleman-Weinberg type potential, is given by

$$V(\phi) = V_0\Big[\Big(\frac{\phi}{\mu}\Big)^4\Big(\ln\Big(\frac{\phi}{\mu} - \frac{1}{4}\Big) + \frac{1}{4}\Big)\Big] \tag{1.260}$$

This is also the potential for radiatively-induced symmetry breaking in electroweak and grand unified theories, and so this is a particularly popular popular inflationary model. In this model, the scalar field escapes from the maximum due to quantum fluctuations. Energy is homogeneously released as it rolls toward its minimum value.

The third potential is one formulation of chaotic inflation. This is a sub-type of a class of models known as large-field models, where the inflaton field starts off at some large value and evolves to zero at the origin. In chaotic inflation, the potential is dominated by the term

$$V(\phi) = \lambda_p \phi^p \tag{1.261}$$

where λ_p is a coupling constant. This self-coupling should be very small in order for the density fluctuations to be small. These models are known as chaotic models since they may have nearly any set of initial conditions, and this is the broadest possible class of models that satisfy the slow-roll parameters.

Multi-field Models and the Curvaton

Single-field inflation models rely on the inflaton to both inflate the universe and be responsible for the primordial density fluctuations. Multi-field models, as the name suggests, introduce multiple fields, some of which may be responsible for inflation and some of which may be spectator fields. In the *curvaton* model, the inflaton field does the inflating, and a second scalar field — the curvaton — is responsible for the observed density fluctuations. The curvaton is subdominant during inflation, but can later convert its fluctuations into curvature perturbations. The simplest curvaton model features the

potential,

$$V(\phi,\sigma) = \frac{1}{2}m_\phi^2\phi^2 + \frac{1}{2}m_\sigma^2\sigma^2, \tag{1.262}$$

where the curvaton mass $m_\sigma \ll H$. Its equation of motion, using an FRW metric, is given by,

$$\ddot{\sigma} + 3H\dot{\sigma} + V'(\sigma) = 0. \tag{1.263}$$

We can capture the dynamics of the inflaton and curvaton fields by splitting them into their interacting and background values, e.g. $\sigma(t,\mathbf{x}) = \sigma_0(t) + \delta\sigma(t,\mathbf{x})$. We assume that the curvaton kinetic energy is much smaller than that of the inflaton,

$$\epsilon \equiv -\frac{\dot{H}}{H^2} \approx \frac{\dot{\phi}^2}{2H^2M_{\rm pl}^2}, \tag{1.264}$$

and so the curvaton is slowly-rolling along its potential,

$$\dot{\sigma}_0 \approx -\frac{m^2}{3H}\sigma_0. \tag{1.265}$$

Crucially, curvature fluctuations now no longer need to be generated immediately after inflation. As the inflaton reheats following inflation, the curvaton will slowly roll along its potential until its mass reaches the Hubble scale. At this point, the curvaton will oscillate about its potential and dilute like matter. The curvature perturbation is related to the curvaton by,

$$\zeta \approx \left(\frac{3\rho_\sigma}{4\rho_r + 3\rho_\sigma}\right)\zeta_\sigma, \tag{1.266}$$

where ρ_r gives the radiation density, and ζ_σ are the curvature fluctuations generated by the curvaton. The factor out front gives the ratio of the curvaton energy density to the total energy density. The dimensionless power spectrum of the curvaton will be scale-invariant,

$$\Delta_\sigma^2(k) \sim \left(\frac{H_*}{2\pi}\right)^2, \tag{1.267}$$

and the power spectrum of curvature perturbations is,

$$\mathcal{P}_\zeta = \left(\frac{3\rho_\sigma}{4\rho_r + 3\rho_\sigma}\right)^2 \left(\frac{H_*}{2\pi\sigma_*}\right)^2 . \tag{1.268}$$

Primordial Non-Gaussianities

All current observations are consistent with Gaussian primordial curvature perturbations. Inflationary models, however, predict the generation of *primordial non-Gaussianities*, deviations from a perfect Gaussian distribution in the statistics of the primordial curvature perturbations. A Gaussian distribution implies that the statistical properties of the perturbations are entirely characterized by the two-point correlation function — or equivalently, the power spectrum. Non-Gaussianities require higher-order correlation functions, such as the three-point function or the four-point function, to fully describe the statistics of the perturbations.

The curvature perturbation $\zeta(x)$ is usually expanded into a sum of Gaussian and non-Gaussian components,

$$\zeta(x) = \zeta_G(x) + f_{\rm NL}\zeta_G^2(x) + \cdots , \tag{1.269}$$

where $f_{\rm NL}$ is a non-linearity parameter that characterizes deviations from Gaussianity, and $\zeta_G(x)$ gives the Gaussian part of the distribution. The first non-Gaussianity is given by the three-point correlation function,

$$\langle\zeta(k_1)\zeta(k_2)\zeta(k_3)\rangle = (2\pi)^3\delta^{(3)}(k_1 + k_2 + k_3)B_\zeta(k_1, k_2, k_3), \tag{1.270}$$

where we explicitly write momentum-conserving delta functions. The *bispectrum* $B_\zeta(k_1, k_2, k_3)$ characterizes the amplitude and shape of the non-Gaussianities, and is commonly parameterized as,

$$B_\zeta(k_1, k_2, k_3) = 2f_{\rm NL}\left[P_\zeta(k_1)P_\zeta(k_2) + P_\zeta(k_2)P_\zeta(k_3) + P_\zeta(k_3)P_\zeta(k_1)\right], \tag{1.271}$$

where P_ζ is the usual power spectrum. Different models of inflation predict different shapes and magnitudes of the bispectrum, which are described by different configurations of k_1, k_2, k_3. The non-Gaussianities are then characterized by the shape of the momentum-conserving triangle formed by k_1, k_2, k_3.

Local non-Gaussianity is generated locally in real space, and the bispectrum is largest when one of the momenta is much smaller than the other; this is the *squeezed limit*. In this limit, the bispectrum is approximated as,

$$B_\zeta(k_1, k_2, k_3) \sim \frac{1}{k_1^3 k_2^3} + \text{permutations} = \tag{1.272}$$

Equilateral non-Gaussianity refers to a bispectrum in which $k_1 = k_2 = k_3$. The bispectrum goes like,

$$B_\zeta(k_1, k_2, k_3) \sim \frac{1}{k_1^3 k_2^3 k_3^3}. \tag{1.273}$$

Finally, folded non-Gaussianity occurs when $k_1 \approx k_2 + k_3$. Each shape is characterized by its own f_{NL} parameter. For arbitrary shapes, we can measure the magnitude of the non-Gaussianity as,

$$f_{\rm NL} \equiv \frac{5}{18} \frac{B_{\mathcal{R}}(k, k, k)}{P_{\mathcal{R}}(k)^2}. \tag{1.274}$$

Non-Gaussianities are also described by higher-point correlation functions, such as the four-point function,

$$\langle \zeta(k_1)\zeta(k_2)\zeta(k_3)\zeta(k_4)\rangle = (2\pi)^3 \delta^{(3)}(k_1 + k_2 + k_3 + k_4) T_\zeta(k_1, k_2, k_3, k_4), \tag{1.275}$$

where $T_\zeta(k_1, k_2, k_3, k_4)$ is the trispectrum.

The latest constraints on non-Gaussianities from Planck are [16],

$$f_{NL}^{\rm local} = -0.9 \pm 5.1,\ f_{NL}^{\rm equil} = -26 \pm 47,\ f_{NL}^{\rm ortho} = -38 \pm 24, \tag{1.276}$$

where the orthogonal shape is a combination of other shapes. These results are consistent with Gaussian primordial perturbations, and no significant detection of non-Gaussianities has so far been made.

Cosmological Collider Physics

A more detailed calculation shows that, in the squeezed limit, three-point function of curvature perturbations depends on momenta as,

$$\langle \zeta(\mathbf{k}_1)\zeta(\mathbf{k}_2)\zeta(\mathbf{k}_3)\rangle \propto e^{-\pi\mu}\frac{1}{k_1^3}\frac{1}{k_3^3}\left(\frac{k_3}{k_1}\right)^{\Delta(m)} e^{i\delta(\mu)}, \tag{1.277}$$

where $k_3 \ll k_1, k_2$ and $\Delta(m)$ depends on the mass m of the interacting (non-inflaton) particle as,

$$\Delta(m) = \frac{3}{2} + i\sqrt{\frac{m^2}{H^2} - \frac{9}{4}}. \tag{1.278}$$

Immediately, we can see that the behavior of the three point function can be divided into three distinct regimes. When the mass is much higher than Hubble, $m \gg H$, the Boltzmann factor in 1.277 dominates, and the signal becomes heavily suppressed. When the mass is much lower than Hubble, $m \ll H$, 1.278 is real, corresponding to an analytic signature. However, in the "Goldilocks zone" of $m \gtrsim \frac{3H}{2}$, 1.278 contains an imaginary part, generating a non-Gaussian oscillatory signal that is directly proportional to the mass. Further, this signature is sensitive to the actual value of m, opening up the possibility of distinguishing interacting particles from one another. If the particle has a nontrivial spin, the proportionality of 1.277 may also contain angle-dependent Legendre polynomials, allowing for further avenues for identification [17]. This possibility of probing particle interactions in the early universe is known as *cosmological collider physics*.

While this program could provide a promising window into the inflationary era, there are many challenges to navigate. First, Gaussianities significantly dominate the observables in the CMB and large-scale structure, to one part in 10^5, implying weak couplings between SM particles and inflationary fields. Thus far, no non-Gaussianities (NG) have been detected, and any measurements hoping to observe these signals must be performed

with an as-yet achieved level of precision. The target NG can also be produced by several sources, namely primordial NG generated during the inflationary era as well as those originating from the non-linear growth of perturbations post-inflation. Thus, our goal is to identify target models that can generate NG that are distinct and large enough to detect. The simplest instance of such models include direct couplings between the inflaton and some particle with a mass $m \sim H$. We explicitly decompose the inflaton field into its classical and and quantum parts,

$$\phi(t, \mathbf{x}) = \xi(t, \mathbf{x}) + \phi_0(t). \tag{1.279}$$

The curvature perturbations can then be written in terms of the interacting component of the inflaton field, $\zeta = \xi/\phi_0$, and the object of interest is the three-point function as $\langle \xi(\mathbf{k}_1)\xi(\mathbf{k}_2)\xi(\mathbf{k}_3)\rangle$. Such interactions can come from tree-level diagrams, With

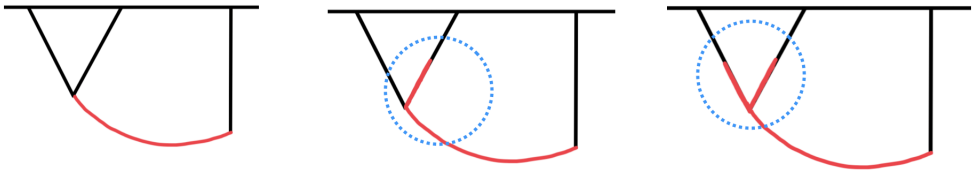

Figure 1.4: The single, double, and triple exchange diagrams, respectively. Note the inclusion of a mixed propagator.

$k_3 \ll k_1, k_2$, this leads to the operator product expansion (OPE) regime, in which the spatial coordinates obey the structure $|\mathbf{x}_1 - \mathbf{x}_2|/ll|\mathbf{x}_1 - \mathbf{x}_3|$. In this sense, the diagrams above "factorize" into two sub-diagrams, In the limit that $k_3 \to 0$, the soft mode of this sub-diagram is approximately constant over distances $|\mathbf{x}_1 - \mathbf{x}_2$. We can then reorganize our diagrams to contain one "effective vertex," which serves to vastly simplify our calculations. Our job, then, is to identify models that include the inflationary couplings that can be encoded in f_{NL}. The computation of the diagrams then proceeds in the familiar way, with the caveat that our spacetime is now dS. Computing amplitudes in dS

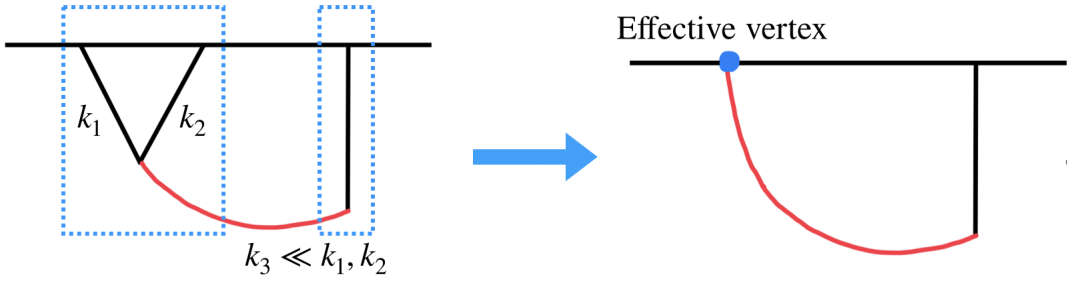

Figure 1.5: By evaluating these diagrams in the squeezed limit, the computation can be split into two sub-diagrams. Computing the diagram on the left leads to an effective vertex for this diagram.

is notoriously difficult, as lack of knowledge of the outgoing states and their evolution invalidates the usual "in-out" formalism of flat-space QFT. This necessitates the use of the "in-in" formalism first introduced by Schwinger.

Starting in the interaction picture, the expectation value of an operator $\mathcal{Q}$ can be written as

$$\langle \mathcal{Q} \rangle = \langle \Omega | U^\dagger \mathcal{Q} U | \Omega \rangle = \langle 0 | \bar{T} e^{+i \int_{-\infty(1-i\epsilon)}^{\eta} \mathcal{H}(\eta') dt'} \mathcal{Q}(\eta) T e^{-i \int_{-\infty(1+i\epsilon)}^{\eta} H(\eta'') dt''} | 0 \rangle. \tag{1.280}$$

This is known as the "master equation" for cosmological collider physics. This equation sums over our ignorance of the outgoing states by computing the "in-in" state: evolving the operator forward and then backward in time. This can bee seen via the use of the time ordering and anti-time ordering operators T and $\bar{T}$, respectively. $\mathcal{H}$ is understood to be the Hamiltonian in the interaction picture, and we note that the time evolution occurs with respect to two bulk times η' and η'', with the final answer evaluated at a time η_0 on the boundary.

The cosmological collider can be considered in analogy with an particle accelerator like the LHC, in which particle collisions lead to signals like leptons or jets, which in turn leave a map of energy deposition on a calorimeter. A cosmological "collider", on the other hand, uses inflationary expansion in the early universe as the mechanism for colliding particles.

These then leave primordial density perturbations which may manifest in cosmological datasets such as the CMB, the 21-cm line, and maps of large-scale structure. Crucially, because the scale of Hubble during inflation may be as high as 10^{13} GeV, this could open up a window into exploration of physics at extremely high energies, far above what any of our current experiments can probe.

1.4.4 Where is Gravity? (and Other Concerns)

There are many more puzzles and questions that lead us to expect the existence of new physics in addition to the ones already discussed. While not by any means a comprehensive list, we will give an overview of a few more problems that suggest that the SM and ΛCDM are by themselves insufficient to capture the phenomena of particle physics and cosmology.

Where is gravity? The most glaring shortcoming of the SM is that it does not include gravity. Quantum gravity is expected to kick in at the Planck scale, and the development of this field is an active area of research in high energy theory. While there have been many attempts to quantize gravity — the most successful being, arguably, string theory — we have no testable predictions to verify a theory of quantum gravity.

What is dark energy? While the ΛCDM model does an excellent job of fitting with data, it does not say what the Λ (or, for that matter, the CDM) actually is.

What is the neutrino mass mechanism? In the SM, neutrinos are massless, but everybody knows that this cannot be the case. The discovery of neutrino oscillations — in which neutrinos change between the electron, muon, and tau flavors they travel— showed that neutrinos must have small, nonzero masses. This is because oscillation is only

possible if the mass eigenstates of neutrinos are different from their flavor eigenstates, which requires the mass eigenstates to have different, nonzero masses. This phenomenon requires an extension of the SM, as it cannot account for neutrino masses. Possible explanations include the seesaw mechanism and the introduction of sterile neutrinos, though the exact origin of neutrino mass is still unknown.

What is the source of the baryon asymmetry? The baryon asymmetry problem refers to the observed imbalance between matter and antimatter in the universe. According to the SM and Big Bang theory, equal amounts of matter and antimatter should have been produced in the early universe, leading to their mutual annihilation. However, the universe is dominated by matter. Explaining this requires mechanisms that violate CP symmetry, baryon number, and out-of-equilibrium processes — the *Sakharov conditions* — but the Standard Model's CP violation is too small to account for the observed asymmetry, suggesting new physics is needed.

Leptogenesis is one mechanism that explains the baryon asymmetry of the universe by first generating an asymmetry in the lepton sector. This occurs through the decay of heavy right-handed neutrinos, which violate lepton number and CP symmetry. The lepton asymmetry is then converted into a baryon asymmetry through electroweak sphaleron processes, which violate baryon and lepton numbers but conserve their difference.

What is the nature of the Higgs? While we've confirmed the existence of the Higgs and measured its mass, questions remain about its exact nature, such as whether it is an elementary particle or part of a more complex system (e.g., part of a composite or extended Higgs sector).

What is the origin of the hierarchical structure of the flavor sector? The flavor sector contains the majority of the SM's free parameters, including the quark and lepton

masses, quark mixing angles, and a CP-violating phase. Giving neutrinos mass in order to align with experiment requires still more parameters. The question of the origin of these parameters and their hierarchical structure comprises the flavor puzzle.

Importantly, although the mass parameters of the flavor sector display significant hierarchies, these parameters are 't Hooft natural. This is the SM is invariant under a global flavor symmetry when the Yukawa couplings are set to zero. The flavor problem is not severe in the sense that it is not a naturalness problem, but the expectation of physics beyond the SM motivates a search for its explanation.

Chapter 2

P not PQ

2.1 Introduction

The current upper bound on the size of the neutron electric dipole moment (EDM) is $|d_n| < 1.8 \cdot 10^{-26}\ e \cdot \text{cm}$ [18].[1] In turn, this severely constrains the size of the QCD vacuum angle, which is required to be

$$\bar{\theta} = \theta_s + \theta_q \lesssim 10^{-10}\ , \tag{2.1}$$

where θ_q is the argument of the determinant of the quark mass matrix, and θ_s the coefficient of the $G\tilde{G}$ operator,

$$\mathcal{L} \supset \frac{\theta_s \alpha_s}{4\pi} \text{tr}\left(G^a \tilde{G}^a\right)\ . \tag{2.2}$$

In the Standard Model (SM), $\theta_q = \arg\det(y_u y_d)$, with $y_{u,d}$ the Yukawa matrices in the up- and down-quark sectors. $\bar{\theta}$ provides a physical, basis-independent measurement of CP-violation in the strong sector of the SM.

That $\bar{\theta}$ is constrained to be so tiny is one of the most puzzling features of the SM, and it is known as the strong CP problem. It stands alongside the electroweak hierarchy problem

[1] Here we quote the direct limit; the inferred bound $|d_n| < 1.6 \cdot 10^{-26}\ e \cdot \text{cm}$ from the ^{199}Hg EDM limit [19] is comparable assuming no additional contributions to the atomic EDM.

and the cosmological constant problem as one of the three great naturalness puzzles that remain unsolved. Although numerically the strong CP problem is orders of magnitude less severe than either of its siblings, it is considerably more robust against anthropic arguments.[2] As such, it has drawn renewed attention during an era in which LHC null results are challenging naturalness-based approaches to the electroweak hierarchy.

Although one could argue that $\theta_s = 0$ on the basis that QCD interactions otherwise preserve CP, a similar argument cannot be made for a vanishing θ_q. For example, if CP were a good symmetry of the Yukawa sector then the Yukawa matrices would need to be real. However, real Yukawas would lead to a vanishing phase in the CKM matrix, in direct conflict with the $\mathcal{O}(1)$ CP-violation observed in the electroweak sector of the SM. Besides CP, a non-zero $\bar{\theta}$ also violates P. Again, the fact that the strong sector preserves parity may allow us to impose $\theta_s = 0$. P invariance in the Yukawa sector would require the Yukawa matrices to be hermitian, in which case $\theta_q = 0$ too, while still allowing for a non-zero CKM phase. However, the fact that P is maximally violated by the electroweak interactions severely weakens this line of reasoning as an attempt to argue for a small $\bar{\theta}$.

So although $\bar{\theta}$ is a measurement of both P and CP violation by strong dynamics, the above discussion highlights how the origin of the strong CP problem in the SM ultimately lies in the features of the *electroweak* sector. It is the fact that electroweak interactions maximally violate both P and CP that precludes an understanding of the bound in Eq.(2.1) based on the underlying symmetries of the SM.

With this in mind, it is natural to attempt an understanding of the smallness of $\bar{\theta}$ in the context of theories with an extended electroweak sector. If either P or CP are good symmetries of the extended theory, then $\bar{\theta}$ will be forced to vanish. Of course, to account for the P and CP violation we observe in nature, they must eventually be broken, and a non-zero $\bar{\theta}$ will be radiatively generated. If the induced $\bar{\theta}$ is small

[2] See, however, [20] for arguments to the contrary.

enough, this class of theories offer a symmetry-based solution to the strong CP problem. Concrete implementations of this idea based on spontaneously broken P and CP were first proposed in [21–23] (building on [24]) and [25, 26] respectively. It is the former that will be the focus of this work.[3]

There is another good reason to consider solutions to strong CP based on the restoration of spacetime symmetries, namely that these may be realized as gauge symmetries in the context of string theory [30, 31]. As such, they can only be broken spontaneously and not explicitly. Depending on the scale of spontaneous symmetry breaking, the apparent lack of P and CP violation in the strong sector could therefore be fully, or partially, explained in this context. Clearly, a resolution to the strong CP problem along these lines would be especially attractive: it would allow us to understand the smallness of $\bar{\theta}$ as an accident resulting from the underlying structure of the UV-completion, as opposed to being the result of a model-building effort specially designed to address Eq.(2.1).

From the bottom-up, there are a number of ways the SM can be extended to accommodate spontaneously broken P. However, in order to address the strong CP problem, a necessary feature of all of them is the presence of an $SU(2)_R$ gauge factor, as well as an extended matter content that mirrors that of the SM. Crucially, the $SU(3)$ quantum numbers of SM fermions and their mirror counterparts must be the same, making the presence of additional colored particles an irreducible feature of these models. With this extended field content, parity enforces the Yukawa couplings in the two sectors to be identical. To be phenomenologically acceptable, parity must be broken at some scale v' above the weak scale, with the additional gauge bosons and mirror quarks being sufficiently heavy to evade experimental constraints. Naïvely, bounds on the mass of colored particles would seemingly require $y_u v' \gtrsim 1$ TeV [32, 33], in turn setting a lower bound $v' \gtrsim 10^8$ GeV. But a parametric separation of scales between v and v' entails an irre-

[3] For recent exploration along these lines, see also [27–29].

ducible amount of fine-tuning $\Delta^{-1} \simeq 2v^2/v'^2$, which would become $\Delta^{-1} \lesssim 10^{-12}$ for such a stringent bound on v'. Considering that the goal is to naturally explain a number of $\mathcal{O}(10^{-10})$, parity would hardly seem to remain an attractive solution to strong CP.

In this paper, we show that the conclusion of the previous paragraph is premature, and that a parity-breaking scale as low as 18 TeV is consistent with all experimental constraints. This significantly improves the level of fine-tuning, and leaves an open window for symmetry-based solutions to strong CP that are based on spontaneously broken parity. The leading constraint on the low-tuning version of these models comes not from bounds on colored particles, but from direct searches for Z' and W' resonances at the LHC [34,35]. Future searches for heavy gauge bosons at current and future colliders are the most promising probes of this class of theories, with a 100 TeV proton collider guaranteed to make a discovery if the level of fine-tuning is better than $\Delta^{-1} \sim 10^{-5}$ [36, 37]. Overall, the viability of these parity-based models makes collider experiments a central testing ground for solutions to strong CP.

Another attractive feature of this class of solutions to the strong CP problem is that they are robust against the effects of symmetry-breaking higher dimensional operators (HDOs) that may arise from short-distance physics associated with a gravitational UV completion. If parity is a gauge symmetry of the underlying theory, we are led to consider only those HDOs proportional to the source of spontaneous symmetry breaking. On the other hand, if parity were global, the expectation that quantum gravity violates all global symmetries [38–50] suggests we should include all HDOs that explicitly violate P. Although the nature of the operators is different in the gauge and global implementations, the conclusion will be the same: in both cases, the leading HDOs with $\mathcal{O}(1)$ coefficients may be present without destabilizing the solution to strong CP.

This stands in stark contrast with the reality of what has traditionally been the most popular solution to the strong CP problem: the QCD axion [51–57]. In this case, the

parameter $\bar{\theta}$ is promoted to the status of dynamical field, the axion, which is a pseudo-Nambu-Goldstone boson of a spontaneously broken $U(1)_{PQ}$ *global* symmetry. A potential for the axion is induced non-perturbatively by QCD dynamics, and its vacuum expectation value (vev) adjusts such that $\bar{\theta} = 0$, thereby solving strong CP. To work, the QCD axion potential must dominate to 1 part in 10^{10}, overwhelming any other contributions that may arise from additional degrees of freedom. New dynamics responsible for, say, dark matter, baryogenesis, or addressing the hierarchy problem, cannot significantly contribute to the axion potential. Similarly, Planck-suppressed HDOs that break $U(1)_{PQ}$ must be exceptionally suppressed [58–61]. The mechanism is not robust. The need for $U(1)_{PQ}$ to be a high quality global symmetry has become known as the "axion quality problem". Attempts to turn the QCD axion into a high quality axion are valuable [62–69], but hardly helpful in making a small $\bar{\theta}$ appear natural.

The goal of this work is to identify the most natural parity-based solution to the strong CP problem, and highlight its experimental consequences. We do so by following a strategy that combines the traditional notion of naturalness with the expectation that gravity violates all global symmetries. The former singles out a specific implementation of the spectrum of parity-symmetric models, and underscores the central role of collider experiments in exploring solutions to strong CP. The latter opens up an entirely new avenue of exploration for parity solutions to the strong CP problem, ranging from EDM experiments to gravitational wave observatories, depending on the degree to which the symmetry remains approximate.

This chapter is organized as follows. In section 2.2 we review the main features of parity-based solutions to strong CP, and discuss how a low symmetry breaking scale can be realized while complying with experimental constraints. We focus on the main phenomenological signatures of these models that are relevant for collider and flavor experiments in section 2.3. In section 2.4, we discuss the size of radiative corrections

to both $\bar{\theta}$ and the EDM of elementary fermions, including charged leptons, depending on the details of the parity-breaking sector. We explore the effect of Planck-suppressed HDOs on this class of models in section 2.5, paying special attention to a potential gravitational wave signal from the spontaneous breaking of parity. Section 2.6 contains our conclusions. Finally, a series of appendices contain results that have been crucial in our analysis, but may be skipped on a first reading of the manuscript.

2.2 P to solve strong CP

In this section, we introduce the main features of symmetry-based solutions to the strong CP problem based on parity. In 2.2.1 we review the basic idea, as first introduced in [21–23]. We focus on the scalar potential in 2.2.2, with an emphasis on the implications for fine-tuning of the weak scale that arise as a result of the breaking of parity. In 2.2.3 we discuss how the scale of additional colored particles can be decoupled from the parity-breaking scale, in turn minimizing the level of fine-tuning.

2.2.1 Parity as a solution to the strong CP problem

A symmetry-based solution to the strong CP problem based on parity requires extending the SM both in terms of matter content and gauge interactions. The minimal implementation of this idea is based on the gauge group

$$SU(3) \times SU(2)_L \times SU(2)_R \times U(1)_{\hat{Y}}, \tag{2.3}$$

as well as a doubling of the matter content of the SM into a 'mirror' sector with identical quantum numbers, except that $SU(2)_L$ doublets are now doublets of $SU(2)_R$. Table 2.1 summarizes the gauge charges in the quark and Higgs sectors of the theory. (Analogous charge assignments apply in the lepton sector, which we don't make explicit.) Crucially,

the Higgs sector of the theory does not introduce additional sources of CP-violation — indeed, the freedom to perform both $SU(2)_L$ and $SU(2)_R$ gauge transformations allows us to expand around a vacuum where both vevs are real. [4]

	$Q = \begin{pmatrix} u \\ d \end{pmatrix}$	$U^\dagger$	$D^\dagger$	H	$Q'^\dagger = \begin{pmatrix} u'^\dagger \\ d'^\dagger \end{pmatrix}$	U'	D'	H'^*
$SU(3)$	**3**	**3**	**3**	·	**3**	**3**	**3**	·
$SU(2)_L$	**2**	·	·	**2**	·	·	·	·
$SU(2)_R$	·	·	·	·	**2**	·	·	**2**
$U(1)_{\hat{Y}}$	$\frac{1}{6}$	$\frac{2}{3}$	$-\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{6}$	$\frac{2}{3}$	$-\frac{1}{3}$	$\frac{1}{2}$

Table 2.1: Quantum numbers in the quark and Higgs sectors. Mirror sector fields are distinguished with a prime. We use notation such that all of Q, U, and D (as well as their mirror counterparts Q', U', and D') are left-handed, two-component Weyl fermions, whereas daggered fields are always right-handed.

With this additional field content, the theory admits an alternative definition of parity that combines the action of the 'ordinary' parity transformation with an internal symmetry that exchanges the fields of the SM and mirror sectors. Explicitly, in the gauge, quark, and Higgs sectors:

$$\mathbf{W}_L^\mu \leftrightarrow \mathbf{W}_{R\mu}, \tag{2.4}$$

$$Q, U, D \leftrightarrow Q'^\dagger, U'^\dagger, D'^\dagger, \tag{2.5}$$

$$H \leftrightarrow H'^*, \tag{2.6}$$

[4] The model we have just introduced is *not* the minimal parity-symmetric extension of the SM, but rather the minimal extension that solves the strong CP problem. The most minimal extension of the SM that admits a generalized parity symmetry was introduced in [70], and does not require a doubling of the fermion sector. The scalar sector, however, requires the addition of an additional Higgs field transforming as a bifundamental of the two $SU(2)$ factors, in order to write Yukawa couplings in a gauge invariant fashion. As discussed in [23], the vev of this field will in general be complex, precluding a vanishing θ_q.

and similarly for leptons. Since $SU(3)$ and $U(1)_{\hat{Y}}$ interactions are not mirrored, the corresponding gauge fields transform as usual under parity. Unlike ordinary parity in the SM, this 'generalized' parity transformation is now a good symmetry of the gauge sector of the theory, thanks to the extended electroweak sector and matter content.

In this context, the strong CP problem is solved as follows. On the one hand, parity requires that $\theta_s = 0$, just as one may argue in the SM based on the properties of the strong sector alone. On the other hand, the presence of additional colored particles results in an extended quark mass matrix. In particular, Yukawa terms can be written for both the SM and mirror sectors, of the form

$$\mathcal{L} \supset -\left\{(y_u)_{ij} Q_i H U_j + (y'_u)_{ij} Q'^{\dagger}_i H'^{*} U'^{\dagger}_j\right\} + \text{h.c.}, \tag{2.7}$$

and similarly for down-type quarks and leptons. As a result, the tree-level value of θ_q in these models is given by

$$\theta_q = \arg\,\det(y_u y_d) + \arg\,\det(y'^{*}_u y'^{*}_d). \tag{2.8}$$

Crucially, demanding that Yukawa interactions preserve parity, which is now a good symmetry of the extended electroweak sector, enforces the Yukawa couplings in the two sectors to be identical, i.e.

$$y'_f = y_f. \tag{2.9}$$

In turn, this implies $\theta_q = 0$, as per Eq.(2.8), forcing $\bar{\theta}$ to vanish at tree-level in parity-symmetric models.

With the field content outlined in table 2.1, the theory admits an additional fermion mass term involving only the $SU(2)$-singlets, of the form

$$\mathcal{L} \supset -(\mathcal{M}_u)_{ij} U_i U'_j + \text{h.c.} \tag{2.10}$$

(with analogous terms for down-type quarks and leptons), where invariance under generalized parity requires that the vector-like mass matrix be hermitian, i.e. $\mathcal{M}_f^\dagger = \mathcal{M}_f$.[5] In general, the expression for θ_q can be conveniently written as

$$\theta_q = \arg\det(\mathbb{M}_u \mathbb{M}_d), \tag{2.11}$$

where $\mathbb{M}_u$ and $\mathbb{M}_d$ are 6×6 matrices, of the form

$$\mathbb{M}_f = \begin{pmatrix} \not{\vdash} & \frac{v'}{\sqrt{2}} y_f'^* \\ \frac{v}{\sqrt{2}} y_f^T & \mathcal{M}_f \end{pmatrix}, \qquad \text{for} \qquad f = u, d. \tag{2.12}$$

Due to the zero in the upper-left block of the overall 6×6 mass matrix, the expression for θ_q remains as in Eq.(2.8). As we will discuss in 2.2.3, the presence of vector-like masses is crucial in implementing a version of the model with low fine-tuning.[6]

To obtain a phenomenologically viable model, parity must be broken, with different vevs in the mirror and SM Higgs sectors. This will induce a non-vanishing $\bar{\theta}$ beyond tree-level, which must be small enough if the theory is to remain a bona-fide solution to strong CP. The size of radiative corrections depends on the details of how parity is spontaneously broken. If P is broken without breaking CP, then the radiatively induced $\bar{\theta}$ will be no larger than in the SM [23], where $\bar{\theta} < 10^{-19}$ [71]. On the other hand, if CP is also spontaneously broken (e.g. through the vev of a pseudo-scalar) then a larger $\bar{\theta}$, as well as a neutron EDM *independent* of $\bar{\theta}$, may be radiatively generated. Even in this latter case, we will see that radiative corrections can be small enough to remain compatible with experimental constraints. Given that the final size of $\bar{\theta}$ is a somewhat

[5] Note that a non-hermitian mass matrix is compatible with softly broken parity; we will explore the consequences of such soft breaking in section 2.4.1.

[6] A variation on the model we have so far discussed entails extending the gauge group in Eq.(2.3) with an additional $U(1)$, as first discussed in [23]. In this case, SM and mirror fields are charged under different $U(1)$ factors, which transform into each other under parity. Although this seems like a minimal modification of the model presented here, this two-$U(1)$ version does not allow for the vector-like mass terms of Eq.(2.10), in turn precluding the implementation of a low parity-breaking scale.

model-dependent feature of this class of models, we defer a more detailed discussion of this issue to section 2.4.

More generally, discussing the leading effect of broken parity on the fine-tuning of the electroweak sector does not require committing to a specific implementation of spontaneous symmetry breaking. For this purpose, it will be enough to focus on the features of the Higgs sector, to which we now turn.

2.2.2 Scalar sector and fine-tuning

For the time being, we will parametrize the necessary breaking of parity through an explicit soft term in the scalar potential. Of course, such soft breaking should ultimately be the result of some spontaneous symmetry breaking dynamics, as we will make more explicit in section 2.4. In this spirit, the most general scalar potential involving the $SU(2)_L$ and $SU(2)_R$ Higgs doublets takes the form

$$V(H,H') = -m_H^2(|H|^2+|H'|^2) + \lambda(|H|^2+|H'|^2)^2 + \kappa(|H|^4+|H'|^4) + \mu^2|H|^2. \quad (2.13)$$

At this level, Eq.(2.13) is identical to the scalar potential of theories of Neutral Naturalness, such as Twin Higgs [72, 73]. The first two terms respect both parity and a larger accidental $SU(4)$ (really, $O(8)$) symmetry, while κ respects the former but not the latter. The parameter μ^2 softly breaks parity. In the interest of a non-trivial vacuum structure, we take $m_H^2 > 0$. Depending on the relative signs and sizes of the quartic couplings, the tree-level vacua for $\mu^2 = 0$ either preserve parity (with $v' = v$) or spontaneously break parity (with $v \neq 0, v' = 0$ or $v = 0, v' \neq 0$). A vacuum with $v' \gg v \neq 0$ may be obtained by deforming the theory away from the parity-symmetric vacuum with nonzero μ^2 and $\lambda, \kappa > 0$. At tree-level, the vevs in the SM and mirror Higgs sectors are then given by

$$v^2 = \frac{m_H^2 - \mu^2(1+\lambda/\kappa)}{2\lambda+\kappa}, \qquad \text{and} \qquad v'^2 = \frac{m_H^2 + \mu^2\lambda/\kappa}{2\lambda+\kappa}, \quad (2.14)$$

where $v^2 \ll v'^2$ is necessary in order to obtain a phenomenologically viable model. After spontaneous symmetry breaking, the spectrum of the theory contains two scalar fields, h and h', with masses $m_h \simeq 2\sqrt{\kappa}v$ and $m_{h'} \simeq \sqrt{2\lambda}v'$ respectively, as well as six Goldstones that become the longitudinal components of the gauge bosons of our extended electroweak sector. The physical gauge boson spectrum contains Z' and W' resonances, which are heavier than their SM counterparts by a factor of v'/v. We defer further details of the scalar and gauge sectors to appendix A.1.1.

It is clear from Eq.(2.14) that to obtain a hierarchy of scales between v and v' we need to introduce a tree-level tuning between the parity-preserving and parity-breaking mass-squared terms. Heuristically, the necessary fine-tuning is given by

$$\Delta^{-1} \equiv \frac{m_H^2 - \mu^2(1+\lambda/\kappa)}{m_H^2} \simeq \frac{(2\lambda+\kappa)v^2}{(\lambda+\kappa)v'^2} \simeq \frac{2v^2}{v'^2}. \tag{2.15}$$

This is an irreducible contribution to the fine-tuning in this class of models. Insofar as it involves the sensitivity of the weak scale v to underlying parameters, it may be classified as a tuning associated with the electroweak hierarchy problem, although it is not necessarily the only such contribution. For example, a hierarchy of scales $v'^2 \ll M_{Pl}^2$ would constitute an additional source of fine-tuning in the absence of a stabilizing mechanism. Similarly, additional hierarchies of scales or couplings in the sector responsible for spontaneously breaking P might necessitate similar accurate cancellations. However, these are issues that could, at least *in principle*, be addressed at some higher scale above v', provided the necessary dynamics do not spoil the smallness of $\bar{\theta}$ [74]. In contrast, Eq.(2.15) is forced on us independently of the UV-completion. Although it is tempting to attach the tuning in Eq.(2.15) to the electroweak hierarchy problem and attribute it to anthropic selection (the perspective advocated in e.g. [28, 29]), this necessarily entails some favorable assumptions about the properties of an anthropic landscape. Here we prefer to render unto strong CP the things that are strong CP's, and take the irreducible

tuning in Eq.(2.15) at face value as a measure of the degree to which a parity model naturally explains the small value of $\bar{\theta}$ without reintroducing tuning elsewhere.

With this in mind, in this paper we focus on implementations of parity solutions to strong CP where the level of fine-tuning, as parametrized in Eq.(2.15), is as mild as possible. This will concentrate our attention on a specific mechanism to generate fermion masses that in turn endows these models with characteristic phenomenology, as we discuss next.

2.2.3 Fermion masses and a low parity-breaking scale

Fermion mass terms arise from the Yukawa couplings of Eq.(2.7), as well as from the vector-like mass involving the $SU(2)$-singlets. In total:

$$\mathcal{L} \supset -\left\{ \frac{v}{\sqrt{2}}(y_u)_{ij} u_i U_j + \frac{v'}{\sqrt{2}}(y_u)^*_{ij} u'_i U'_j + (\mathcal{M}_u)_{ij} U_i U'_j \right\} + \text{h.c.}, \tag{2.16}$$

where we have already set $y'_u = y_u$, as mandated by generalized parity, and analogous mass terms are present both for down-type quarks and leptons.

The structure of Eq.(2.16) allows for two limiting realizations of the fermion spectrum. If the overall scale of the vector-like mass matrix is $M \ll v, v'$, then fermion masses are generated mainly through the Yukawa terms, as in the SM. In this case, mirror fermions would be an exact copy of the SM, just heavier by a factor of v'/v. Demanding that the lightest mirror quark is heavy enough to comply with current experimental constraints requires $m_u \times v'/v \gtrsim 1$ TeV [32, 33], in turn setting a lower bound $v' \gtrsim 10^8$ GeV. As advertised in the Introduction, this sets the level of fine-tuning in the electroweak sector to $\Delta^{-1} \lesssim 10^{-12}$. The phenomenology of parity solutions to strong CP in this regime was discussed recently in [28].

On the other hand, the limit $M \gg v, v'$, allows for a see-saw realization of the fermion spectrum, consisting of three light (SM-like) fermions, and three heavy fermions with

mass of order M. A sufficiently high scale for the mass of additional colored particles can now be achieved by increasing M, not v'. This allows for a much lower parity-breaking scale, and therefore a much better level of fine-tuning. See-saw implementations of fermion masses, for both quarks and leptons, are discussed in [75–78], and it was in fact in this context that a parity-based solution to the strong CP problem was first proposed [21,22]. It is this second realization of the fermion spectrum that we concentrate on in this work. [7]

The up-quark sector requires special consideration, since the see-saw mechanism cannot be applied to the top quark while maintaining perturbative Yukawas. So let us discuss the down-quark and lepton sectors first. (We will use notation appropriate to the down-quark sector, but emphasize that the same results apply for leptons.) To leading order in both v/M and v'/M, the masses of the light and heavy fermions are obtained by diagonalizing the 3×3 hermitian matrices

$$\frac{vv'}{2} y_d^* \mathcal{M}_d^{-1} y_d^T, \qquad \text{and} \qquad \mathcal{M}_d, \tag{2.17}$$

respectively. We make the simplifying assumption that there are no significant hierarchies in the eigenvalues of $\mathcal{M}_d$, and therefore the heavy quarks appear at a common scale $\sim M$. Parametrically, light quark masses are then of the form $m_{d_i} \sim |y|^2 vv'/M$. The see-saw mechanism generates fermion masses $m_{d_i} \ll v$ while allowing for much larger Yukawa couplings than in the SM, which is obviously one of the main attractions of this class of models. Generating the b quark mass through the see-saw mechanism while maintaining perturbativity sets an upper bound on the ratio M/v', parametrically:

$$m_b \sim |y|^2 \frac{vv'}{M} \lesssim \frac{vv'}{M} \qquad \Rightarrow \qquad \frac{M}{v'} \lesssim \frac{v}{m_b} \sim 10^2. \tag{2.18}$$

[7] The vector-like masses of Eq.(2.16) provide a soft breaking of generalized parity, precluding the existence of degrees of freedom that would be stabilized by the internal part of this symmetry. As a result, these models do not feature a natural dark matter candidate whose presence is linked to the resolution of the strong CP problem, unlike some of the viable parameter space of the QCD axion.

Rotating from the flavor to the mass eigenbasis in the fermion sector can be conveniently performed step by step at each order in perturbation theory, and we present a detailed discussion of this procedure in appendix A.1.2. At zeroth order in $v^{(\prime)}/M$, it is necessary to perform unitary transformations acting separately on the $SU(2)$-singlet and doublet fields, of the form:

$$d \to \mathcal{O}_d^\dagger d\,,\ d' \to \mathcal{O}_d^T d'\ , \qquad \text{and} \qquad D' \to \mathcal{O}_{D'}^\dagger D'\,,\ D \to \mathcal{O}_{D'}^T D\,. \tag{2.19}$$

$\mathcal{O}_d$ and $\mathcal{O}_{D'}$ are 3×3 unitary matrices acting on flavor space that diagonalize the first and second matrices of Eq.(2.17), respectively. At first order in $v^{(\prime)}/M$, a further rotation is required that mixes the $SU(2)$-singlet and doublet fields as follows:

$$\begin{pmatrix} d \\ D' \end{pmatrix} \to \begin{pmatrix} \nVdash_3 & \epsilon_d^\dagger \\ -\epsilon_d & \nVdash_3 \end{pmatrix} \begin{pmatrix} d \\ D' \end{pmatrix}, \qquad \text{and} \qquad \begin{pmatrix} d' \\ D \end{pmatrix} \to \begin{pmatrix} \nVdash_3 & \epsilon_d'^\dagger \\ -\epsilon_d' & \nVdash_3 \end{pmatrix} \begin{pmatrix} d' \\ D \end{pmatrix}, \tag{2.20}$$

where ϵ_d and ϵ_d' are 3×3 matrices with entries of $\mathcal{O}(v/M)$ and $\mathcal{O}(v'/M)$ respectively, and whose explicit expressions are given in Eq.(A.13).

Using Dirac notation, the left- and right-handed components of the light and heavy mass eigenstates are then given by

$$d_{iL} = \begin{pmatrix} d_i \\ 0 \end{pmatrix},\ d_{iR} = \begin{pmatrix} 0 \\ d_i'^\dagger \end{pmatrix}, \qquad \text{and} \qquad D_{iL} = \begin{pmatrix} D_i' \\ 0 \end{pmatrix},\ D_{iR} = \begin{pmatrix} 0 \\ D_i^\dagger \end{pmatrix}. \tag{2.21}$$

In particular, notice that the right-handed components of the light (SM-like) fermions consist of the corresponding component of the $SU(2)_R$-doublets, up to corrections of $\mathcal{O}(v'/M)$. This feature plays a crucial role in the phenomenology of these models. In particular, it leads to unsuppressed couplings between $SU(2)_R$ gauge bosons, and the right-handed currents of the SM-like fermions. As we will discuss in 2.3.1, this leads to the most stringent bound on the parity-breaking scale.

As far as the up-quark sector is concerned, the see-saw mechanism can be implemented for the u and c quarks, with the corresponding heavy partners appearing at the scale $\sim M$.

The mass eigenstates for the first two generations are as in Eq.(2.21). The top sector, on the other hand, cannot be significantly "see-sawed". Instead, it consists of light and heavy top partners with tree-level masses $m_t \simeq y_t v/\sqrt{2}$ and $m_{t'} \simeq m_t \times v'/v$, respectively. In Dirac notation, and at zeroth order in $v^{(\prime)}/M$, the mass eigenstates are now purely made of SM and mirror sector fields, i.e.

$$t_L = \begin{pmatrix} u_3 \\ 0 \end{pmatrix}, \; t_R = \begin{pmatrix} 0 \\ U_3^\dagger \end{pmatrix}, \qquad \text{and} \qquad t'_L = \begin{pmatrix} U'_3 \\ 0 \end{pmatrix}, \; t'_R = \begin{pmatrix} 0 \\ u_3'^\dagger \end{pmatrix}. \tag{2.22}$$

As usual, rotation matrices in the quark sector are constrained by the requirement that the CKM matrix is reproduced appropriately, which in this case implies $V = \mathcal{O}_u \mathcal{O}_d^\dagger$, up to corrections of $\mathcal{O}(v^2/M^2)$. Further details concerning the mass diagonalization procedure in the fermion sector can be found in appendix A.1.2.

This finalizes our discussion of the main characteristics of parity solutions to strong CP that feature low fine-tuning in the electroweak sector. Before moving on, we include in figure 2.1 a schematic representation of the typical spectrum of these models. Amusingly, the combination of parity and the see-saw mechanism leads to a spectrum of partner particles strikingly reminiscent of a "natural" left-right Twin Higgs model [73, 79] with light top and W/Z partners.

2.3 Dial P for Phenomenology

We now turn to the phenomenology of natural parity solutions to strong CP, beginning with direct bounds from the LHC in section 2.3.1 before turning to indirect flavor constraints in 2.3.2. The collider and flavor phenomenology of similar left-right models has been the topic of previous work [21, 22, 78–82], and our focus here will be on those "irreducible" signatures that are mandated by the structure of the theory in its capacity as a solution to strong CP. A more in depth study of collider and flavor signatures in

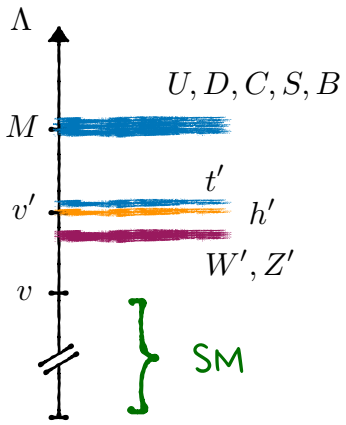

Figure 2.1: Schematic illustration of the particle spectrum of parity solutions to the strong CP problem in their least tuned version. The lightest exotic particles are W' and Z' resonances, followed by a top partner appearing at a scale of order v'. A mirror Higgs is also part of the low-lying spectrum, with $m_{h'} \simeq \sqrt{2\lambda}v'$. (For illustration, we have chosen $\lambda = \mathcal{O}(1)$, but note that h' could be much lighter if $\lambda \ll 1$.) Additional mirror quarks have masses of order the see-saw scale, $M \gg v'$. The lepton sector must also be "see-sawed", with mirror leptons similarly appearing well above v' (although we emphasize that the see-saw scales in the quark and lepton sectors need not coincide).

light of forthcoming data can illuminate the additional structure of these models, and it is a worthwhile direction for continued study.

2.3.1 Collider bounds

The doubling of the electroweak sector gives rise to a plethora of experimental signatures at colliders, ranging from additional vector bosons (the W' and Z' of spontaneously broken $SU(2)_R$) to vector-like quarks (the $SU(2)$-singlet fermions) to additional Higgs bosons. Ultimately, given that the Z' and W' gauge bosons acquire masses exclusively from $SU(2)_R$ breaking, and inherit couplings to the SM-like quarks and leptons, collider searches for these additional vectors place the most solid and strongest direct bounds on the models under consideration.

Neutral currents

The Z' resonance inherits couplings to both the left- and right-handed currents of SM-like fermions. In the down-type quark and lepton sectors, these are both flavor diagonal and generation universal. After rotating to the mass eigenbasis in both the gauge and fermion sectors, as outlined in appendix A.1, we find:

$$\mathcal{L} \supset g Z'_\mu \sum_{i=1}^{3} \left(z'_{d_R} \bar{d}_{iR} \gamma^\mu d_{iR} + z'_{e_R} \bar{e}_{iR} \gamma^\mu e_{iR} + z'_{\nu_R} \bar{\nu}_{iR} \gamma^\mu \nu_{iR} \right) + \{R \to L\} \ . \tag{2.23}$$

As discussed in section 2.2.3, the see-saw implementation of fermion masses leads to unsuppressed couplings between the $SU(2)_R$ gauge bosons and right-handed fermions. Up to corrections of $\mathcal{O}(\sin^2\theta_w)$, these are identical to the couplings between the SM Z and left-handed currents. Specifically:

$$z'_{d_R} \simeq -\frac{g}{2}\left(1 + \mathcal{O}(\mathrm{s}_w^2)\right), \quad z'_{e_R} \simeq -\frac{g}{2}\left(1 + \mathcal{O}(\mathrm{s}_w^2)\right), \quad \text{and} \quad z'_{\nu_R} \simeq \frac{g}{2}\left(1 + \mathcal{O}(\mathrm{s}_w^2)\right), \tag{2.24}$$

where $\mathrm{s}_w \equiv \sin\theta_w$, and we have ignored corrections of $\mathcal{O}(v^2/v'^2)$. On the other hand, couplings of the Z' to left-handed currents are now suppressed:

$$z'_{d_L} \simeq -\frac{g}{6}\frac{\mathrm{s}_w \mathrm{t}_w}{\sqrt{\cos 2\theta_w}} = \mathcal{O}(\mathrm{s}_w^2), \quad \text{and} \quad z'_{e_L} = z'_{\nu_L} = -3 z'_{d_L}. \tag{2.25}$$

The situation in the up-quark sector is somewhat different. Now, couplings between the Z' and the right-handed fermion currents are no longer universal. Instead, we have:

$$\mathcal{L} \supset g Z'_\mu \left(\sum_{i=1}^{3} z'_{u_L} \bar{u}_{iL} \gamma^\mu u_{iL} + \sum_{i=1}^{2} z'_{u_R} \bar{u}_{iR} \gamma^\mu u_{iR} + z'_{t_R} \bar{t}_R \gamma^\mu t_R \right). \tag{2.26}$$

As before, Z' couplings to first and second generation right-handed currents are unsuppressed, and are given by

$$z'_{u_R} \simeq \frac{g}{2}\left(1 + \mathcal{O}(\mathrm{s}_w^2)\right) \ , \tag{2.27}$$

whereas those to left-handed fermions, as well as to the right-handed top, now read

$$z'_{u_L} = \frac{z'_{t_R}}{4} \simeq -\frac{g}{6}\frac{\mathrm{s}_w \mathrm{t}_w}{\sqrt{\cos 2\theta_w}} = \mathcal{O}(\mathrm{s}_w^2) \ . \tag{2.28}$$

Bounds on the Z' mass from its production at the LHC will therefore be similar to those found in the so-called Sequential Standard Model, which features a Z' resonance that is just a heavy copy of the SM Z. In the present model, couplings of the Z' to SM fermions are similar to those of the Z after the replacement $L \leftrightarrow R$ — a replacement that does not affect either the production cross section or the decay rates into light fermions. The most constraining limits thus come from [34], where a search focused on leptonic final states sets a lower bound $m_{Z'} \gtrsim 5$ TeV. In turn, this translates into a lower limit on the scale of parity breaking of order $v' \gtrsim 13$ TeV.

Charged currents

W' gauge bosons interact with the right-handed SM fermions in a way that mirrors the interactions between their left-handed counterparts and SM W. In the lepton sector:

$$\mathcal{L} \supset \frac{g}{\sqrt{2}} \sum_{i,j=1}^{3} \left(B_{ij} W_\mu^+ \bar{\nu}_{iL} \gamma^\mu e_{jL} + B'_{ij} W_\mu'^+ \bar{\nu}_{iR} \gamma^\mu e_{jR} \right) + \text{h.c.}, \tag{2.29}$$

where $B = B' = \mathcal{O}_\nu \mathcal{O}_e^\dagger$, up to corrections of order ${v^{(\prime)}}^2/M^2$. As far as the quark sector is concerned, the up-type sector again requires special consideration. We find:

$$\mathcal{L} \supset \frac{g}{\sqrt{2}} \sum_{i,j=1}^{3} W_\mu^+ V_{ij} \bar{u}_{iL} \gamma^\mu d_{jL} + \text{h.c.}, \tag{2.30}$$

with $V = \mathcal{O}_u \mathcal{O}_d^\dagger + \mathcal{O}(v^2/M^2)$, whereas

$$\mathcal{L} \supset \frac{g}{\sqrt{2}} W_\mu'^+ \sum_{j=1}^{3} \left(\sum_{i=1}^{2} V'_{ij} \bar{u}_{iR} \gamma^\mu d_{jR} + \Delta V'_{3j} \bar{t}_R \gamma^\mu d_{jR} \right) + \text{h.c.} \tag{2.31}$$

Up to corrections of $\mathcal{O}(v'^2/M^2)$, we have $V' = V$, and $\Delta V'_{3j} = (\epsilon'^*_u V)_{3j}$. The 3×3 matrix ϵ'_u, whose entries are suppressed by a factor of $v'/M \ll 1$, is given explicitly in Eq.(A.19).

As with the Z', we expect bounds on the W' to be comparable to those in the Sequential Standard Model. Current direct searches set stringent constraints on such W' resonances, of order $m_{W'} \gtrsim 6$ TeV [35]. In turn, this sets the strongest limit on the scale

of parity breaking: $v' \gtrsim 18$ TeV. Although direct searches for vector-like quarks and additional Higgs bosons are also germane, they lead to significantly weaker bounds on the scale of parity breaking compared to W' and Z' searches. For example, null results in searches for vector-like top partners [32, 33] lead to $v' \gtrsim 2$ TeV, with comparable bounds coming from searches for SM-singlet scalars.

Looking to the future, a 100 TeV pp collider such as the proposed FCC-hh should be sensitive to W' and Z' bosons as heavy as ~ 40 TeV [36, 37], corresponding to $v' \gtrsim 120$ TeV. This would comprehensively cover the most natural parameter space consistent with current data, and the non-observation of heavy vectors at such a collider would suggest that parity solutions are tuned at the $\Delta^{-1} \sim 10^{-5}$ level. In this respect, future colliders provide a decisive test of parity solutions to the strong CP problem.

2.3.2 Flavor constraints

In the SM, flavor-changing neutral currents (FCNCs) are absent at tree-level, appearing only at one-loop, and being additionally suppressed by the GIM mechanism. As a result, precision measurements of flavor-violating processes often imply stringent constraints on extensions of the SM. In the class of models under consideration, FCNCs arise already at tree-level, mediated by the Z and Z' gauge bosons, as well as the scalars h and h'. However, their size is suppressed by factors of the Yukawa couplings of the relevant fermions, making their effect negligible. At one-loop, FCNCs proceeding via box diagrams involving W' gauge bosons and mirror up-type quarks can lead to deviations in kaon properties, in turn setting the leading constraints on the flavor structure of these models.

Tree-level FCNCs

Rotating from the gauge to the mass eigenbasis in the fermion and gauge boson sectors, as specified in appendix A.1, leads to the presence of flavor-changing interactions between the Z and the SM-like fermions. For example, in the down-quark sector there are new interactions of the form

$$\mathcal{L} \supset \frac{g}{2c_w}(\epsilon_d^\dagger \epsilon_d)_{ij} Z_\mu \bar{d}_{iL} \gamma^\mu d_{jL}, \tag{2.32}$$

where ϵ_d is a 3×3 matrix acting on flavor space whose explicit form is given in Eq.(A.13). Integrating out the Z, the effective hamiltonian relevant to describe $|\Delta F| = 1$ processes, such as the leptonic decay of B mesons, now contains additional terms, of the form

$$\Delta\mathcal{H}_{\text{eff}} \simeq -\sqrt{2} G_F (\epsilon_d^\dagger \epsilon_d)_{32} \cos(2\theta_w)(\bar{b}_L \gamma^\mu s_L)(\bar{\mu}_L \gamma_\mu \mu_L) + \text{h.c.} \tag{2.33}$$

(An analogous term involving right-handed muons is also present, but suppressed by a factor of s_w^2, so we neglect it in the subsequent discussion.)

The deviation with respect to the SM prediction for the branching fraction of the process $B_s^0 \to \mu^+\mu^-$ as a result of the operator in Eq.(2.33) can be written as

$$r_{\mu\mu} \equiv \frac{\text{BR}(B_s^0 \to \mu^+\mu^-)_{\text{BSM}}}{\text{BR}(B_s^0 \to \mu^+\mu^-)_{\text{SM}}} - 1 \simeq \frac{|C_{10}^{(\text{SM})} + C_{10}^{(\text{BSM})}|^2}{|C_{10}^{(\text{SM})}|^2} - 1, \tag{2.34}$$

where $C_{10}^{(\text{SM})}$ and $C_{10}^{(\text{BSM})}$ are the SM and BSM contributions to the Wilson coefficient of the four-fermion operator $(\bar{b}_L \gamma^\mu s_L)(\bar{\mu}\gamma_\mu \gamma^5 \mu)$. In the SM

$$C_{10}^{(\text{SM})} = \frac{G_F}{2\sqrt{2}} \frac{\alpha}{4\pi} (V_{tb}^* V_{ts}) \tilde{C}_{10}^{(\text{SM})}, \tag{2.35}$$

with $\tilde{C}_{10}^{(\text{SM})} \simeq 4.41$ [83], whereas from Eq.(2.33) we have

$$C_{10}^{(\text{BSM})} \simeq \frac{G_F}{\sqrt{2}} \cos(2\theta_w)(\epsilon_d^\dagger \epsilon_d)_{32}. \tag{2.36}$$

A stringent upper bound on the size of the $(\epsilon_d^\dagger \epsilon_d)_{32}$ coefficient arises from the requirement that the masses of the down-type quarks are correctly reproduced in this model. From Eq.(A.13), we have

$$(\epsilon_d^\dagger \epsilon_d)_{32} = \frac{v^2}{2} \sum_i \frac{(\tilde{y}_d)_{3i}(\tilde{y}_d)^*_{2i}}{m^2_{D_i}} \sim \frac{v^2}{M^2} \sum_i (\tilde{y}_d)_{3i}(\tilde{y}_d)^*_{2i} \lesssim \frac{v}{M} \frac{\sqrt{m_b m_s}}{v'} \ll 1, \tag{2.37}$$

where in the last step we have made use of the upper bound in Eq.(A.12). We then have, parametrically

$$r_{\mu\mu} \sim \frac{2|C_{10}^{(\mathrm{BSM})}|}{|C_{10}^{(\mathrm{SM})}|} \lesssim 10^{-3} \left(\frac{18 \text{ TeV}}{v'}\right)^2 \left(\frac{v'}{M}\right). \tag{2.38}$$

This effect is much smaller than the theoretical and experimental errors on $\mathrm{BR}(B_s^0 \to \mu^+\mu^-)$, which are both on the order of 10% [84].[8]

The effects of Z-mediated FCNCs on other processes are even more suppressed. For example, $|\Delta F| = 2$ processes such as kaon mixing require two insertions of the (tiny) flavor-violating coefficient. In the lepton sector, even $|\Delta F| = 1$ decays are virtually unobservable, as the effect is now suppressed by the masses of the relevant leptons. FCNCs mediated by the SM Higgs are similarly negligible, since the corresponding Wilson coefficients feature the same suppression as those from Z exchange, on top the smaller coupling between the Higgs and light fermions. Flavor-changing interactions mediated by the Z' and h' are further suppressed by an additional factor of $m_Z^2/m_{Z'}^2$ and $m_h^2/m_{h'}^2$ respectively, making them irrelevant. Overall, the strong suppression of the tree-level FCNCs that occurs naturally in these models makes their effects negligible.

One-loop FCNCs

Another source of FCNCs beyond those present in the SM arises at one loop. In these models, the familiar box diagram that describes meson mixing in the SM is now

[8] The effective operator $(\bar{b}_R \gamma^\mu s_R)(\bar{\mu}\gamma_\mu\gamma^5\mu)$ is also generated after integrating out the Z, with a Wilson coefficient $C_{10}^{\prime(\mathrm{BSM})} \sim C_{10}^{(\mathrm{BSM})}$ that enters into Eq.(2.34) in a similar manner. The presence of this operator does not quantitatively affect our analysis.

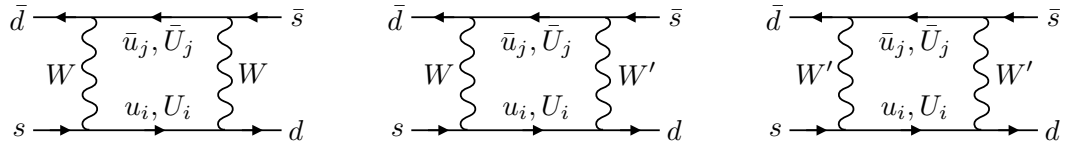

Figure 2.2: Additional box diagrams contributing to kaon mixing in the models under consideration include any of the up-type quarks propagating inside the loop, as well as (left) two W bosons, (center) one W and one W', and (right) two W's. Quantitatively, it is diagrams with one W and one W' that are the most relevant.

accompanied by similar diagrams that include W's as well as the additional (heavy) up-type quarks running inside the loop, as we show in figure 2.2. Given the level of experimental precision in measurements of kaon mixing parameters, even a modification to this process at the loop level can be a significant source of constraints.

The relevant interactions are those involving the SM-like down-type quarks and both the W and W' gauge bosons. They are given by

$$\mathcal{L} \supset \frac{g}{\sqrt{2}} W^+_\mu \sum_{j,i=1}^{3} \left(V_{ij} \bar{u}_{i_L} \gamma^\mu d_{j_L} + \Delta V_{ij} \bar{U}_{i_L} \gamma^\mu d_{j_L} \right) + \text{h.c.}, \tag{2.39}$$

where $U_{3L} = t'_L$ here and $\Delta V = \epsilon_u V$, as well as

$$\begin{aligned} \mathcal{L} \supset \frac{g}{\sqrt{2}} W'^+_\mu \sum_{j=1}^{3} \Bigg(& \sum_{i=1}^{2} V_{ij} \bar{u}_{i_R} \gamma^\mu d_{j_R} + V_{3j} \bar{t}'_R \gamma^\mu d_{j_R} \\ & + \sum_{i=1}^{2} \Delta V'_{ij} \bar{U}_{i_R} \gamma^\mu d_{j_R} + \Delta V'_{3j} \bar{t}_R \gamma^\mu d_{j_R} \Bigg), \end{aligned} \tag{2.40}$$

with $\Delta V' = \epsilon'^*_u V$. The entries of the 3×3 matrices ϵ_u and ϵ'_u are $\mathcal{O}(v/M)$ and $\mathcal{O}(v'/M)$ respectively, and explicit expressions can be found in Eq.(A.19).

The detailed expressions, including loop functions, relevant to estimate the contributions to the kaon mixing parameters Δm_K and $|\epsilon_K|$ can be found in appendix A.3. Additional box diagrams including two Ws or two W's always lead to a contribution which is much smaller than that of the SM, and can therefore be neglected. The leading contribution arises from diagrams including one W and one W'. In this case, there is an "irreducible" contribution to both parameters (irreducible in the sense that it can only

be "turned-off" by increasing v'), which comes from the u and c quarks, whose couplings to the W' gauge boson are set to be equal to those of the CKM matrix as a result of generalized parity. The size of this correction reads

$$(\Delta m_K)_{u,c} \approx -6 \cdot 10^{-16}\ \text{GeV} \left(\frac{6\ \text{TeV}}{m_{W'}}\right)^2, \qquad \text{and} \qquad |\epsilon_K|_{u,c} \approx 7 \cdot 10^{-5} \left(\frac{6\ \text{TeV}}{m_{W'}}\right)^2, \tag{2.41}$$

which in both cases is an order of magnitude below the theoretical error in the corresponding SM prediction, for values of $m_{W'}$ consistent with the direct bounds discussed in 2.3.1.

Contributions from box diagrams that involve additional members of the up-quark sector additionally depend on the see-saw scale M, as well as the size of both diagonal and off-diagonal entries in the up-type Yukawa matrices. As far as Δm_K is concerned, the leading contribution comes from diagrams where the u and c quarks propagate inside the loop, and so it is roughly equal to the result in Eq.(2.41), even for a see-saw scale M that sits only slightly above v'. In contrast, the contribution to $|\epsilon_K|$ can be large, and it is dominated by diagrams where the t quark propagates inside the loop. Choosing the individual entries in the Yukawa couplings to saturate the upper bound given in Eq.(A.18), $|\epsilon_K|$ sets a lower bound on M than can range between 750 TeV and 1000 TeV (depending on whether the leading contribution interferes destructively or constructively with the SM result) for $v' \sim 18$ TeV. This value of M sits comfortably within the upper bound $M \lesssim 10^2 v'$, which follows from the requirement of perturbative Yukawas, as discussed around Eq.(2.18). Alternatively, even for $v' = 18$ TeV and $M = 40$ TeV, an additional suppression by a factor of $\mathcal{O}(0.1)$ in the off-diagonal elements of the up-type Yukawas with respect to their upper bound is enough to bring the predicted value of $|\epsilon_K|$ within the allowed range.

Overall, the class of parity solutions to the strong CP problem that we focus on in this

work can comfortably satisfy existing constraints from flavor physics. Flavor-changing processes are, nevertheless, an interesting probe of the structure of these models, and a more in-depth investigation is a promising avenue for future work.

2.4 Broken parity and the neutron EDM

As we discussed in section 2.2, parity-symmetric theories predict a vanishing $\bar{\theta}$, therefore offering a potential solution to the strong CP problem. However, the breaking of parity that is necessary for phenomenological reasons implies that, although zero at tree-level, a non-zero $\bar{\theta}$ may be generated radiatively. In this section, we investigate in detail the size of radiative corrections to both $\bar{\theta}$, and the EDM of elementary fermions. We focus on the effect of non-gravitational interactions, and leave gravitational considerations to section 2.5.

The size of radiative corrections to the $\bar{\theta}$ parameter is a somewhat model-dependent question, as it depends on the details of how parity is broken. For instance, we could regard generalized parity to be a global symmetry that is only broken softly by dimensionful parameters, as in Eq.(2.13). More realistically, we might expect that the breaking of parity is spontaneous, and not explicit. This must certainly be the case if, for example, parity were a gauge symmetry of the UV theory. Even in this case, there are two qualitatively different options: either parity is broken without breaking CP (e.g. through a symmetry-breaking sector with two scalar fields that obtain asymmetric vevs); or both parity and CP are broken simultaneously (e.g. through the vev of a pseudo-scalar). The former situation is quantitatively similar to the global case. In the latter, however, the symmetry-breaking sector can introduce an additional source of CP-violation beyond that present in the SM, and a non-vanishing $\bar{\theta}$ can arise already at one loop.

In the remainder of this section, we discuss the three qualitatively different possibil-

ities for the breaking of parity, with a focus on the implications for the size of radiative corrections to the neutron EDM.

2.4.1 Softly broken parity

We will first discuss the possibility of parity being broken softly, only as a result of dimensionful parameters. Performing this analysis will give us an understanding of the irreducible effects that will be present in any theory where the breaking of parity happens dynamically.

There are two potential sources of soft breaking. One corresponds to the μ^2 term in the scalar potential of Eq.(2.13), which splits the Higgs vevs in the SM and mirror sectors. As anticipated in the introduction, if this was the only source of parity-violation, radiative corrections to $\bar{\theta}$ would be no larger than in the SM [23]. Another potential source of soft breaking are the vector-like mass matrices of Eq.(2.10). Relaxing the requirement that these be hermitian introduces a soft breaking of both generalized parity and CP. In this case, a correction to the EDMs of elementary charged fermions (both quarks and leptons) arises already at one loop, whereas $\bar{\theta}$ remains zero both at the tree- and one-loop levels. In turn, this translates into a contribution to the neutron EDM *independent* of $\bar{\theta}$.

Taking the vector-like mass matrices of the $SU(2)$-singlets to be general complex matrices, WLOG we may write them as

$$\mathcal{M}'_f = \mathcal{M}_f + i\Delta\mathcal{M}_f, \tag{2.42}$$

where both $\mathcal{M}_f^\dagger = \mathcal{M}_f$ and $\Delta\mathcal{M}_f^\dagger = \Delta\mathcal{M}_f$. If $\Delta\mathcal{M}_f$ is non-vanishing, $\mathcal{M}'_f$ is no longer hermitian, therefore (softly) breaking both generalized parity and CP. At one-loop, a non-zero $\Delta\mathcal{M}_f$ leads to a non-vanishing contribution to the EDM of elementary fermions, with the relevant diagrams depicted in figure 2.3. The result is dominated by diagrams where the mirror Higgs, h', and the heavy mirror fermions propagate inside the loop. We

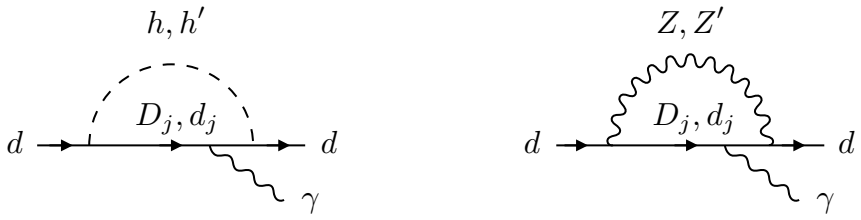

Figure 2.3: One-loop diagrams generating a non-zero EDM for the d quark in the presence of soft breaking of generalized parity through non-hermitian vector-like mass matrices for the $SU(2)$-singlets. The leading contribution arises from the diagram where h', and the heavy mirror quarks, D_j, propagate inside the loop. Analogous diagrams are present in both the up-quark and lepton sectors.

present a detailed calculation in appendix A.2.1. For any of the light SM fermions, we find

$$\frac{d_f}{e} \simeq \frac{n_f Q_f m_f}{32\pi^2 M^2} \times \mathcal{O}\left(\frac{|\Delta\mathcal{M}|}{M}\right). \tag{2.43}$$

where n_f is the number of mirror fermions appearing at the see-saw scale in each fermion sector (i.e. $n_d = n_e = 3$, and $n_u = 2$), and $|\Delta\mathcal{M}|$ refers to the typical size of the entries in the $\Delta\mathcal{M}$ matrix.

Taking the soft-breaking through $\Delta\mathcal{M}$ to be $\mathcal{O}(1)$, we find, parametrically,

$$d_u, d_d \sim 10^{-28} \left(\frac{40 \text{ TeV}}{M}\right)^2 e \cdot \text{cm}. \tag{2.44}$$

In turn, this will translate into an EDM for the neutron of approximately the same size. For illustration, we have normalized the above expression to a value of M that is roughly a factor of two larger than the current lower bound on v'. The corresponding result lies two orders of magnitude below the current experimental bound on d_n, and could fall within reach of future experiments depending on the value of the see-saw scale (see e.g. [85] for a survey of prospective molecule-based searches promising orders-of-magnitude improvement in sensitivity to hadronic CPV).

The see-saw mechanism must also be implemented in the charged lepton sector. If it were not, the mirror partner of the SM electron would appear at a scale $m_{e'} \simeq m_e \times v'/v$,

which would be as low as ~ 40 MeV for the least fine-tuned version of the model where $v' \simeq 18$ TeV. Since mirror fermions carry the same electromagnetic charge as their SM counterparts, this possibility is obviously ruled out. As a result, a non-zero electron EDM is also a generic prediction of this class of theories. Parametrically

$$d_e \sim 10^{-29} \left(\frac{90 \text{ TeV}}{M}\right)^2 e \cdot \text{cm}, \tag{2.45}$$

where we have chosen the see-saw scale in the lepton sector so as to saturate the current upper bound on the electron EDM, which is $|d_e| < 1.1 \cdot 10^{-29}\ e \cdot \text{cm}$ [86].

Although an EDM is generated at one loop for the various elementary fermions, $\bar{\theta}$ remains zero at this order. At tree-level, it is easy to see that $\bar{\theta} = 0$, even in the presence of non-hermitian vector-like mass matrices. Working in the flavor basis, the full 6×6 mass matrices in both the up- and down-quark sectors are only modified with respect to Eq.(2.12) by replacing $\mathcal{M}_f$ with $\mathcal{M}'_f$ in the bottom-right block. We then have

$$\det \mathbb{M}_f = \det \begin{pmatrix} \not{<}_3 & \frac{v'}{\sqrt{2}} y'^*_f \\ \frac{v}{\sqrt{2}} y_f^T & \mathcal{M}'_f \end{pmatrix} \propto \det \left(y'^*_f y_f^T\right) \qquad \text{for} \qquad f = u, d, \tag{2.46}$$

which is real regardless of $\mathcal{M}'_f$, since $y'_f = y_f$. This is clearly an accidental consequence of the zero appearing in the upper-left corner of the quark mass matrix — the gauge structure of the theory does not allow for relevant operators with the appropriate quantum numbers to fill that block. The vanishing of $\bar{\theta}$ at one-loop is less immediately obvious. The relevant calculation was performed in [22], and it is also apparent as a byproduct of our EDM calculation in appendix A.2.1. As already emphasized in [22], a non-zero correction to $\bar{\theta}$ could appear at the two-loop order, and would lead to an additional contribution to the neutron EDM that could be comparable in size to the one discussed here.

2.4.2 Spontaneously broken parity and CP

Perhaps more compellingly — and necessarily, if parity is a gauge symmetry — the breaking of parity can be accomplished through the vacuum expectation value of an additional field. The most minimal realization actually entails the breaking of both parity and CP through the vev of a pseudo-scalar field ϕ. The soft term in Eq.(2.13) is generated by pseudo-scalar couplings to the Higgs sector, of the form

$$V \supset \mu_\phi \phi(|H|^2 - |H'|^2) + \lambda_\phi \phi^2(|H|^2 + |H'|^2) \ . \tag{2.47}$$

The first term above splits the two vevs, and to obtain $v' \gg v$ entails

$$\mu_\phi v_\phi \sim \kappa v'^2 \ . \tag{2.48}$$

A natural possibility is to take $v_\phi \sim \mu_\phi \sim v'$, with $\kappa = \mathcal{O}(1)$. However, $\kappa \ll 1$ is also possible, especially since this coupling breaks the otherwise accidental $SU(4)$ global symmetry of the scalar potential in Eq.(2.13). Indeed, the quartic coupling of the SM-like Higgs is $\lambda_{SM} \sim 2\kappa$, suggesting $\kappa \lesssim 0.1$ and thus a pseudo-scalar vev v_ϕ that is numerically somewhat smaller than v'.

Crucially, there is an additional operator consistent with all symmetries that involves ϕ and the $SU(2)$-singlet fermions, of the form [74]:

$$\mathcal{L} \supset i(\bar{y}_d)_{ij} \phi D_i D'_j + \text{h.c.}, \tag{2.49}$$

and similarly for up-type quarks and leptons. The $\bar{y}_f$ matrices must be hermitian in order to respect generalized parity. When ϕ gets a vev, this term breaks both parity *and* CP. In the notation of section 2.4.1, a non-hermitian contribution to the vector-like masses in the fermion sector is generated, of the form $\Delta\mathcal{M}_f = \bar{y}_f v_\phi$. More importantly, new interactions involving the pseudo-scalar lead to a non-zero contribution to $\bar{\theta}$ already at one loop, which sets stringent constraints on the size of these couplings. The relevant

diagrams are those on the left of figure 2.3, minus the external photon line, and allowing for ϕ to propagate inside the loop. In appendix A.2.2, we present a detailed calculation, performed in the mass eigenbasis, of the one-loop correction to the quark mass matrix, and the corresponding correction to $\bar{\theta}$, in the context of left-right models with a see-saw fermion structure. In the remainder of this section, we will reproduce the parametric contribution to $\bar{\theta}$ from the down-quark sector using a spurion analysis that the reader might find more instructive.

If we parametrize the one-loop correction to the 6×6 mass matrix in the down-quark sector in terms of 3×3 blocks, as follows

$$\Delta \mathbb{M}_d \equiv \begin{pmatrix} \Delta_{d'd} & \Delta_{d'D'} \\ \Delta_{Dd} & \Delta_{DD'} \end{pmatrix}, \tag{2.50}$$

then the corresponding contribution to $\bar{\theta}$ from the down-quark sector can be written as

$$\begin{aligned} \theta_d &\equiv \arg\ \det\left(\mathbb{M}_d + \Delta\mathbb{M}_d\right) \\ &\simeq \mathrm{Im}\,\mathrm{Tr}\left(\mathbb{M}_d^{-1}\Delta\mathbb{M}_d\right) \\ &= \mathrm{Im}\,\mathrm{Tr}\left\{-\left(\frac{vv'}{2}y_d'^{*}\mathcal{M}_d^{-1}y_d^T\right)^{-1}\Delta_{d'd} + \left(\frac{v}{\sqrt{2}}y_d^T\right)^{-1}\Delta_{Dd} + \left(\frac{v'}{\sqrt{2}}y_d'^{*}\right)^{-1}\Delta_{d'D'}\right\}. \end{aligned} \tag{2.51}$$

Notice that the $\Delta_{DD'}$ block does not contribute to θ_d at this order, which again is a consequence of the zero in the upper-left corner of $\mathbb{M}_d$.

We will now estimate the size of the Δ matrices appearing in Eq.(2.51) through a spurion analysis, as follows. The lagrangian will remain invariant under $SU(3)$ flavor transformations on the various quark fields, of the form

$$d \to \mathcal{R}_Q d, \qquad d' \to \mathcal{R}_{Q'} d', \qquad \text{and} \qquad D \to \mathcal{R}_D D, \qquad D' \to \mathcal{R}_{D'} D', \tag{2.52}$$

provided the various Yukawa couplings, as well as the vector-like mass matrix, similarly transform in an appropriate manner. The correct transformation rules for these objects

are given by

$$\begin{aligned} y_d^T &\to \mathcal{R}_D^* y_d^T \mathcal{R}_Q^\dagger \,, & y_d'^* &\to \mathcal{R}_{Q'}^* y_d'^* \mathcal{R}_{D'}^\dagger \,, \\ \mathcal{M}_d &\to \mathcal{R}_D^* \mathcal{M}_d \mathcal{R}_{D'}^\dagger \,, & \bar{y}_d &\to \mathcal{R}_D^* \bar{y}_d \mathcal{R}_{D'}^\dagger \,. \end{aligned} \tag{2.53}$$

On the other hand, the Δ matrices of Eq.(2.51) must similarly transform as follows:

$$\Delta_{d'd} \to \mathcal{R}_{Q'}^* \Delta_{d'd} \mathcal{R}_Q^\dagger \,, \quad \Delta_{d'D'} \to \mathcal{R}_{Q'}^* \Delta_{d'D'} \mathcal{R}_{D'}^\dagger \,, \quad \text{and} \quad \Delta_{Dd} \to \mathcal{R}_D^* \Delta_{Dd} \mathcal{R}_Q^\dagger \,. \tag{2.54}$$

It is now straightforward to identify the leading objects that transform as in Eq.(2.54) and contain a single insertion of $\bar{y}_d$. These are of the form

$$\Delta_{d'd} \sim \frac{v v' v_\phi}{16\pi^2} \left(y_d'^* \mathcal{M}_d^{-1} \bar{y}_d \mathcal{M}_d^{-1} y_d^T \right) , \tag{2.55}$$

whereas

$$\Delta_{d'D'} \sim \frac{v' v_\phi}{16\pi^2} \left(y_d'^* \mathcal{M}_d^{-1} \bar{y}_d \right) , \qquad \text{and} \qquad \Delta_{Dd} \sim \frac{v v_\phi}{16\pi^2} \left(\bar{y}_d \mathcal{M}_d^{-1} y_d^T \right) . \tag{2.56}$$

We have also included numerical factors to account for the loop suppression, as well as to take into account that the contribution to $\bar{\theta}$ must not diverge in the limits where either v or v' vanish. Plugging this back into Eq.(2.51), we find that all three terms give a contribution of the same size. Parametrically:

$$\theta_d \sim \frac{v_\phi}{16\pi^2} \mathrm{Tr} \left(\mathcal{M}_d^{-1} \bar{y}_d \right) \sim \frac{|\bar{y}_d| v_\phi}{16\pi^2 M} . \tag{2.57}$$

This result is consistent with the more detailed calculation of the contribution to $\bar{\theta}$ from the quark sector presented in appendix A.2.2.

Requiring that $\bar{\theta} \lesssim 10^{-10}$ sets an upper bound on the typical size of the entries of the $\bar{y}$ matrices in the quark sector, of the form

$$\bar{y} \lesssim 10^{-8} \frac{M}{v_\phi} \lesssim 10^{-6} , \tag{2.58}$$

where in the last step we have assumed that $v_\phi \sim v'$, and have taken into account the upper bound on the see-saw scale M as given in Eq.(2.18).

This result bring us to the following conclusion: if the spontaneous breaking of parity also implies breaking CP, then any interaction between the quark and symmetry breaking sectors must be extremely weak. Fortunately, if $\bar{y} = 0$ at tree-level, a non-zero value of $\bar{y}$ will not be generated radiatively. Indeed, $\bar{y}$ and the vector-like mass matrices $\mathcal{M}$ are the only two parameters that violate the $\mathbb{Z}_2$ symmetry acting on the matter fields of the mirror sector. The breaking through $\mathcal{M}$, however, is soft, and therefore will not translate into a non-zero $\bar{y}$ at the loop order. In this sense, a vanishing $\bar{y}$ is technically natural.

2.4.3 Spontaneously broken parity alone

A less minimal possibility is to spontaneously break parity while preserving CP through the addition of two scalar fields, σ and σ', whose vevs differ. This can be achieved if this symmetry breaking sector has a scalar potential of the form

$$V_\sigma = -\frac{m_\sigma^2}{2}(\sigma^2 + \sigma'^2) + \frac{\lambda_1}{4}(\sigma^2 + \sigma'^2)^2 + \frac{\lambda_2}{4}\sigma^2\sigma'^2 \ . \tag{2.59}$$

where for simplicity we have forbidden cubic terms by imposing an additional $\mathbb{Z}_2$ symmetry. If $\lambda_2 > 0$, the vacua lie at $\langle\sigma\rangle = 0$, $\langle\sigma'\rangle = \pm\sqrt{m_\sigma^2/\lambda_1}$ and viceversa. This option is not viable for the Higgs potential itself, which requires both v and v' to be nonzero, but is perfectly adequate for an additional scalar sector.

Parity breaking can then be translated into the Higgs sector by writing appropriate couplings of the form

$$V \supset \lambda_\sigma \left(\sigma^2|H|^2 + \sigma'^2|H'|^2\right) + \lambda'_\sigma(\sigma'^2|H|^2 + \sigma^2|H'|^2) \ . \tag{2.60}$$

These terms are compatible with the generalized parity introduced in section 2.2, acting additionally as $\sigma \leftrightarrow \sigma'$. Provided $\lambda_\sigma \neq \lambda'_\sigma$, this will generate the soft term in Eq.(2.13) proportional to $\lambda_\sigma - \lambda'_\sigma$. For example, in the vacuum with $\langle\sigma'\rangle \neq 0$, $v' \gg v$ corresponds to

$$(\lambda'_\sigma - \lambda_\sigma)\langle\sigma'\rangle^2 \sim \kappa v'^2. \tag{2.61}$$

As this scenario breaks P without breaking CP (and the additional $\mathbb{Z}_2$ symmetry acting on the σs forbids marginal couplings between σ, σ' and fermion bilinears), there are no significant additional contributions to the neutron EDM. There is, of course, the possibility of collider signatures coming from the Higgs portal coupling in Eq.(2.60), most notably mixing between the Higgs and the scalar that acquires a vev, as well as invisible decays of the Higgs if kinematically allowed. The two scalars acquire masses of order $\sqrt{2\lambda_1}\langle\sigma'\rangle \sim \sqrt{\lambda_1}v'$ and $\sqrt{\lambda_2/2}\langle\sigma'\rangle \sim \sqrt{\lambda_2}v'$, respectively, and it is certainly possible for one to be lighter than half the Higgs mass depending on the values of $\lambda_{1,2}$.

2.5 Strong CP and quantum gravity

As we discussed in the Introduction, the strong CP problem arises out of the difficulty of reconciling the smallness of $\bar{\theta}$ with the $\mathcal{O}(1)$ violation of both parity and CP by the electroweak sector. In turn, all attempts to address this puzzle are themselves based on the introduction of an additional symmetry beyond those of the SM. However, there is strong evidence that within a theory of quantum gravity, global symmetries cannot be exact — they must be either broken, or gauged. The origin of this statement goes back a long way [38–46], and to some extent it has recently been established [47, 48]. Of course, the single most pressing issue for phenomenology is to establish a lower bound on the amount of global symmetry violation that must be present in the IR. Attempts at finding such a "universal" lower bound have been made [49, 50], but a fully satisfactory answer remains elusive. Absent a full understanding of how quantum gravity affects global symmetries at low energies, we can at least attempt to assess the robustness of an EFT against global symmetry violation by considering the impact of HDOs suppressed by the appropriate power of M_{Pl}. This both constrains the viable parameter space of parity solutions to strong CP and illustrates the sense in which P, rather than $U(1)_{PQ}$,

provides a solution to the strong CP problem that is robust against the expected intrusion of quantum gravity. Beyond imposing constraints, these higher-dimensional operators also lead to new experimental signatures associated with the spontaneous breaking of parity, which we explore in section 2.5.2.

2.5.1 Constraints from Planck-suppressed operators

The observation that the breaking of global symmetries by quantum gravity can have a profound impact on the validity of the QCD axion solution to strong CP was first made in [58–61]. Planck-suppressed HDOs that violate the $U(1)_{PQ}$ symmetry carried by the field Φ, the phase of which is the axion, are of the form

$$\mathcal{L} \supset \frac{\eta}{M_{Pl}^{d-4}} |\Phi|^{d-n} \Phi^n + \text{h.c.} \tag{2.62}$$

Here, d is the operator dimension, n its units of $U(1)_{PQ}$ charge (so $n \geq 1$ in order to break the symmetry), and η a coupling that will in general feature arbitrary real and imaginary parts. HDOs of this form contribute to the axion potential, and, in general, will displace the axion vev away from the value leading to a small $\bar{\theta}$. Following [59], requiring that the shift in the axion vev is small enough so as not to spoil the solution to strong CP translates into the following upper bound

$$|\eta| \left(\frac{f_a}{\sqrt{2} M_{Pl}} \right)^d \lesssim 10^{-81} \bar{\theta} \lesssim 10^{-91}, \tag{2.63}$$

where f_a is the scale of $U(1)_{PQ}$ spontaneous symmetry breaking (alternatively, the axion decay constant), which is experimentally constrained to be between 10^8 and 10^{17} GeV [84]. Focusing on operators of dimension $d = 5$, this translates into an upper bound on the size of η, of the form

$$|\eta| \lesssim 10^{-55} \left(\frac{10^{12}\text{GeV}}{f_a} \right)^5 \left(\frac{\bar{\theta}}{10^{-10}} \right) . \tag{2.64}$$

In other words, for all experimentally allowed values of the axion decay constant, the $U(1)_{PQ}$ symmetry must remain an approximate global symmetry to an exceptional degree. This is clearly one of the most significant drawbacks of the axion solution to strong CP.

In the remainder of this section we study the effect of Planck-suppressed HDOs on parity solutions to the strong CP problem. We consider separately the cases where parity is global or gauged. The nature of the HDOs under consideration will be different, but in both cases we will see that even $\mathcal{O}(1)$ coefficients are compatible with solving strong CP.

Parity as a global symmetry

If we regard parity as a global symmetry, then we must consider the effect of HDOs that explicitly violate P. The relevant dimension-5 HDOs were already identified in [87], and they are of the form

$$\mathcal{L} \supset \frac{1}{M_{Pl}} \left[(\alpha_u)_{ij} (H' Q'_i)(H Q_j) + (\alpha_d)_{ij} (H'^\dagger Q'_i)(H^\dagger Q_j) \right] + \text{h.c.} \tag{2.65}$$

Notice that if $\alpha_f = \alpha_f^\dagger$ then the above terms would be parity-symmetric. In general, however, the α_f's will *not* be hermitian, and it is under this assumption that we proceed.

Setting the Higgs to their vevs, Eq.(2.65) leads to a correction to the quark mass matrix that, for arbitrary α_f's, does not respect generalized parity. The leading contribution to $\bar{\theta}$ will come from the contributions to the up- and down-quark masses, which are of the form

$$\delta m_u \simeq \frac{v v' (\alpha_u)_{11}}{2 M_{Pl}}, \qquad \text{and} \qquad \delta m_d \simeq \frac{v v' (\alpha_d)_{11}}{2 M_{Pl}}. \tag{2.66}$$

In turn,

$$\theta_q \simeq \frac{\text{Im}(\delta m_u)}{m_u} + \frac{\text{Im}(\delta m_d)}{m_d} \sim 10^5 \frac{|\alpha| v'}{2 M_{Pl}}, \tag{2.67}$$

where in the last step we have used $m_u/v \sim m_d/v \sim 10^{-5}$. Requiring that the above contribution is smaller than the current bound on $\bar{\theta}$ translates into an upper bound on the parity breaking scale:

$$v' \lesssim \frac{20 \text{ TeV}}{|\alpha|} \left(\frac{\bar{\theta}}{10^{-10}} \right). \tag{2.68}$$

Notice this upper bound is (just) compatible with the lower bound $v' \gtrsim 18$ TeV from direct searches of W' gauge bosons, as discussed in section 2.3.1. As a result, if *global* generalized parity is responsible for solving strong CP, an $\mathcal{O}(1)$ violation of the symmetry due to gravitational effects would imply a contribution to $\bar{\theta}$ accessible in near-future experiments.

Parity as a gauge symmetry

If parity is instead a gauge symmetry of the underlying theory, HDOs that explicitly violate P are therefore not allowed. Planck-suppressed operators such as those in Eq.(2.65) might still be generated, but only with $\alpha_f = \alpha_f^\dagger$, and therefore will not contribute to $\bar{\theta}$. Instead, the operators of interest must be proportional to the source of spontaneous symmetry breaking. If the latter takes place via the vev of a pseudo-scalar, as discussed in section 2.4.2, then there are two dimension-5 HDOs that satisfy this requirement, namely:

$$\mathcal{L} \supset \eta_s \frac{\phi \alpha_s}{4\pi M_{Pl}} \text{Tr}\left(G^a \tilde{G}^a \right), \tag{2.69}$$

and

$$\mathcal{L} \supset \frac{i\phi}{M_{Pl}} \left\{ (\zeta_u)_{ij} Q_i H U_j + (\zeta'_u)_{ij} Q'_i H' U'_j + (\zeta_d)_{ij} H^\dagger Q_i D_j + (\zeta'_d)_{ij} H'^\dagger Q'_i U'_j \right\} + \text{h.c.}, \tag{2.70}$$

with $\eta_s \in \mathbb{R}$, and $\zeta'_f = \zeta_f^*$ so as to satisfy generalized parity.

The operator of Eq.(2.69) will generate a contribution to θ_s after spontaneous symmetry breaking, of the form

$$\theta_s \simeq \frac{\eta_s v_\phi}{M_{Pl}}. \tag{2.71}$$

Assuming that $v_\phi \sim v'$, demanding that this contribution is smaller than the current bound on $\bar{\theta}$ leads to an upper bound on the parity breaking scale

$$v_\phi \sim v' \lesssim \frac{10^9 \text{ GeV}}{\eta_s} \left(\frac{\bar{\theta}}{10^{-10}} \right), \tag{2.72}$$

which is clearly well above current bounds on v'.

At the same time, once ϕ gets its vev, the operator of Eq.(2.70) leads to an extra contribution to the Yukawa couplings of the up- and down-type quarks in the SM and mirror sectors that are *not* parity-symmetric. In turn, this will lead to an additional contribution to the mass eigenvalues of the light quarks which will in general contain an imaginary component. For example, in the down quark sector

$$\text{Im}(\delta m_{d_i}) \simeq v v' \,\text{Im} \left\{ \frac{i v_\phi}{M_{Pl}} \sum_j \frac{(y_d)^*_{ij} (\zeta_d)_{ij}}{m_{D_j}} \right\} \sim \frac{v v'}{M} \frac{|\zeta_d| v_\phi}{M_{Pl}} |(y_d)_{i\star}|, \tag{2.73}$$

where $|(y_d)_{i\star}|$ refers to the typical size of the entries in the i-th row of the y_d matrix. The leading contribution to θ_q will come from the up and down quarks. In total:

$$\theta_q \simeq \frac{\text{Im}(\delta m_u)}{m_u} + \frac{\text{Im}(\delta m_d)}{m_d} \simeq \frac{v_\phi}{M_{Pl}} \frac{v v'}{M} \left\{ \frac{|\zeta_u| (y_u)_{1\star}}{m_u} + \frac{|\zeta_d| (y_d)_{1\star}}{m_d} \right\}. \tag{2.74}$$

Taking into account the upper bound on the entries of the Yukawa couplings necessary to reproduce the light quark masses (see Eq.(A.12)), as well as the requirement that $v' \lesssim M$ in order to implement the see-saw mechanism, the previous equation implies

$$\theta_q \lesssim 10^2 \frac{|\zeta| v_\phi}{M_{Pl}}, \tag{2.75}$$

where we have set $m_u/v \sim m_d/v \sim 10^{-5}$, and have assumed that $|\zeta_u| \sim |\zeta_d|$. In turn, taking $v_\phi \sim v'$, this sets an upper bound on the scale of spontaneous symmetry breaking:

$$v_\phi \sim v' \lesssim \frac{10^7 \text{ GeV}}{|\zeta|} \left(\frac{\bar{\theta}}{10^{-10}} \right). \tag{2.76}$$

As before, this is fully compatible with current experimental bounds on the parity-breaking scale, even for $\mathcal{O}(1)$ coefficients of the corresponding HDOs.

2.5.2 Gravitational waves from the spontaneous breaking of parity

Beyond providing additional constraints on the parameter space of parity solutions to strong CP, the expected effects of gravity also introduce new experimental signatures. Here we highlight one possibility, namely the impact of HDOs when parity is a spontaneously broken global symmetry. The spontaneous breaking of discrete symmetries can lead to the formation of a network of domain walls in the early universe, provided the reheating temperature after inflation is above the scale of spontaneous symmetry breaking [88] [9]. If the spontaneously broken symmetry is global, but otherwise exact, a domain wall configuration interpolates between two distinct vacua that are degenerate, making these defects topologically stable objects. The formation of such networks can be fatal on two grounds. On the one hand, the energy density in domain walls redshifts slower than that of matter or radiation, and would eventually dominate the universe's energy budget. If this happened before the current epoch, the rapid expansion of the subsequent domain-wall-dominated era would be at odds with observation. On the other hand, even if only a subdominant component of the total energy density was in the form of domain walls today, their effect on large-scale density fluctuations rules out defects with characteristic scales above ~ 1 MeV [91]. These considerations are often referred to as the "domain wall problem" of theories with spontaneously broken discrete symmetries.

These problems are largely solved when we take into consideration that, within a theory of quantum gravity, we expect all symmetries to be either broken or gauged [38–48] — an expectation that includes spacetime symmetries [47, 48]. In this context, the domain wall network is unstable, rendering its earlier formation largely unproblematic (see

[9] This statement relies on the restoration of the spontaneously broken symmetry at high temperatures. Scenarios where symmetry restoration does not take place have been explored in [89, 90]. In these cases, topological defects would not form via the mechanism of [88].

e.g. [92], and also [89, 90]). Moreover, the significant amount of gravitational radiation emitted in the process results in a stochastic gravitational wave background that may be within reach of current and future observatories. We discuss this possibility in the remainder of this section.

We will focus first on the scenario where parity is a global symmetry that is only explicitly broken by gravitational effects. At low energies, the symmetry-breaking dynamics will enter the effective potential for ϕ through HDOs that violate parity. One such operator is of the form

$$V \supset \epsilon \frac{\phi^5}{M_{Pl}} \ . \tag{2.77}$$

This breaks the degeneracy between the two previously degenerate vacua, corresponding to $\langle\phi\rangle = \pm v_\phi$. Parametrically, the energy difference now reads

$$\delta V \sim \frac{\epsilon v_\phi^5}{M_{Pl}} \ . \tag{2.78}$$

If the reheating temperature is above the scale of spontaneous symmetry breaking, then we expect that a network of domain walls will be formed once the temperature of the universe drops bellow $T \sim v_\phi$ [88]. Numerical [93–97] and analytical [98, 99] studies suggest that, shortly after formation, the network evolves according to a scaling solution, with $\rho_{\rm DW}(t) \simeq \sigma/t$, and typical domain wall size comparable to the Hubble scale $H(t)^{-1}$. σ corresponds to the tension of the domain walls, which in our model is of the form $\sigma \sim \sqrt{\kappa_\phi} v_\phi^3$, where κ_ϕ refers to the quartic coupling in the ϕ potential. Two competing effects determine the network's subsequent evolution. On the one hand, the pressure difference between the two vacua exerts a force per unit area of order $\sim \delta V$. On the other, the tension per unit area acting on a wall with curvature radius R is $\sim \sigma/R$. In the scaling regime, $R(t) \sim H(t)^{-1} \sim t$ (assuming the universe is radiation dominated), and therefore the effect of tension decreases with time. Eventually, the pressure difference

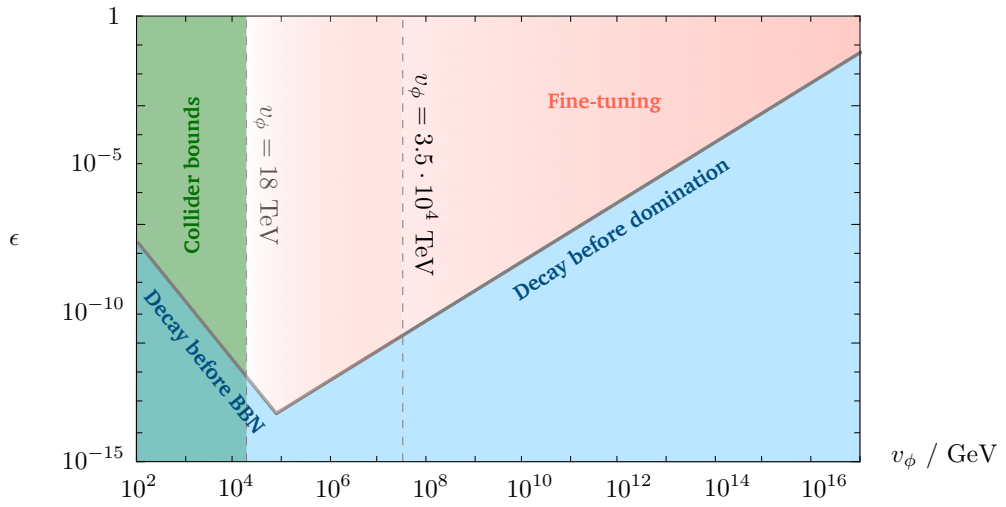

Figure 2.4: Constraints on the size of the coefficient of the Planck-suppressed HDO of Eq.(2.77), as a function of the scale of spontaneous symmetry breaking, v_ϕ. The region $v_\phi \lesssim 18$ TeV (green) is in conflict with direct bounds on the mass of W' and Z' resonances, under the assumption that $v_\phi \sim v'$, as discussed in section 2.3.1. Values of ϵ that are too small (blue) do not destabilize the domain wall network early enough to either avoid a domain-wall-dominated era, or to ensure collapse before the onset of BBN, and are therefore ruled out. The region of parameter space in pink is experimentally allowed, but the level of fine-tuning in the electroweak sector worsens as v_ϕ is increased (corresponding to a darker shade). The dashed line at $v_\phi \simeq 3.5 \cdot 10^4$ TeV corresponds to a fine-tuning of $\mathcal{O}(10^{-10})$ in the electroweak sector. (For illustration, we have set the quartic coupling in the pseudo-scalar potential to be $\kappa_\phi = 1$ in this plot.)

between the two vacua dominates, causing the network to collapse at a time of order

$$t_{\text{coll.}} \sim \frac{\sigma}{\delta V} \sim \frac{\sqrt{\kappa_\phi} M_{Pl}}{\epsilon v_\phi^2} \ . \tag{2.79}$$

Clearly, the domain wall network could be very long-lived if $\epsilon \lll 1$. The requirement that collapse takes place either before the universe becomes domain wall dominated, or before the start of BBN (so as to avoid energy injection into the SM plasma that would disrupt light element formation), sets a lower bound on ϵ as a function of the spontaneous symmetry breaking scale. This is depicted in figure 2.4, where the BBN and domain-wall-domination restrictions dominate for values of v_ϕ below and above ~ 73 TeV respectively. As can be appreciated in the figure, in the region of parameter space

where the fine-tuning is better than 10^{-10} (that is, $v_\phi \sim v' \lesssim 3.5 \cdot 10^4$ TeV), ϵ may be as small as $\mathcal{O}(10^{-13})$.

The collapse of a domain wall network leads to the production of gravitational waves [100, 101]. On dimensional grounds, one would expect the energy density in gravitational radiation to be of the form $\rho_{\rm gw} \sim G_N \sigma^2$ (the mandatory power of G_N times the necessary factors of σ to make up dimensions), an expectation that is largely upheld by numerical analysis [102–105]. The resulting gravitational wave spectrum has an extended shape, peaking at a frequency corresponding to the Hubble size at the time of collapse (corresponding to the typical size of the domain walls), and falling off as $1/f$ for larger frequencies. At the present epoch, the peak frequency of the gravitational wave signal is given by

$$\begin{aligned} f_* &\simeq 10^{-9}\ \text{Hz} \left(\frac{T_{\rm coll.}}{10^{-2}\ \text{GeV}}\right) \left(\frac{g_*(T_{\rm coll.})}{10}\right)^{1/6} \\ &\sim 10^{-9}\ \text{Hz} \left(\frac{v_\phi}{18\ \text{TeV}}\right) \left(\frac{\epsilon}{10^{-12}}\right)^{1/2} \left(\frac{1}{\kappa_\phi}\right)^{1/4}, \end{aligned} \tag{2.80}$$

and the energy density in gravitational radiation at frequency peak reads [106]

$$\begin{aligned} \Omega_{\rm gw} h^2(f_*) &\simeq 2 \cdot 10^{-10} \left(\frac{\sigma}{(20\ \text{TeV})^3}\right)^2 \left(\frac{10^{-2}\ \text{GeV}}{T_{\rm coll.}}\right)^4 \left(\frac{10}{g_*(T_{\rm coll.})}\right)^{4/3} \\ &\sim 10^{-10} \left(\frac{v_\phi}{18\ \text{TeV}}\right)^2 \left(\frac{10^{-12}}{\epsilon}\right)^2 \left(\frac{\kappa_\phi}{1}\right)^2, \end{aligned} \tag{2.81}$$

where $T_{\rm coll.}$ refers to the temperature of the SM plasma at a time $t_{\rm coll.}$, and we have assumed that network collapse takes place during radiation domination.[10]

Figure 2.5 shows the region that can be spanned by the peak of the stochastic gravitational wave background in the f_* vs. $\Omega_{\rm gw} h^2(f_*)$ plane, together with the sensitivity curves of a number of gravitational wave experiments. The lower bound on ϵ depicted in figure 2.4 translates into a lower bound on f_* for each value of the symmetry breaking

[10] In the second steps of Eq.(2.80) and (2.81), we have substituted $T_{\rm coll.}$ by the corresponding expression in terms of the model's fundamental parameters, while ignoring a weak dependence on g_*.

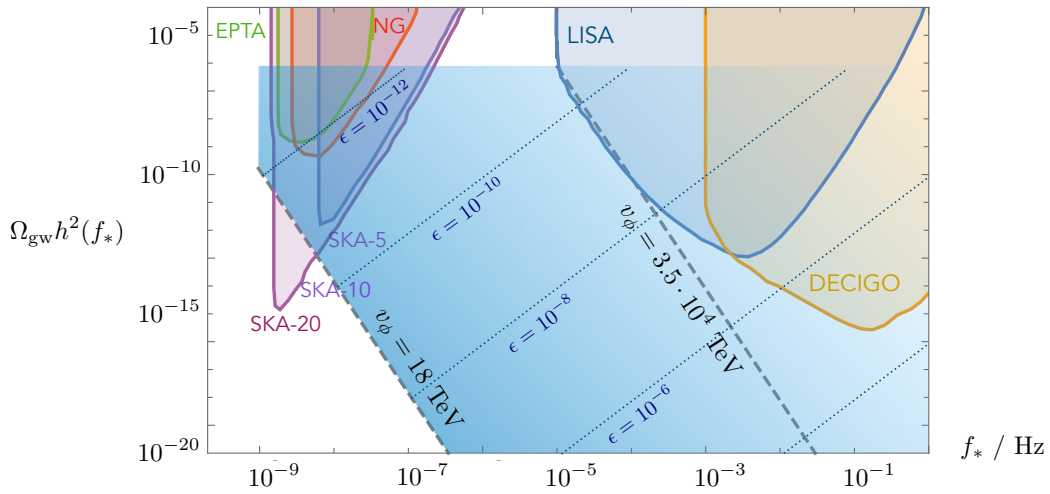

Figure 2.5: In blue, region of parameter space where the frequency peak of the stochastic gravitational wave signal, f_*, as well as the corresponding energy density, $\Omega_{\rm gw}h^2(f_*)$, can fall in light of the experimental constraints on the various model parameters summarized in figure 2.4. Dotted lines correspond to constant ϵ. The region to the right of the dashed line corresponding to $v_\phi = 3.5 \cdot 10^4$ TeV features a level of fine-tuning worse than 1 part in 10^{10}, and it is therefore less attractive. Sensitivity curves for a variety of gravitational wave experiments are shown, including the pulsar timing arrays EPTA [109], NANOGrav [108], SKA [107] (observation time of 5, 10 and 20 years as indicated), as well as the space-based interferometers LISA [110], and DECIGO [111]. (For illustration, we have set the quartic coupling in the pseudo-scalar potential to be $\kappa_\phi = 1$ in this plot.)

scale (e.g. $f_* \gtrsim 10^{-9}$ Hz for $v_\phi \simeq 18$ TeV). As can be seen in figure 2.5, a region of parameter space with low v' falls within reach of gravitational wave observatories probing the low frequency regime such as SKA [107], NANOGrav [108], and the EPTA [109]. As ϵ is increased, the collapse of the domain wall network occurs earlier, further suppressing the current value of the energy density in gravitational radiation by the corresponding redshift factor.

Our discussion so far applies in the context of global discrete symmetries provided that they either do not descend from a continuous symmetry, or that, if they do, the symmetry breaking scale of the continuous factor is above the reheating temperature, so that a network of cosmic strings is not formed in the early universe. On the other hand,

if the reheating temperature is larger than the scale set by the tension of the strings, μ, then a string network will be formed first, with the strings later joined by domain walls. The entire string-wall network now evolves together, and the problem features an additional time scale, given by

$$t_* \sim \frac{\mu}{\sigma} \,. \tag{2.82}$$

At $t \sim t_*$, the force per unit length on a string of radius $R(t) \sim t$, given by $\sim \mu/t$, becomes comparable to the wall tension. The system then becomes dominated by the tension of the domain walls, causing the network to shrink, and break down into pieces that will further decay into gravitational waves (or, potentially, also massive particles, depending on their size and the relevant particle spectrum) [100, 101]. If this timescale is shorter than $t_{\text{coll.}}$, the earlier destruction of the network of defects could move any potential gravitational wave signal into an unobservable regime.

The discussion of the previous paragraph is especially relevant if parity is instead realized as a gauge symmetry, for which explicit breaking is no longer allowed. Naïvely, one would hope that the gauge case would be cosmologically more benign: the gauge equivalence of the two vacua makes them no longer distinct, eliminating the topological stability of the domain walls. Indeed, domain walls can be destroyed by a process in which a string loop is nucleated on the wall, further growing to destroy the entire defect. However, the corresponding nucleation probability is proportional to $e^{-\mu^3/\sigma^2}$ [112], which will be exceedingly small for any reasonable separation of scales between the string and wall tensions, therefore rendering gauge domain walls effectively stable. It is therefore crucial that the reheating temperature is above the string tension scale, so that a string network is formed that can later result in the entire collapse of the subsequent string-wall network. The cosmological implications, as well as potential gravitational wave signatures, of a discrete parity symmetry that is gauged will be further explored in

future work.

2.6 Conclusions

The strong CP problem remains one of the great naturalness problems of the Standard Model, and is perhaps the most compelling in light of its resistance to straightforward anthropic explanations. Fully satisfying solutions to the problem remain elusive given the expected violation of global symmetries in a theory of quantum gravity, which demands extensive effort to protect the Peccei-Quinn symmetry underlying axion-based approaches. In this work we have pursued a possibility that is more transparently robust against the effects of quantum gravity, revisiting parity-based solutions to the strong CP problem. Our approach highlights the experimental signatures associated with the most natural regions of parameter space in these models, as well as ancillary signatures that are dependent upon the detailed mechanism of parity breaking.

The notion of naturalness within this parameter space is governed by the tuning associated with the separation of scales of $SU(2)_L$ and $SU(2)_R$ breaking, which are related by generalized parity. Given this tuning, "see-saw" vector-like masses for the $SU(2)$-singlet fermions play a key role in allowing the scale of $SU(2)_R$ breaking to be lowered toward its most natural value consistent with experimental constraints. Within this framework, the LHC provides the strongest test of natural parity-based solutions to the strong CP problem, probing the scale of $SU(2)_R$ breaking through searches for W' and Z' vector bosons as well as vector-like quarks and additional Higgs bosons. This leaves parity solutions tuned at the $\sim 10^{-3}$ level, which while not fully natural remains a significant improvement in explaining the observed $\bar{\theta} \lesssim 10^{-10}$. The extended reach for heavy resonances at future colliders such as FCC-hh will decisively test these parity solutions at the level of $\sim 10^{-5}$ tuning. Constraints on new sources of flavor violation

play a complementary role, with additional sensitivity to the scale of vector-like fermions and the underlying model of flavor.

The detailed mechanism of parity breaking gives rise to additional signatures within reach of near-future tabletop experiments and gravitational wave observatories. Soft parity and CP-violating terms give rise to EDMs for elementary fermions at one-loop, both quarks and charged leptons, which provide a pathway to discovery in precision searches for CP-violation in molecular systems. Spontaneous violation of parity and CP through the vev of a pseudo-scalar gives rise to additional one-loop contributions to $\bar{\theta}$, which provides an additional pathway to discovery and already requires the source of parity violation to be sequestered from the quark sector (albeit in a technically natural way). The expected violation of global symmetries in a theory of quantum gravity further shapes the viable parameter space and potential experimental signatures through the impact of various Planck-suppressed operators whose form depends on the underlying parity-breaking mechanism. If parity is a global symmetry that is broken both spontaneously (by a pseudo-scalar vev) and explicitly (by gravitational effects), collapse of the domain wall network associated with the spontaneous breaking of parity can generate a gravity wave signal accessible at low-frequency gravitational wave observatories. In this respect, the violation of global symmetries by gravitational effects is a feature of parity-based solutions to the strong CP problem, rather than a bug. Taken together, these experimental opportunities warrant further exploration of generalized parity as a solution to strong CP.

Chapter 3

Dark Matter in a Mirror Solution to the Strong CP Problem

3.1 Introduction

The nature of DM and the strong CP problem are two of the most compelling puzzles of particle physics. In the last decades, cosmological observations of the cosmic microwave background (CMB), distant supernovae, large samples of galaxy clusters, and baryon acoustic oscillations have firmly established a standard cosmological model in which DM accounts for about 85% of the matter content of the Universe, and about 27% of the global energy budget [113]. Another puzzle stems from the observational absence of the neutron electric dipole moment, $d_n < 10^{-26}$ e $\cdot$ cm [114], which constrains the amount of CP violation in the strong interactions. (C stands for charge conjugation, P for spacetime parity and CP for their combination.) In the QCD Lagrangian, there is only one CP violating term [115],

$$\mathcal{L} \supset \bar{\theta}_{\rm QCD} \frac{g_s^2}{32\pi^2} G^a_{\mu\nu} \tilde{G}^{a,\mu\nu} \,, \tag{3.1}$$

where $G_a^{\mu\nu}$ is the gluon field strength tensor, g_s the strong coupling constant, $\tilde{G}_{a,\mu\nu} \equiv \frac{1}{2}\epsilon_{\mu\nu\alpha\beta}G_a^{\alpha\beta}$ and $\bar{\theta}_{\rm QCD}$ is an angle $\in [0, 2\pi]$. In the SM, $\bar{\theta}_{\rm QCD}$ combines the bare θ-angle $\theta_{\rm QCD}$ with an anomalous contribution from the quark mass matrix M into a flavor-invariant quantity, $\bar{\theta}_{\rm QCD} \equiv \theta_{\rm QCD} + \arg\det(M)$. The constraint above results in the upper limit $\bar{\theta}_{\rm QCD} \lesssim 10^{-10}$ [116–118], so that $\bar{\theta}_{\rm QCD}$ is by far, and inexplicably, the smallest dimensionless parameter of the SM. This strong CP problem is made even more surprising by the fact that weak interactions violate CP with a phase of order unity.

Three approaches to this problem have received considerable attention in the literature: a massless colored fermion [119–122] which makes $\bar{\theta}_{\rm QCD}$ unphysical (and whose minimal incarnation is now excluded by lattice data [123]), spontaneously broken P or CP symmetries [21, 22, 25, 26] that fix $\bar{\theta}_{\rm QCD} = 0$ in the UV and rely on its extremely suppressed renormalization [71, 124–127], and a spontaneously broken global anomalous symmetry à la Peccei-Quinn [51, 52] which relaxes $\bar{\theta}_{\rm QCD}$ to 0 through the QCD-induced dynamics of the predicted QCD axion [53, 54]. The QCD axion additionally turns out to be a natural DM candidate [128–130]. (See Ref. [131] for a review on the QCD axion and a large set of references.)

It was recognized in the 1970s already that parity might solve the strong CP problem [24, 132], and several models have been analyzed in the literature since then (see e.g. [21–23, 28, 29, 133–140]). Those models require extending the gauge group and particle content of the SM. However, unlike the QCD axion, the presence of a good DM candidate is not guaranteed by a suitable choice of parameters and cosmological history, as P relates the additional couplings and masses to the SM ones so that there is little freedom to find an appropriate parameter space. Extra fields can be added to play the role of DM [134, 135], while within minimal models the possibility that DM is made out of the mirror neutrinos or mirror electrons predicted by parity has been explored in Refs. [141, 142] and [136], respectively. For mirror neutrinos, justifying the necessary

small parameters requires extra ingredients, while only non-thermal production mechanisms were known to work for mirror electrons. In this work, we study whether it is possible to thermally produce DM within a minimal setup where, as with the axion, its mass and abundance can be computed in terms of parameters already required to solve the strong CP problem.

Concretely, we analyze the model proposed by some of us in Ref. [143], based on a mirror copy of the SM. In the process, we automatically explore other setups studied in the literature [23, 136], to which our model reduces in the appropriate limit. Parity fixes almost all coefficients given the measured SM parameters, so that one can genuinely scan the full parameter space. Generically, the only DM candidate is the mirror electron e', i.e. the new fermion paired with the SM electron by parity. We argue that production from a dark sector in thermal equilibrium with the SM is not allowed by experimental constraints. Moreover, the only viable thermal production mechanism from the SM bath, that can adequately suppress the dangerous relic abundance of mirror up quarks u', is (sequential) freeze-in[1]. Mirror up quarks are electrically neutral and colored and hence can bind into fractionally-charged exotic hadrons. We perform a thorough numerical analysis and identify the range of model parameters where one finds consistent e' DM. This fixes all scales of the model within a few orders of magnitude, dramatically increasing the predictivity beyond the constraints from parity alone.

As we were completing this work, a different cosmological history for the same model was presented in Ref. [139], where e' DM is obtained through freeze-out and subsequent dilution, while fractionally-charged exotic hadrons are argued to be sufficiently suppressed. While the bounds on such hadrons [136] are based on fluxes that have uncertainties [145], they appear to exclude this cosmological scenario by several orders of

[1]Freeze-in of mirror electron DM has been considered in Twin Higgs models [144]. There, neutral naturalness does not lead to dangerous exotic hadrons and freeze-in through kinetic mixing is allowed.

magnitude, calling into question its viability.

The paper is organized as follows. In Sec. 3.2, we summarize the relevant features and mechanisms of the model proposed in Ref. [143] to solve the strong CP problem. In Sec. 3.3, we discuss the running of masses and gauge couplings in this model, while in Sec. 3.4 we study the induced kinetic mixing between the SM and mirror photon, which strongly impacts the direct detection (DD) prospects of the model. In Sec. 3.5 we explore whether this model provides a solution to the DM problem, and in Sec. 3.7 we comment on the compatibility between DM and the Higgs parity mechanism, which allows parity to be exact instead of softly broken in the UV without adding new degrees of freedom. Finally, we conclude in Sec. 3.8. In App. B.1, we compute the freeze-in distribution of mirror photons, allowing us to compute the e' DM abundance in Sec. 3.5.

3.2 Mirror Solutions to the Strong CP problem

Here, we summarize the model presented in Ref. [143].

The full SM gauge group is mirrored to $SU(3) \times SU(2) \times U(1)_Y \times SU(3)' \times SU(2)' \times U(1)'$, and the matter content is doubled to include mirror copies of the fermion and Higgs fields. One set of particles has the usual SM quantum numbers under $SU(3) \times SU(2) \times U(1)_Y$ and is a singlet of $SU(3)' \times SU(2)' \times U(1)'$, while the converse is true for the other set. Each of the two Higgs fields is responsible for the breaking of the electroweak sector of its own "world". This setup can solve the strong CP problem when the two worlds are related via a $\mathbb{Z}_2$ symmetry composed with the usual action of P, so that the (C)P-odd θ angles of $SU(3)$ and $SU(3)'$ satisfy the relation $\theta = -\theta'$. In Ref. [143], it was shown that breaking the two $SU(3)$ gauge groups to their diagonal subgroup, which is identified with $SU(3)_{\rm QCD}$, provides a solution to the strong CP problem, as one finds $\theta_{\rm QCD} = \theta + \theta'$. In the limit of exact or softly broken P in the UV, the two θ angles cancel completely. In

the IR, despite P being broken, $\theta_{\rm QCD}$ receives negligible corrections.

The mirror Higgs acquires a vacuum expectation value (VEV) $v' \gg v$, where v denotes the SM Higgs VEV, which breaks P spontaneously. A phenomenologically-viable vacuum is achieved either by a soft P breaking term in the tree-level scalar potential or by radiative corrections at the scale v' [29]. In the latter case, v' cannot be chosen arbitrarily and depends on the other couplings of the model. We discuss it further in Sec. 3.7. In the rest of this paper, we treat v' as an independent parameter, which can for instance be achieved through the aforementioned soft breaking of parity. In either case, the theory does not address the hierarchy problem. Moreover, in this paper, we consider the simplest $SU(3) \times SU(3)' \to SU(3)_{\rm QCD}$ breaking mechanism, provided by the VEV of a bifundamental scalar field Σ. Although the solution to the strong CP problem does not depend on the specific breaking mechanism (see Ref. [143] for additional examples), the phenomenology of the model is sensitive to it. This will be commented on when necessary.

The interest of mirror solutions to the strong CP problem stems from their simplicity and predictivity: since the gauge and Yukawa couplings are fixed by P, the only two free parameters with respect to the SM are v' and v_3, the energy scale at which $SU(3)\times SU(3)'$ is broken. The constraints on the parameter space of this setup from collider searches and from the requirement that it solves the strong CP problem are shown in Fig. 1 of Ref. [143]. The allowed region roughly reads

$$2 \cdot 10^8\,\text{GeV} \lesssim v' \lesssim 10^{14}\,\text{GeV}; \quad v_3 \gtrsim 5\,\text{TeV}. \tag{3.2}$$

3.3 Masses, couplings and their running

As alluded to above, parity forces the gauge and Yukawa couplings in the two worlds to be equal at scales above v',

$$g_{G'} = g_G \ , \quad y_{f'} = y_f \ , \tag{3.3}$$

where $G = 1, 2, 3$ denotes a simple factor of the SM gauge group of coupling g_G and similarly for the mirror gauge couplings $g_{G'}$, while y_f are the SM Yukawa couplings in the up, down and charged lepton sectors and similarly for the mirror Yukawa couplings $y_{f'}$. The breaking of parity will generate deviations from these relations below v' in a calculable fashion. Knowing the precise values will be of extreme importance for the discussion of DM below, so we discuss here the renormalization group equations (RGEs) of our model.

For a choice of v' and v_3, the values of the parameters at all scales and in both worlds can be found. Indeed, the solution to the RGEs is fully fixed by the following constraints:

1. Matching to the measured SM parameters at low energies,

2. Continuity at v_3, except for
$$\frac{1}{\alpha_{\rm QCD}} = \frac{1}{\alpha_s} + \frac{1}{\alpha'_s} \tag{3.4}$$
(plus possible threshold corrections at loop level), where $\alpha_s^{(\prime)} = \frac{g_s^{(\prime)2}}{4\pi}$ with $g_s^{(\prime)}$ the (mirror) color coupling constant above v_3,

3. Parity at v', i.e. the relations of Eq. (3.3).

For given v' and v_3, these constraints constitute sufficient boundary data to fix all integration constants in the RGEs. We focus on the one-loop RGEs [146] and use the values in the modified minimal subtraction ($\overline{\rm MS}$) scheme given in Ref. [147]. We have checked that using the two-loop RGEs for the strong couplings affect the result at the percent

level, yielding a smaller source of uncertainty than that on the SM up quark mass (see below). The same applies to the use of the pole mass instead of the $\overline{\text{MS}}$ mass. In this paper, we use the latter everywhere. We also include the effect of the bifundamental scalar field Σ on the RGEs. For simplicity, we assume that all components of that scalar field acquire masses of order v_3 (this can be achieved upon suitably choosing the parameters of its scalar potential).

The most straightforward case is when the lightest colored mirror fermion, namely the mirror up quark u', is heavy. Specifically, when $m_{u'} \geq v_3$. In this case, the modification of the SM RGEs is minimal all the way to v': One simply needs to replace $\alpha_{\text{QCD}} \to \alpha_s$ and add the contribution of three flavors of fundamental scalars to the $SU(3)$ running. One can then pick a random value of $\alpha_s(v_3)$ (which lies between[2] $\alpha_{\text{QCD}}(v_3)$ and $2\alpha_{\text{QCD}}(v_3)$), run the SM world couplings to v', fix the mirror world parameters at v' and run them down to v_3 to check whether the relation of Eq. (3.4) holds. Spanning over all (or many) choices of $\alpha_s(v_3)$, one can identify the appropriate value.

When $m_{u'} < v_3$, the situation is more intricate. The boundary values for the mirror quark masses depend on the SM quark masses, which depend on the running of the QCD coupling constant, which itself depends on the mirror quark masses and their contribution to the β function. In this case, one needs to numerically solve the full system of RGEs of the two coupled mirror worlds. In practice, we implement this for all values of $m_{u'}/v_3$, and we cross-check the results with the previous method when applicable. We checked that our numerical precision is such that the constraints imposed by matching at v_3 and parity at v' are satisfied at the sub-per-mille level. We illustrate some results of this procedure in Fig. 3.1. In the right panel, we see that the loop correction to the ratio of $m_{u'}$ to $m_{e'}$, the mirror electron e' mass, depends on v_3 and can be noticeable, which will

[2]The lower bound directly follows from Eq. (3.4), while the upper bound also takes into account the fact that the mirror gauge coupling runs faster at lower energies, since the mirror fermions are heavier. Therefore, since the two gauge couplings are equal at v', $\alpha_s(v_3) < \alpha'_s(v_3)$.

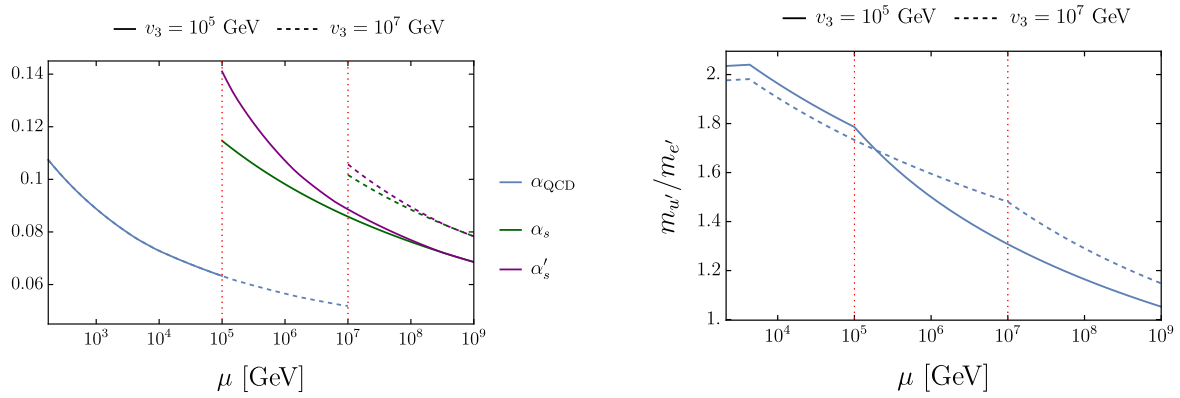

Figure 3.1: $\overline{\rm MS}$ running parameters for $v' = 10^9$ GeV and two choices of v_3, to which the red vertical lines correspond. Left panel: running gauge couplings of the color groups. Right panel: ratios of the u' and e' masses. The curves stop at $\mu = m_{e'}(\mu)$, and the running of $m_{u'}$ is frozen below $\mu = m_{u'}(\mu)$.

prove crucial when we discuss DM later on.

The outputs of this procedure are the running gauge couplings and fermion masses. They are used as inputs in a second step in which we compute the running kinetic mixing and Higgs quartic, which we discuss below. The RGEs of the gauge couplings and of the Higgs potential have been studied using similar techniques in Ref. [139].

A further comment on the input value for m_u is in order. In the following, we find the DM production to be extremely sensitive to $m_{u'}/m_{e'}$, whose value is obtained from the low energy determination of the SM up-quark and electron masses, as described above. The electron mass is known with the astonishing precision of ~ 0.1 ppb, while the PDG reports $m_u(2\,\text{GeV}) = 2.16^{+0.49}_{-0.26}$ MeV [148]. This is the largest source of uncertainty in our result, as discussed in Sec. 3.5. The central value is used for the right panel of Fig. 3.1. Note, that a recent combination of lattice determinations of m_u reports a smaller uncertainty of $\sim 4\%$ [123], but is not without controversy and remains under active investigation.

3.4 Kinetic Mixing

Parity is compatible with a kinetic mixing between the gauge boson of $U(1)_Y$ and its mirror copy,

$$\mathcal{L} \supset \epsilon F^{\mu\nu} F'_{\mu\nu} \,. \tag{3.5}$$

This term was not relevant for the solution to the strong CP problem in Ref. [143], but it constitutes an important source of cosmological constraints on the parameter space of the model. Indeed, once this term is introduced, all fermions charged under $U(1)'$ become charged also under $U(1)_Y$. If this is the case for DM, very stringent constraints on its charge from direct detection experiments apply. Anticipating the discussion in Sec. 3.5, the mirror electron is the only DM candidate of the model and kinetic mixing plays an important role in assessing its viability.

Removing the tree-level contribution to the kinetic mixing is a reasonable assumption. For instance, in UV completions where one $U(1)$ belongs to a larger gauge group, the interaction term in Eq. (3.5) is not allowed by gauge invariance. Therefore, it vanishes above the scale at which the two $U(1)$ gauge theories emerge. However, it will be generated at loop-level below that scale. In the current setup, the first loop-level contribution to this term arises at 4-loop [136], and could be further suppressed by an appropriate arrangement of the energy scales of the model, which we will discuss next. Nevertheless, due to the tight constraints, it is worth studying in detail the 4-loop contribution to ϵ, as it depends only on the two free parameters of the model, v' and v_3.

The leading diagram is shown in Fig. 3.2. As noted in Ref. [136], the renormalization group equation of the kinetic mixing parameter can be read off from the four-loop beta function of QCD [149], yielding

$$\frac{\mathrm{d}\epsilon}{\mathrm{d}\ln\mu} = \frac{e^2 g_s^6}{(4\pi)^8} \left(-\frac{1760}{27} + \frac{1280}{9}\zeta(3) \right) \sum_{ij} q_i q'_j \,. \tag{3.6}$$

It is clear that this contribution is only present below v_3, as above that scale the internal gluons do not couple simultaneously to SM and mirror quarks. Above v_3, we need higher-loop diagrams to connect γ and γ', which are further suppressed. Moreover, below their

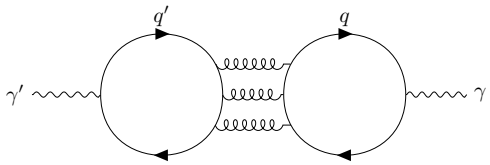

Figure 3.2: 4-loop diagram contributing to kinetic mixing.

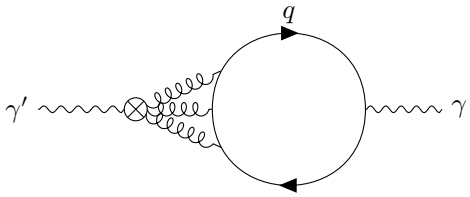

Figure 3.3: Diagram contributing to kinetic mixing below the lightest mirror quark mass.

masses, the mirror quarks are integrated out at one-loop and the three-loop diagram of Fig. 3.3 arises, where the effective vertex on the left connects one mirror photon to three gluons[3] and is therefore suppressed by $1/m_{u'}^4$ (note that, due to parity, $m_{u'}$ is the smallest mass scale among the UV fields). Dimensional analysis dictates that such contributions must come with an extra factor $\sim (\frac{p}{m_{u'}})^4$, where p is the largest momentum flowing in the quark loops, while one retains the $e^2 g_S^6$ and four-loop suppression (one for the one-loop matching at $\mu \sim m_{u'}$ and three for the diagram in the IR theory). Therefore, one can safely neglect the running below $m_{u'}$. In conclusion, 4-loop contributions to the kinetic mixing are relevant below the scale v_3 and when there is at least one mirror quark below that scale.

[3]There is no effective operator connecting one mirror photon to any combination of two gauge fields, due to symmetry (see the discussion on the "X^3" class of operators in section 5 of [150]).

3.5 Dark Matter

Given the unbroken mirror QED symmetry of the model of Ref. [143], it is natural to expect the presence of a good DM candidate. However, it turns out that this happens only in a specific region of parameter space and through a non-standard production mechanism. This is the topic of this section.

Although little is known about the particle nature of DM, a good DM candidate must:

1. Have a relic abundance today that matches the observed value $\Omega_{DM} = \rho_{DM}/\rho_{\rm crit} = 0.265(7)$, where $\rho_{\rm crit} = 8.5(1) \times 10^{-30}\,{\rm g\,cm^{-3}}$ is the critical density of the Universe [148],

2. Be cosmologically stable so as to agree with current experimental observations [151–154],

3. Become non-relativistic well before matter-radiation equality, as required by the formation of large scale structures in the Universe [155–157],

4. Have zero, or very small, electromagnetic charge, as required by searches for stable charged particles [158, 159],

5. Have limited self-interactions, constrained by the observed DM halo profiles, cluster collisions and the CMB spectrum [160].

In the present model, all mirror particles are electrically neutral, up to kinetic mixing. Through mirror Yukawa and electroweak couplings fixed by parity, the mirror particles have decay channels similar to their SM counterparts, hence the only stable massive particles are the mirror electron e' and up quark u'. They cannot decay to the SM because of the unbroken mirror electromagnetic charges that they carry. The mirror photon remains massless and contributes to dark radiation (by an amount far below

current limits in the scheme to be described), whereas mirror neutrinos are heavy but quickly decay to the SM unless further model building is invoked, as we discuss below. Finally, the physical components of the bifundamental scalar field Σ which breaks the color groups are generically of mass $\sim v_3$ and quickly decay to quarks and gluons [161].

Being the lightest stable mirror fermion, and charged only under mirror gauge groups, the mirror electron is the best DM candidate in this class of models. However, in Sec. 3.3 we have seen that the u' has a mass close to the e', and so its abundance has to also be carefully considered. This will turn out to be of utmost importance for the viability of a DM production mechanism, given the bounds on the u' relic density. Another important constraint on the cosmology stems from the fact that the potential of Σ displays an accidental $\mathbb{Z}_3$ symmetry, leading to domain walls (DWs) if the universe is reheated above the phase transition temperature $\sim v_3$. (This symmetry appears to be accidental to all mass dimensions, preventing one from introducing a bias to collapse the domain walls.) However, over most of the parameter space, this requirement is weaker than the one associated with the relic density of u'.

3.5.1 Bounds on mirror quarks

An extensive discussion of the cosmological history of the mirror quarks in these models can be found in Ref. [136]. The heavier mirror quarks decay into the lighter ones, while the latter hadronize after the QCD phase transition, forming bound states by combining with other colored particles. Rearrangements mediated by scattering processes then lead to the presence of two kinds of exotic bound states:

1. Hadrons made of u' and SM quarks: $u'qq$, $u'u'q$, $u'\bar{q}$. We denote these by h'.

2. Hadrons made of three u'.

The abundance of these states, relative to the abundance of u', has been estimated in Ref. [136],

$$\begin{aligned} Y_{u'u'u'} &\simeq Y_{u'} \times \begin{cases} Y_{u'}/Y_{\rm crit} & Y_{u'} < Y_{\rm crit} \\ 1 & Y_{u'} > Y_{\rm crit} \end{cases} , \\ Y_{h'} &\simeq Y_{u'} \times \begin{cases} 1 & Y_{u'} < Y_{\rm crit} \\ Y_{\rm crit}/Y_{u'} & Y_{u'} > Y_{\rm crit} \end{cases} , \\ Y_{\rm crit} &\simeq 2 \times 10^{-13} \left(\frac{m_{u'}}{\rm GeV}\right)^2 Y_{\rm DM} \, . \end{aligned} \tag{3.7}$$

While bound states made of only u' can in principle be a component of DM, the abundance of h' is strongly constrained by searches for nuclear and electric recoil at deep underground detectors [162–164], and by tracks of ionizing particles in bulk matter, on earth [165–168] as well as in meteorites [169]. Again we refer the reader to the discussion in Ref. [136].

Collider bounds on fractionally charged heavy stable states give $m_{u'} \gtrsim 1.5\,\text{TeV}$ [143], while bounds from higher dimensional operators that can spoil the solution to the strong CP problem imply $m_{u'} \lesssim 10^6$ TeV. In this mass range, one obtains $Y_{h'} \lesssim [10^{-14} - 10^{-8}]\ Y_{\rm DM}$ from the bounds of the MACRO [166], ICRR [167] and Baksan experiments [168], as well as $Y_{h'} \lesssim m_{u'}/\text{GeV}\, 10^{-15}\ Y_{\rm DM}$ from searches of ionizing tracks in the Hoba meteorite [169]. The lower bound on $m_{u'}$ and the definition of $Y_{\rm crit}$ imply that $Y_{\rm crit} \gtrsim 4 \times 10^{-7}\ Y_{\rm DM}$, hence $Y_{h'} < Y_{\rm crit}$ given the bounds above. This shows that the only viable scenario is

$$Y_{u'} < Y_{\rm crit}, \quad Y_{h'} \simeq Y_{u'}, \quad Y_{u'u'u'} \simeq \frac{Y_{u'}^2}{Y_{\rm crit}} \, . \tag{3.8}$$

In the literature, it has been debated whether strongly interacting particles can reach terrestrial experiments such as MACRO, ICRR and Baksan, because of their interactions in the Earth's atmosphere and crust. However, even assuming a geometrical cross

section for DM-nucleus scattering (which is the largest possible value for such a cross section[4] [170]), it has been shown that this is the case for neutral DM bound states [171] as well as charged ones [136]. The determination of the h' flux in the galactic disk from its charged massive particle nature [145] is not straightforward, hence the systematic errors on the bounds for $Y_{h'}$ are difficult to estimate. Therefore, in the following, we show different exclusion scenarios. Nonetheless, the constraints are so stringent that, even taking a conservative stand, they play an important role in our results.

Concerning bound states made of only u', one finds from the relations above that $Y_{u'u'u'} \lesssim 2 \times 10^{-10}\, Y_{\rm DM}$, and the DM fraction of fully mirror bound states is negligible.

3.5.2 Bounds on e' DM

Since bound states of mirror quarks are constrained to provide only an extremely small contribution to DM, we are left with the possibility of mirror electron DM. We remind the reader that in this model the mirror electron is charged only under the mirror gauge groups $SU(2)' \times U(1)'$ and has a mass larger than $\sim 750\,\text{GeV}$[5]. Before studying the production mechanism in the early universe, we discuss the bounds from DM direct and indirect searches.

Being coupled with the mirror photon, the e' experiences a long-range force. However, its mass is too heavy to appreciably self-scatter and disrupt the DM halo profile [160] nor the spectrum of the CMB. On the other hand, the kinetic mixing discussed in Sec. 3.4 leads to an interaction with SM particles and can produce an observable signal. The strongest bound comes from searches for DM-nucleus scattering. The results of the

[4] A larger cross section can be obtained through long range interaction or, in the case of composite DM, by the presence of resonances or threshold bound states [170].

[5] This limit is obtained from the bound on the mirror quark masses which, due to parity, implies a bound on the mirror electron mass with a mild dependence on the parameters of the model.

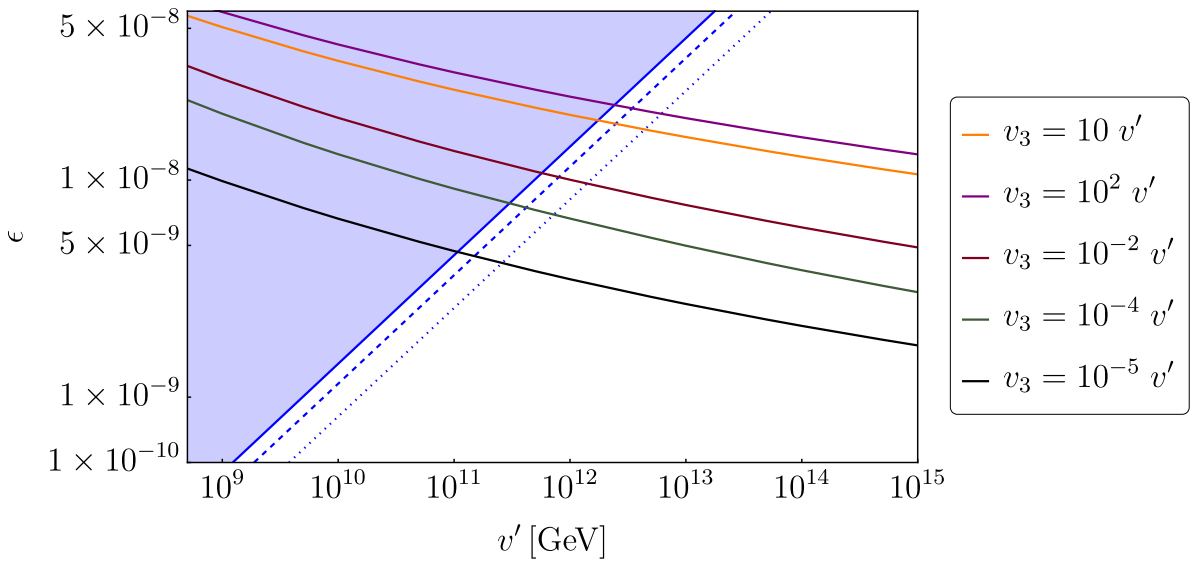

Figure 3.4: Limit on ϵ vs v' from DM direct searches (blue region) from XENON1T [163] (solid line), XENONnT [172] (dashed line) and LUX-ZEPLIN [173] (dotted line) together with predictions for different values of v_3/v'.

experiment XENON1T [163] can be recast [136] into the bound

$$m_{e'} > 10^6\,\text{GeV}\left(\frac{\epsilon}{10^{-8}}\right)^2 . \tag{3.9}$$

Results from XENONnT [172] and LUX-ZEPLIN (LZ) [173] lead to similar bound, up to the respective replacement of 10^6 by 1.5×10^6 and 3×10^6. Since the mass of the mirror electron depends only on the VEV of the mirror Higgs with a very good precision, this limit translates into a bound on ϵ and v', as shown in Fig. 3.4. The same figure also shows the predictions of ϵ for different values of v_3, assuming no UV kinetic mixing as discussed in Sec. 3.4. As anticipated there, the smaller v_3, the weaker the bound on v'. For v_3 below the lightest mirror quark mass, the kinetic mixing is extremely suppressed and this bound disappears. For very low values of v_3 (above $m_{u'}$), we find the lower limit $v' \gtrsim 10^{11}\,\text{GeV}$. This complements the other bounds on the parameters of the model. In fact, Ref. [143] finds the bound $v' \lesssim 10^{14}\,\text{GeV}$ from higher-dimensional operators that can spoil the solution to the strong CP problem. Therefore, in the scenario $v_3 > m_{u'}$,

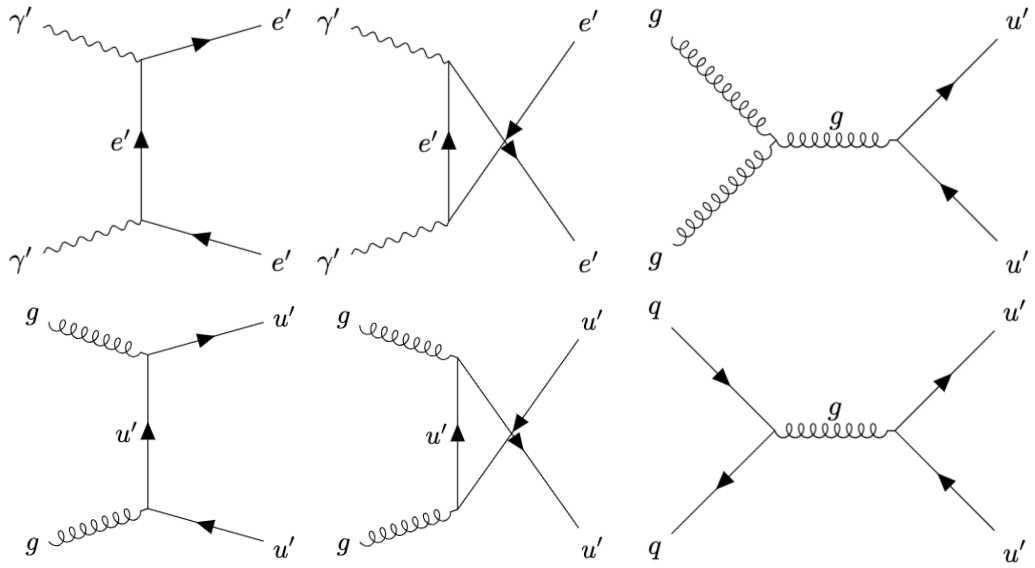

Figure 3.5: Most relevant tree-level diagrams for the production of e' and u' in the early universe when the temperature is smaller than v_3.

this leaves only 2-3 orders of magnitude for v'.

3.5.3 Freeze-out

The most relevant processes for the interaction of e' and u' with the thermal bath are shown in Fig 3.5 when the temperature is smaller than v_3. At temperatures higher than its mass, e' is efficiently produced by/annihilated into two mirror photons. In the freeze-out scenario, the resulting e' yield reads $Y_{e'} \simeq Y_{\rm DM}\left(\frac{v'}{10^8\ \rm GeV}\right)$ and the point in parameter space which reproduces the right DM abundance is already excluded by collider searches, as mentioned in Sec. 3.2. Refs. [139, 174] studied the possibility of heavier e', whose abundance is diluted trough the decay of the mirror neutrinos. However, we discussed the running of the mirror fermion masses in Sec. 3.3 and we saw that the ratio $m_{u'}/m_{e'}$ can be at most as large as ~ 2 for $m_{u'}$ close to v_3. This implies that the abundance of e' and u' would be similar. Indeed, for u' the mirror photon channel is also active, but it is even dominated by the production/annihilation through gluons or SM quarks. For

a temperature above v_3, the gluons become mirror gluons, and the bifundamental scalar Σ is also in equilibrium and can annihilate into a u' pair. It then appears to us that the bounds discussed in Sec. 3.5.1 are too strong and prevent the freeze-out scenario from being viable. Weakening the bounds on the u' abundance by a few orders of magnitude to account for the large uncertainties is not sufficient to change the conclusion.

3.5.4 Freeze-in

We therefore need to assume that the mirror fermions are never in thermal equilibrium and are produced through a freeze-in mechanism. If $m_{u'} > m_{e'} \gg T$, a factor of a few in the mass leads to an exponential suppression in the abundances, $Y_{u'}/Y_{e'} \sim \exp(\frac{m_{u'}-m_{e'}}{T})$, due to the Boltzmann suppression of the thermal bath particles which are energetic enough to produce e' and u'.

In the following, we assume that the reheating temperature at the end of inflation, T_R, is the highest temperature $T_{\max}$ reached in the cosmological history of the universe and is smaller than $m_{e'}$. We are thus adding a new parameter to the model, whose cosmology is now determined by three quantities: v', v_3, T_R. In addition, we need to assume that the inflaton does not directly decay to mirror fermions, or with a very reduced rate, and that it does not produce too many mirror photons either, as those would subsequently generate a population of mirror fermions[6]. None of this happens if we assume that the inflaton primarily decays to the SM sector, or that it mostly produces gluons and mirror gluons. The former is not in contradiction with exact parity due to the different electroweak VEVs in the two sectors, for instance if the inflaton couples to matter through a parity-symmetric Higgs portal [175]. For a P-odd inflaton, such a coupling and an inflaton VEV can generate a soft P breaking leading to asymmetric electroweak scales in the

[6]We note that an initial population of mirror photons is an interesting possibility, but we choose to focus here on the minimal, purely thermal DM production from the SM bath.

two sectors, even though a severe tuning is needed. We elaborate on our assumptions regarding reheating in Sec. 3.5.5.

In this setup, the only mirror species that can ever be in thermal equilibrium with the SM is the mirror photon, which, below v_3, interacts with the SM gluons through the 1-loop diagrams in Fig. 3.6. Again, the relevant production processes for e' and u' below v_3 are shown in Fig. 3.5. Above v_3, Σ thermalizes gluons and mirror gluons, so that the relevant diagrams become those of Fig. 3.6 and Fig. 3.5 where gluons are replaced by mirror ones.

γ' in thermal equilibrium

The cross sections for the processes in Fig. 3.6 read

$$\sigma_{gg\to\gamma'\gamma'} \simeq 8\times 10^{-8}\,\frac{g_s^4 e'^4 s^3}{m_{u'}^8} + O\left(\frac{s^4}{m_{u'}^{10}}\right), \tag{3.10}$$

$$\sigma_{gg\to g\gamma'} \simeq 7.5\times 10^{-8}\,\frac{g_s^6 e'^2 s^3}{m_{u'}^8} + O\left(\frac{s^4}{m_{u'}^{10}}\right), \tag{3.11}$$

where $s=(p_1+p_2)^2$ and $p_{1,2}$ are the four-momenta of the external gluons. In the center of mass frame $s=4E_1E_2$.

Following Ref. [176], the rate for these processes is[7]

$$\begin{aligned}\Gamma_{\gamma'} =& n_g \frac{\int \sigma v_{\rm Mol}\, e^{-E_1/T} e^{-E_2/T} d^3p_1 d^3p_2}{\int e^{-E_1/T} e^{-E_2/T} d^3p_1 d^3p_2} \\ \simeq& 1.5\times 10^{-3}\,\frac{(g_s^4 e'^4 + g_s^6 e'^2)\,T^9}{m_{u'}^8},\end{aligned} \tag{3.12}$$

and the condition for having γ' in thermal equilibrium is

$$\Gamma_{\gamma'} \gtrsim H \simeq \frac{T^2}{M_P} \Rightarrow T_R \gtrsim 26\,\text{GeV}\left(\frac{m_{u'}}{\text{TeV}}\right)^{8/7}. \tag{3.13}$$

[7] Note that we integrated over all energies in Eq. (3.12) although we used the simple formulae of Eqs. (3.10)-(3.11) which are not valid at energies larger than $m_{u'}$. The exponential suppression above $E\sim T$ ensures that the result is not noticeably affected. For instance, accounting for resonant $u'\bar{u}'$ bound states formation [177–180] yields an increase in the actual cross section that is comparable to that obtained from Eqs. (3.10)-(3.11) when $s\sim m_{u'}^2$. We checked that this region and larger energies essentially do not contribute to the result.

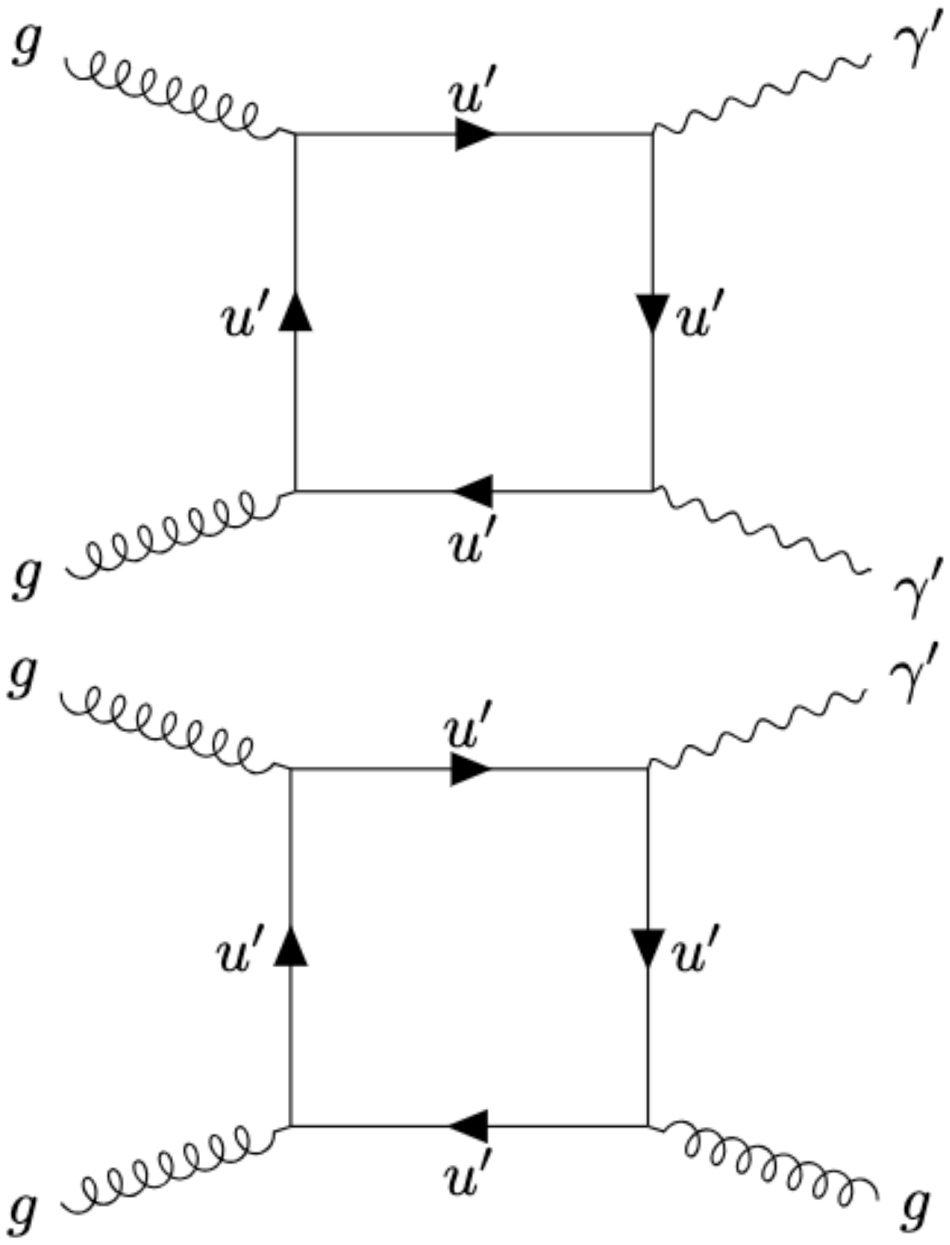

Figure 3.6: 1-loop portal between SM gluons and γ'.

The power of $m_{u'}$ indicates that T_R grows faster than $m_{u'}$ (i.e., than v') and tends to bring mirror fermions in thermal equilibrium, which is problematic as shown above. With this result, it is actually impossible to find a region of parameter space which gives the

correct relic abundance for e'. In fact, we find

$$x_{e'} \equiv \frac{m_{e'}}{T_R} \lesssim 38 \, \frac{m_{e'}}{m_{u'}} \left(\frac{\text{TeV}}{m_{u'}} \right)^{1/7} . \tag{3.14}$$

As mentioned above, collider bounds require $m_{u'} \gtrsim 1.5\,\text{TeV}$, implying that $v' \geq 2 \times 10^8$ GeV. For such values, we find numerically that the ratio $m_{e'}/m_{u'}$ is such that $x_e \lesssim 16$ all over the allowed parameter space, bringing the mirror electrons in thermal equilibrium with the mirror photons and overshooting the DM relic abundance by several orders of magnitude. (For instance, $v' = 10^9$ GeV and $v_3 = 10^5$ GeV yield x_e close to the maximal value, as can be seen in Fig. 3.1.) We verified this estimate solving the Boltzmann equation numerically, and we indeed found that the mirror electrons reach thermal equilibrium. Importantly, the result holds even assuming a large uncertainty $\mathcal{O}(20\%)$ on $m_{u'}$, in which case $x_{e'}$ remains smaller than ~ 22.

γ' out of equilibrium

To reduce the abundance of e' and even further of u', T_R needs to be lower than the limit in Eq. (3.13). This implies that also γ' are out of equilibrium and DM is produced via a sequential freeze-in process [181, 182]. In App. B.1 we describe how to solve the Boltzmann equation for the momentum distribution of the mirror photons. We were able to solve it analytically in the limit where the number of degrees of freedom is constant and for high-energy mirror photons (i.e. $E_{\gamma'} \gg T$). These are consistent assumptions: the SM degrees of freedom are all in the bath until much lower temperatures than T_R where most of the e' are produced, and making e' (such that $m_{e'} \gg T_R$ for freeze-in) requires highly energetic photons. Soft photons do not contribute to this process since they need to scatter with a very energetic one, whose number density is extremely suppressed. We checked the consistency of this argument, as detailed in App. B.1.

With the mirror photon distribution, we then numerically solve the Boltzmann equa-

tion for e' and u', making no further assumption and using the full non-thermal distributions of the mirror photons obtained in App. B.1. Fig. 3.7 and Fig. 3.8 show our results in the $v' - T_R/v'$ plane, for two benchmark cases: when $v_3 > v'$, and when v_3 has the smallest value allowed by collider searches ($v_3 = 5\,\text{TeV}$), respectively. We discuss the extrapolation of intermediate cases later on, while for $v_3 > v'$ the result is independent of v_3. (We fixed it to be $10\,v'$ in our code to generate Fig. 3.7. The actual value of $v_3 > v'$ only matters for kinetic mixing, whose RG running kicks in below v_3 as explained in Sec. 3.4.) The points which provide the right yield of e' are shown with a solid orange line. The blue solid and dotted lines show the experimental bounds on u', discussed in Sec. 3.5.1. Given the uncertainty on these bounds, we show two benchmarks: $Y_{h'} < 10^{-8}\,Y_{\text{DM}}$ (solid blue) and $Y_{h'} < 10^{-12}\,Y_{\text{DM}}$ (dotted blue). For comparison, the region where the mirror photons are in equilibrium is shown in green, confirming that it is incompatible with e' DM. Moving to the right of the plot, the ratio T_R/v' increases as well as the abundances of both e' and u'. A few values of $m_{u'}/m_{e'}$ are also shown along the DM line, from which one sees that the conservative bounds on the u' abundance give the constraint $m_{u'}/m_{e'} \gtrsim 1.6 - 1.7$ (with a small dependence on the value of v_3) to get a viable DM model. In the two figures, this translates to $v' \lesssim 5 \times 10^{10}\,\text{GeV}$ and $T_R \lesssim 5\,\text{TeV}$. Such a low reheating temperature raises the question of the baryogenesis mechanism at play in the early universe. We discuss this point in Sec. 3.6.

Anticipating the results of the numerical analysis for all values of v' and v_3, the prediction for the kinetic mixing in Fig. 3.4 then suggests that models with $v_3 \gtrsim m_{u'}$ do not contain a good DM candidate produced by freeze-in. This applies in particular to the models of Refs. [23, 136] (which can be obtained from ours in the $v_3 \to \infty$ limit), and it can be seen in Fig. 3.9, which shows the ratio $m_{u'}/m_{e'}$ as a function of v_3 and v'. The observed trend is understood as follows: the ratio is larger for smaller $m_{u'}$, i.e. smaller v', as the u' feels the effect of a larger strong coupling constant which grows in

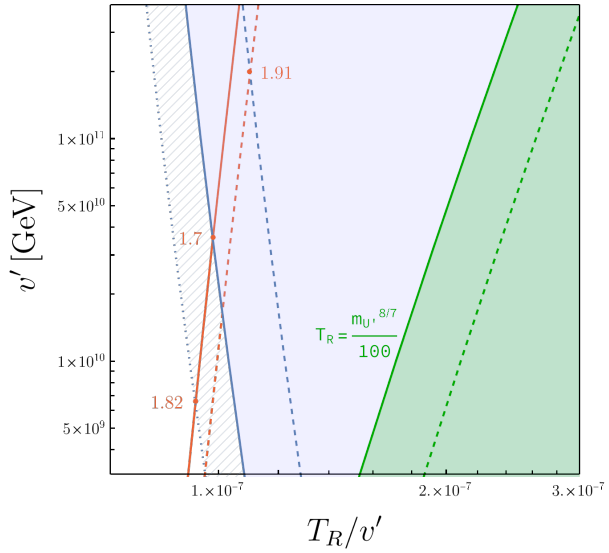

Figure 3.7: Values of T_R and v' reproducing the correct DM relic abundance for e' (solid orange line) in the scenario $v_3 > v'$. Requiring $Y_{h'} < 10^{-8}\, Y_{\rm DM}$ ($Y_{h'} < 10^{-12}\, Y_{\rm DM}$) rules out the blue region to the right of the solid blue (dotted blue) line. The region where the mirror photons distribution is thermal is shown in green. Dashed lines are analogous to solid ones, but using a 20% higher value for m_u.

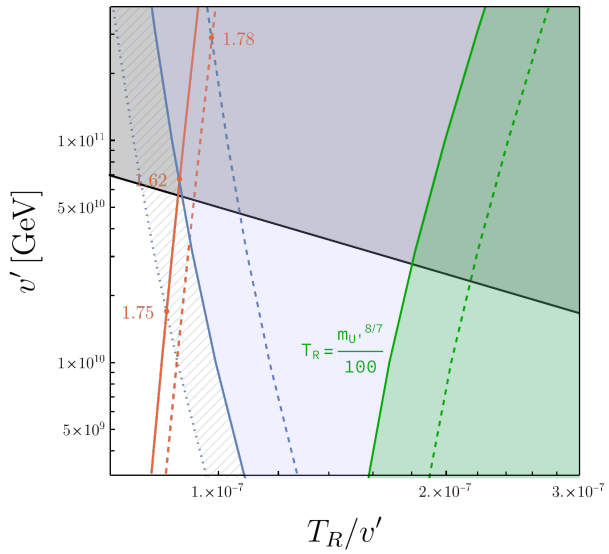

Figure 3.8: Values of T_R and v' reproducing the correct DM relic abundance for e' (solid orange line) in the scenario $v_3 = 5\,\text{TeV}$. Requiring $Y_{h'} < 10^{-8}\, Y_{\rm DM}$ ($Y_{h'} < 10^{-12}\, Y_{\rm DM}$) rules out the blue region to the right of the solid blue (dotted blue) line. The region where the mirror photons distribution is thermal is shown in green. Dashed lines are analogous to solid ones, but using a 20% higher value for m_u. The gray region is such that $T_R > v_3$ and DWs are expected to form in the Σ field.

the IR. Similarly, the smaller v_3, the larger its effect on the running of $m_{u'}$ since strong couplings are typically larger above v_3, as illustrated in Fig. 3.1. This effect competes with the fact that larger couplings make m_u run faster, and that the RG running of the strong couplings above v_3 is steeper than the one below. Numerically, we find that the ratio $m_{u'}/m_{e'}$ is maximized for points where $m_{u'} \sim v_3$, corresponding to the peaks in the contour plot of Fig. 3.9. The bound from DM direct detection is also shown as a shaded region. It arises from the kinetic mixing between γ and γ' as discussed in Sec. 3.4 and prevents the possibility of having DM for large values of v_3. We can make the trend

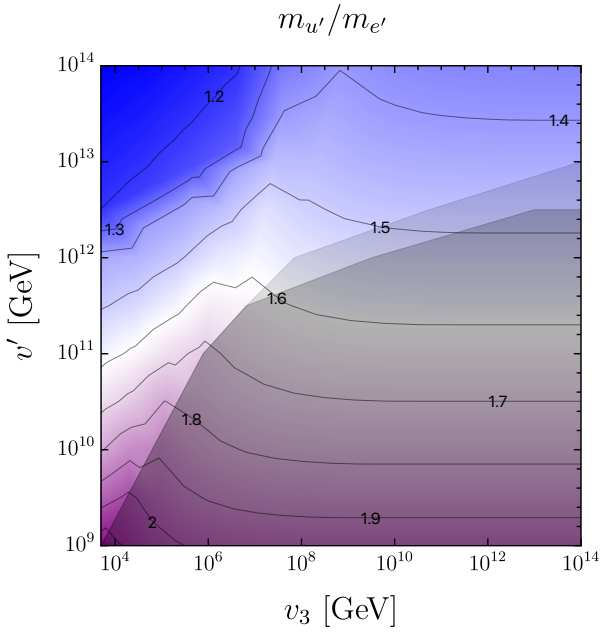

Figure 3.9: Ratio $m_{u'}/m_{e'}$ as a function of v' and v_3. In the shaded region, e' DM is excluded by the irreducible kinetic mixing and its impact on direct detection signals at Xenon1T (darker region) and LZ (lighter region).

displayed in Fig. 3.9 sharper by running our RGEs and Boltzmann codes for all values of v' and v_3. The result is shown in Fig. 3.10, where we therefore identify a precise region of parameter space where the model explains the observed DM abundance while satisfying

all other constraints:

Parameters for Viable DM

$$10^9\,\text{GeV} \lesssim v' \lesssim 10^{11}\,\text{GeV}\,,$$

$$5\times 10^3\,\text{GeV} \lesssim v_3 \lesssim 10^{-5}\,v'\,,$$

$$T_R \simeq 10^{-7}\,v'\,. \tag{3.15}$$

In this parity solution to the strong CP problem, the mirror electron is a good DM candidate with a mass in the range [5 TeV, 250 TeV]. We stress that the requirement of DM greatly reduces the large ranges for v' and v_3 that solve the strong CP problem, shown in Eq. (3.2).

We also investigated the impact of a large error, $\mathcal{O}(20\%)$, on the IR value of m_u. As commented in Sec. 3.3, this is the major source of uncertainty in our result. The results of our numerical analysis are represented by the dashed blue lines in Figs. 3.7 and 3.8. These lines can be roughly reproduced upon rescaling the bounds (solid blue lines) through a 20% shift in T_R/v'. (This amounts to assuming a $\mathcal{O}(20\%)$ shift on $m_{u'}$.) Note however that the DM line shifts to slightly higher T_R (a $\sim 5\%$ effect) (dashed orange lines), due to the dependence of the mirror photon production cross section on $m_{u'}$. Numerically, we find that increasing $m_{u'}$ by 20% extends the e' DM region to $v' \lesssim 2\times 10^{11}\,\text{GeV}$ and $T_R \lesssim 25\,\text{TeV}$. In addition, the lower bound on the ratio $m_{u'}/m_{e'}$ becomes larger as the bound on $Y_{u'}$ becomes stronger. However, for low v_3 most of this region is excluded by the presence of Σ domain walls.

On the other hand, lowering m_u by 20% gives stronger bounds, as expected, constraining v' between $10^9\,\text{GeV}$ and $10^{10}\,\text{GeV}$ and T_R below $10^5\,\text{GeV}$. Overall, this doesn't change the main picture of our result. The summary of results in the (v_3, v') plane can

be found in Fig. 3.10 below.

Finally, a non-thermal population of mirror photons remains until today, but it is far beyond observational prospects. Being more suppressed than thermal at T_R and being subsequently diluted by the SM thresholds, it gives a very small contribution to dark radiation, $\Delta N_{\rm eff} \leq 7 \times 10^{-6}$, which is saturated for the smallest reheating temperatures that allow e' DM.

3.5.5 Comments on Inflation and Reheating

Given the non trivial set of constraints which make DM viable in this model, we briefly comment on their impact on cosmological inflation and the subsequent reheating period. The following conditions are assumed in the predictive scenario for e' production presented above.

- At the end of reheating, when the universe enters a period of radiation domination, only the SM sector is populated and in thermal equilibrium at a temperature T_R in the range [0.1 - 5] TeV. Higher T_R would not generate an appropriate e' relic density (see Figs. 3.7 and 3.8).

- $T_R \leq \mathcal{O}(0.1)\, m_{e'}$ and $m_\phi \lesssim m_{e'}$, where m_ϕ is the mass of the inflaton. The former implies that our freeze-in calculation is applicable and that we do not freeze-out (hence overproduce) e' and u'. The latter prevents the inflaton from directly decaying to e' and u', or its high-energy decay products from scattering and producing e' or u'.

- During reheating, when the universe is dominated by the energy density of the inflaton field which is transferred to radiation, the temperature cannot be larger than T_R to avoid populating the mirror sector[8].

[8] Calling $T_{\rm max}$ the maximum temperature reached by the plasma during reheating, one may wonder if

Furthermore, we require $T_R < v_3$ to avoid domain walls for Σ. It is clear that the possibilities for a viable inflationary model are strongly constrained.

A thorough analysis of inflation and reheating models is beyond the scope of this work; here we just note a few promising possibilities. The requirement that the maximal temperature reached during reheating is T_R is fulfilled in the limit of instantaneous reheating, i.e. when the universe transitions from inflation to a phase of radiation domination in less than a Hubble time. For instance, this is achieved if the inflaton potential makes an almost discontinuous change from being very flat to be very steep. Alternatively, one can deal with a smoother reheating if the temperature of the SM bath increases (or stays constant) throughout this phase, reaching a maximum at T_R. Scenarios of this kind, using dissipative processes other than decays, have been discussed in Ref. [183].

3.6 Mirror neutrinos and leptogenesis

Finally, we discuss mirror neutrinos, which are electrically neutral. In our minimal model, the neutrino masses must arise from Weinberg-type operators. Due to parity, one finds two independent coupling matrices x_ν, x'_ν, the former being symmetric and the latter hermitian [141],

$$\begin{aligned}\mathcal{L}_\nu = &\frac{x_{\nu,ij}}{2\Lambda}(HL_i)(HL_j) + \frac{x^*_{\nu,ij}}{2\Lambda}(H'L'_i)(H'L'_j) \\ &+ \frac{x'_{\nu,ij}}{\Lambda}(HL_i)(H'L'_j)\ .\end{aligned} \tag{3.16}$$

a qualitatively similar DM production could be performed if $m_{e'} > T_{\max} > T_R$. In that case, e' freeze-in happens during (inflaton) matter domination, and a large enough e' relic density needs to be frozen-in in order to compensate the subsequent dilution. This implies that the Boltzmann factor $e^{-m_{e'}/T_{\max}}$ should be increased with respect to the $e^{-m_{e'}/T_R}$ evaluated in Sec. 3.5.4. However, the correlated $e^{-m_{u'}/T_{\max}}$ will also be increased, leading to u' overproduction. In summary, e' freeze-in at $T_{\max} > T_R$ does not work in this model.

Below v', H' is frozen to its vev and the mirror neutrinos acquire a mass matrix, $m_{\nu'}$, and Yukawa coupling matrix to LH, y_ν, of

$$m_{\nu'} = x_\nu^* \frac{v'^2}{\Lambda}, \qquad y_\nu = x'_\nu \frac{v'}{\Lambda}\,. \tag{3.17}$$

It is convenient to work in a basis where $m_{\nu'}$ is real and diagonal. Below the electroweak scale a mass matrix arises also for the SM neutrinos:

$$m_\nu = v^2 \left(\frac{m_{\nu'}}{v'^2} - y_\nu\, m_{\nu'}^{-1}\, y_\nu{}^T \right)\,. \tag{3.18}$$

The spontaneous breaking of parity at scale v' leads to a "direct" neutrino mass term proportional to $m_{\nu'}$ as well as a conventional "seesaw" mass term proportional to $1/m_{\nu'}$. If $x_{\nu,ij}$ and $x'_{\nu,ij}$ are comparable, as expected for example from approximate flavor symmetries, then the direct and seesaw contributions to the light neutrino masses will also be comparable, so that neutrino physics differs from that of just adding right-handed neutrinos to the SM.

With e' DM from sequential freeze-in, the cosmological effects of the mirror neutrinos are highly dependent on the coupling matrices x, x' and the scale Λ. For example, if the matrix elements of x, x' are order unity, and v' is of order 10^{10} GeV, the observed neutrino masses require the scale Λ to be of order 10^{15} GeV. In this case the mirror neutrinos have masses of order 10^5 GeV, well above the reheating temperature of 10^3 GeV, so the mirror neutrinos are not made in the thermal bath and play no cosmological role.

For other parameters the mirror neutrinos are light enough to be produced at reheating, and they decay via the Yukawa coupling $y_\nu\, \nu' LH$ to LH with decay rate

$$\Gamma_{\nu'_i \to \mathrm{SM}} = \frac{1}{8\pi} \left(y_\nu{}^\dagger y_\nu \right)_{ii} m_{\nu'_i}\,. \tag{3.19}$$

Could such decays lead to the cosmological baryon asymmetry via thermal leptogenesis [184]? If there is no degeneracy among mirror neutrinos, the answer is no: for this case

Ref. [140] finds that thermal leptogenesis requires $v' > 10^{12}$ GeV to avoid fine tuning in the SM neutrino masses; Fig. 3.10 shows this is inconsistent with e' DM, even allowing for a large uncertainty in the up quark mass. Furthermore, thermal leptogenesis without ν' degeneracy requires the lightest ν' to be heavier than 10^9 GeV, many orders of magnitude above T_R.

However, it is well known that degeneracy among ν' produces the observed baryon asymmetry for much lower values of $m_{\nu'}$ [185]. In this case, in theories with neutrino masses arising from (3.16), leptogenesis can occur for lower values of v': Fig. 5 of Ref. [140] shows that degeneracies in the range of $10^{-3} - 10^{-6}$, resulting from approximate flavor symmetries, gives successful leptogenesis throughout the allowed range of $v' \sim (10^9 - 10^{11})$ GeV required by e' DM.

An important feature of e' DM from sequential freeze-in, relevant for leptogenesis, is that it requires a very low reheat temperature, $T_R \sim 10^{-7} v' \sim (10^2 - 10^4)$ GeV. Interestingly, this is above the electroweak scale, so any lepton asymmetry produced can be processed to a baryon asymmetry via electroweak sphalerons. A key question is the size of $m_{\nu'_1}$ relative to T_R. If $m_{\nu'_1} \gg T_R$, thermal production of ν'_1, which occurs via the inverse of the decay process (3.19), will be highly Boltzmann suppressed, leading to a negligible lepton asymmetry. On the other hand, if $m_{\nu'_1} \lesssim T_R$ the produced lepton asymmetry will be strongly erased by rescatterings involving ν'_2, which is degenerate with ν'_1. Avoiding such strong washout requires reducing the Yukawa coupling coupling of ν'_2 to the point that, at these low values of $m_{\nu'_1}$, the production of the asymmetry in ν'_1 decays is too small, unless it is boosted by a degeneracy of at least 10^{-8}. This exceeds the natural limit of degeneracy in this theory, 10^{-6}, set by radiative corrections from the tau coupling [140].

Thus the only possibility is that $\nu'_{1,2}$ have masses close to T_R, but sufficiently above that a near thermal abundance of ν'_1 can be produced at reheating, while rescattering via

virtual ν_2' leads to little washout of the asymmetry. We conclude that thermal leptogenesis may occur in the same minimal theory where e' from sequential freeze in accounts for dark matter, but only if $\nu_{1,2}'$ are highly degenerate with mass several times larger than $T_R \sim 10^{-7} v' \sim (10^2 - 10^4)$ GeV.

Other possibilities for leptogenesis exist. The effective theory below v' may contain the coupling $\phi \nu' \nu'$, allowing the inflaton ϕ to decay to mirror neutrinos as well as SM particles. Non-thermal leptogenesis then occurs in ν' decays before they thermalize. Even though strong washout may be avoided by having $m_{\nu'} \gg T_R$, some degeneracy among ν' is still required for a sufficient baryon asymmetry. If the mirror neutrino is the lightest mirror fermion, inflaton decays to other mirror fermions may be kinematically forbidden, so that our previous calculations of the e' and u' abundance still applies. Another possibility is to augment the SM sector with gauge single fermions N and, by parity, SM$'$ is augmented by N'. In this case thermal leptogenesis can result purely in the SM + N sector via N decay, as in conventional minimal leptogenesis. The low reheat temperature again requires significant N degeneracy, but this is much less constrained by radiative corrections. The breaking of parity by v' can lead to N' coupling to ν' in pseudo-Dirac states much heavier than the reheat temperature.

3.7 Higgs Parity

Models that implement a parity solution to the strong CP problem break parity either explicitly, via a soft breaking term in the potential, or spontaneously, with a vacuum having $v' \gg v$ stabilized by Coleman-Weinberg radiative corrections [29]. The latter mechanism, dubbed "Higgs Parity", explains why the SM quartic coupling vanishes when evolved to a high energy scale, and implies that the new physics at this scale is that of parity restoration.

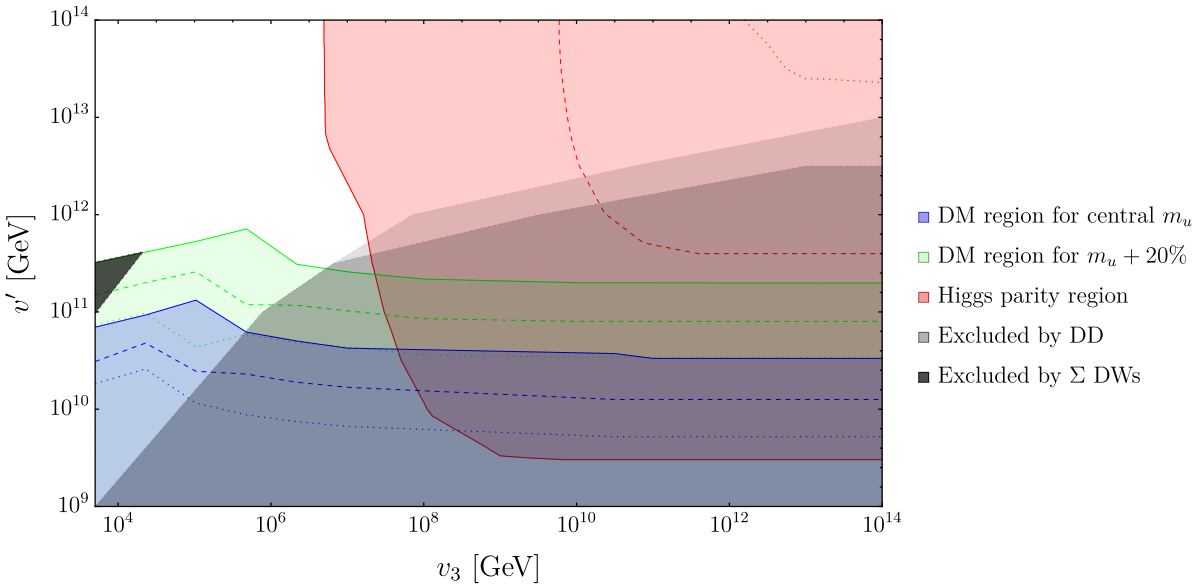

Figure 3.10: In the DM region, there exists a T_R for which e' constitutes all of DM while the u' relics pass the experimental constraints. The solid, dashed and dotted curves assume that those constraints take the form $Y_{h'} < xY_{\rm DM}$ with $x = 10^{-8}, 10^{-10}, 10^{-12}$ respectively. The dark (light) gray-shaded region is excluded by direct detection searches for e' DM by Xenon1T [163] (LZ [173]), and the black--shaded one by the presence of Σ domain walls. In the Higgs parity region, the Higgs quartic coupling vanishes at the scale v' for some values of α_S and m_t within their 2σ contours at m_Z. The dashed red line assumes central values for α_S and m_t, the solid line that they saturate their 2σ contours (in the direction of large m_t and small α_S), and the dotted one that m_t takes its central value while α_S is increased to saturate its 2σ contour.

As described in Refs. [136, 174], the tree-level parity symmetric potential for H and H' has an important feature: in the limit that $v \ll v'$ is imposed, an accidental $U(4)$ global symmetry emerges, with $\mathcal{H} = (H, H')$ forming a fundamental representation so that an $SU(2) \times SU(2)'$ subgroup is gauged. When $\mathcal{H}$ gets a VEV, $\langle \mathcal{H} \rangle = (0, 0, 0, v')$, the $U(4)$ is spontaneously broken to a $U(3)$, H' acquires a mass and at tree-level the SM Higgs arises as a massless Nambu-Goldstone boson. However, radiative corrections to the scalar potential (the leading contribution coming from the SM and mirror top quarks) break explicitly the $U(4)$ global symmetry, giving radiative contributions to the SM Higgs mass and quartic coupling. The large hierarchy v'/v results mainly from fine-tuning and

the SM Higgs quartic coupling λ at the scale v' takes a small value. At lower energies, quantum corrections within the SM renormalize λ so that it grows logarithmically. In this section, we discuss whether the condition $\lambda(v') \simeq -0.001$ [136] is compatible with the parameter space giving rise to e' DM.

The leading contributions to the RGE for λ below the scale v' are

$$\begin{aligned}\frac{d\lambda}{d\ln\mu} =& \frac{24\lambda^2 + 12\lambda y_t^2 - 6y_t^4}{16\pi^2} - \frac{\lambda\,(3\alpha_1 + 9\alpha_2)}{4\pi} \\ &+ \frac{3}{8}\alpha_1^2 + \frac{9}{8}\alpha_2^2 + \frac{3}{4}\alpha_1\alpha_2 \quad ,\end{aligned} \tag{3.20}$$

where y_t is the top Yukawa. The couplings y_t, α_1 and α_2 are computed at all scales, for any given v_3 and v', as discussed in Sec. 3.3. Hence, starting with the low energy value of λ, known from the Higgs boson mass, its value at higher energies is computed in terms of v_3 and v'. The input parameters, in the $\overline{\text{MS}}$ scheme, that we use here are [147]

$$\begin{aligned}\lambda(m_Z) &= 0.13947 \pm 0.00045\,, \\ \alpha_S(m_Z) &= 0.1179 \pm 0.0009\,, \\ m_t(m_Z) &= (168.26 \pm 0.75)\,\text{GeV}.\end{aligned} \tag{3.21}$$

The region of the (v_3, v') plane consistent at 2σ with $\lambda(v') = 0$ is shown in red in Fig. 3.10. That it shuts off for small v_3 is understood as follows. Low v_3 increases the strong coupling constants at a given scale, thereby enhancing the running of y_t, making it decrease faster in the UV than in the SM, hence reducing its impact on λ, which ends up not crossing zero when y_t runs too fast. One can see that, for the central values in Eq. (3.21), the condition $\lambda(v') = 0$ can be satisfied only for $v_3 \gtrsim 10^{10}\,\text{GeV}$ and $v' \gtrsim 6 \times 10^{11}\,\text{GeV}$. Varying the input parameters within their 2σ uncertainty, $\lambda(v') = 0$ can be obtained with v_3 as low as $\approx 10^7\,\text{GeV}$. These conditions for Higgs Parity are however incompatible with the results obtained for e' as DM candidate, using the central value of m_u, as shown by the blue region of Fig. 3.10. (Compare with Eq. (3.15)). As already discussed, a value for m_u larger by 20% would weaken the bounds in Fig. 3.7 and Fig. 3.8, which is however not

sufficient to reconcile Higgs parity and e' DM, as shown by the green region of Fig. 3.10. Higgs Parity in this model would therefore require that measured SM parameters deviate significantly from their central values. Fig. 3.10 neglects the threshold corrections to the quartic at v', but it is found to be very small and negative [136], making the situation slightly worse for Higgs parity. Decreasing the theoretical uncertainty on our prediction through the use of two-loop RGEs would be interesting as well, although we do not expect a different conclusion; we leave this for future work.

A non-minimal theory, with additional heavy fermion states coupled to the Higgs, with mass well below v', would allow $\lambda(v') = 0$ to be consistent with e' DM from sequential freeze-in. Thus parity could be spontaneously broken by the radiative Higgs Parity mechanism in non-minimal theories. Froggatt-Nielsen type theories of flavor contain such heavy fermions; if their masses are well below v' the scale of spontaneous flavor symmetry breaking is relatively low.

3.8 Conclusion

We have shown that certain models based on Parity solutions to the strong CP problem have a DM candidate already embedded in their particle spectrum. Having as a benchmark the model detailed in Ref. [143], where the full gauge group of the SM is copied. We have discussed the parameters of the model, stressing that Parity leaves little freedom, making the model very predictive. There are two free parameters in addition to the SM ones: the scale at which parity is broken, v', which is also the scale of mirror electroweak symmetry breaking, and the scale v_3 at which the two color group break to their diagonal subgroup. We computed the impact of broken parity on the RGEs of the model, and we then studied the unavoidable kinetic mixing between the SM $U(1)_Y$ gauge group and its mirror copy, which plays a relevant role for DM direct detection.

We identified the mirror electron e' as a good DM candidate, while the mirror up quark u' can form fractionally charged bound states with SM quarks after QCD confinement, being therefore excluded by several experimental searches. The closeness in mass of e' and u' strongly determines the DM production mechanism. We find that production of e' DM from the SM bath, with a sufficiently suppressed u' abundance, can occur via a sequential freeze-in mechanism through an out-of-equilibrium bath of mirror photons. This can occur only in the blue wedge-shaped region in the (v_3, v') plane of Fig. 3.10. Hence, the mass of the mirror electron is in the range $[5-250]$ TeV and the $SU(3)\times SU(3)'$ breaking scale is in the range $[5-500]$ TeV. At any point in this blue wedge region, the reheating temperature must be close to $10^{-7}v'$, and hence is predicted to be low, in the range $[0.1-5]$ TeV. We noted that knowing the mass of the up quark with good precision is crucial to make a robust prediction, therefore we commented on the possibility that the precision of lattice determinations is underestimated.

The blue wedge-shaped region of Fig. 3.10 for e' DM is not large and has several observational signals associated with it. Near the vertical edge of the wedge, at $v_3 = 5$ TeV, there are colored states associated with $SU(3) \times SU(3)'$ breaking that may be accessible to future collider experiments, as discussed in [143]. Close to the long sloped edge of the wedge, e' DM may be discovered by direct detection, via kinetic mixing of our photon with the mirror photon. Higher in the wedge v' increases and the u'/e' mass ratio falls; the abundance of the fractionally charged hadrons h, containing u' increases, leading to signals of this exotic DM component as discussed in [145]. Finally, throughout the wedge, e' DM is self interacting, with a long-range mirror electromagnetic force that is precisely predicted, and this may lead to future observational signals in large scale structure.

The low reheating temperature requires a late production of the cosmological baryon asymmetry. The theory satisfies two key requirements for leptogenesis: heavy neutral

fermions (ν') with Yukawa couplings to SM leptons, and a reheat temperature above the electroweak phase transition. Generating sufficient baryons at such late times requires ν' degeneracy, to enhance the asymmetry, and ν' masses somewhat larger than the reheat temperature.

Finally, we discussed whether the mechanism of Higgs Parity, which provides the spontaneous breaking of exact parity, can be realised in these models and is compatible with the parameters leading to a good DM candidate. We showed that the current central values for m_t and α_S clearly exclude this possibility. Stretching these values within their 2σ confidence intervals gets one closer to the region of parameter space where e' can be DM, but overlap also requires a large uncertainty in the up quark mass. If parity is broken spontaneously, either SM parameters are far from their central values, the Higgs Parity theory contains couplings of the Higgs to exotic fermions, or the breaking occurs first in some other sector of the theory and appears as soft breaking in the electroweak sector.

Chapter 4

Gravitational Waves from Stochastic Scalar Fluctuations

4.1 Introduction

The fluctuations observed in the cosmic microwave background (CMB) and large-scale structure (LSS) have given us valuable information about the primordial Universe. As per the standard ΛCDM cosmology, such fluctuations were generated during a period of cosmic inflation (see [15] for a review). While the microphysical nature of inflation is still unknown, well-motivated single-field slow-roll inflationary models predict an approximately scale-invariant spectrum of primordial fluctuations, consistent with CMB and LSS observations. These observations have enabled precise measurements of the primordial fluctuations between the comoving scales $k \sim 10^{-4} - 1$ Mpc^{-1}. However, the properties of primordial density perturbations are comparatively much less constrained for $k \gtrsim$ Mpc^{-1}. In particular, as we will discuss below, the primordial curvature power spectrum Δ^2_ζ can naturally be much larger at such small scales, compared to the value $\Delta^2_\zeta \approx 2 \times 10^{-9}$ observed on CMB scales [186].

Scales corresponding to $k \gtrsim \text{Mpc}^{-1}$ are interesting for several reasons. First, they contain vital information regarding the inflationary dynamics after the CMB-observable modes exit the horizon. In particular, they can reveal important clues as to how inflation could have ended and the Universe was reheated. An enhanced power spectrum on such scales can also lead to overabundant dark matter (DM) subhalos, motivating novel probes (see [187] for a review). Furthermore, if the enhancement is significant, $\Delta^2_\zeta \gtrsim 10^{-7}$, the primordial curvature fluctuations can induce a stochastic gravitational wave background (SGWB) within the range of future gravitational wave detectors [188]. For even larger fluctuations, $\Delta^2_\zeta \gtrsim 10^{-2}$, primordial black holes (PBH) can form, leading to interesting observational signatures [189, 190]. Given this, it is interesting to look for mechanisms that can naturally lead to a 'blue-tilted', enhanced power spectrum at small scales.

In models involving a single dynamical field during inflation, such an enhancement can come, for example, from an inflection point on the inflaton potential or an ultra-slow roll phase [191–195].[1] However, for any generic structure of the inflaton potential, a power spectrum that is blue-tilted at small scales can naturally arise if there are additional light scalar fields other than the inflaton field. One class of such mechanisms involves a rolling complex scalar field where the radial mode φ has a mass of order the inflationary Hubble scale H and is initially displaced away from the minimum [197]. As φ rolls down the inflationary potential, the fluctuations of the Goldstone mode $\propto (H/\varphi)^2$ increase with time. This can then give rise to *isocurvature* fluctuations that increase with k, i.e., a blue-tilted spectrum. This idea was further discussed in [198] to show how *curvature* perturbations can be enhanced on small scales as well, and lead to the formation of PBH. For further studies on blue-tilted isocurvature perturbations, see, e.g., [199–202]. Other than this, models of vector DM [203], early matter domination [204], and incomplete phase transitions [205] can also give rise to enhanced curvature perturbation at small

[1] See also [196] for PBH formation in a multi-field ultra-slow roll inflationary model.

scales.

In this work, we focus on a different mechanism where a Hubble-mass scalar field quantum mechanically fluctuates around the minimum of its potential, instead of being significantly displaced away from it (as in [197, 198]).[2] Hubble-mass fields can naturally roll down to their minimum since the homogeneous field value decreases with time as $\exp(-m^2 t/(3H))$, where m is the mass of the field with $m \lesssim H$. Given that we do not know the *total* number of e-foldings that took place during inflation, it is plausible that a Hubble mass particle was already classically driven to the minimum of the potential when the CMB-observable modes exit the horizon during inflation. For example, for $m^2/H^2 = 0.2$, the field value decreases by approximately a factor of 10^3, for 100 e-foldings of inflation prior to the exit of the CMB-observable modes. For any initial field value $\varphi_{\rm ini} \lesssim 10^3 \langle\varphi\rangle$, this can then naturally localize the massive field near the minimum $\langle\varphi\rangle$. However, the field can still have quantum mechanical fluctuations which tend to diffuse the field away from $\langle\varphi\rangle$. The potential for the field, on the other hand, tries to push the field back to $\langle\varphi\rangle$. The combination of these two effects gives rise to a non-trivial probability distribution for the field, both as a function of time and space.

We study these effects using the stochastic formalism [207, 208] for light scalar fields in de Sitter (dS) spacetime. In particular, such stochastic effects can lead to a spectrum that is blue-tilted at small scales. While we carry out the computation by solving the associated Fokker-Planck equation in detail below, we can intuitively understand the origin of a blue-tilted spectrum as follows. For simplicity, we momentarily restrict our discussion to a free scalar field σ with mass m such that $m^2 \lesssim H^2$. The fluctuation $\sigma_k(t)$, corresponding to a comoving k-mode, decays after horizon exit as $\sigma_k(t) \sim H \exp(-m^2(t - t_*)/(3H))$, where t_* is the time when the mode exits the horizon, $k = a(t_*)H$. We can rewrite

[2] For scenarios where the spectator field fluctuates around the minimum and gives rise to dark matter abundance, see, e.g., [206].

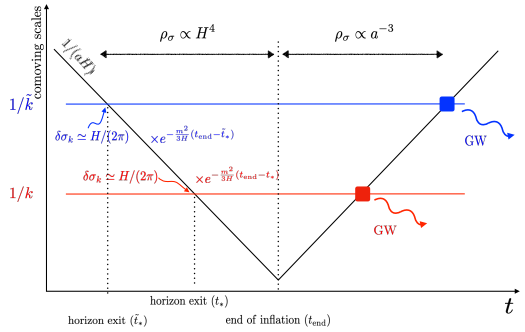

Figure 4.1: Schematic of the mechanism. The comoving horizon $1/(aH)$ decreases during inflation and increases after that. Any k-mode carries a fluctuation of order $H/(2\pi)$ at the time of mode exit. However, modes with larger k (red) exit the horizon later and encounters less dilution compared to modes with smaller k (blue), since $t_* > \tilde{t}_*$. Consequently, modes with larger k source stronger gravitational waves upon horizon re-entry (shown via square box). We also depict the fact that σ carries an energy density $\propto H^4$ during inflation, and dilutes as matter (for our benchmark choices) after inflation ends.

the above by noting that physical momenta redshift as a function of time via $k/a(t) = H\exp(-H(t-t_*))$. Then we arrive at, $\sigma_k(t) \sim H(k/(aH))^{m^2/(3H^2)}$. Therefore, the dimensionless power spectrum, $|\sigma_k|^2k^3 \propto (k/(aH))^{2m^2/(3H^2)}$ has a blue tilt of $2m^2/(3H^2)$. Physically, modes with smaller values of k exit the horizon earlier and get more diluted compared to modes with larger values of k, leading to more power at larger k, and thus a blue-tilted spectrum. This qualitative feature, including the specific value of the tilt for a free field, is reproduced by the calculation described later where we also include the effects of a quartic self-coupling. We summarize the mechanism in Fig. 4.1.

We note that if m is significantly smaller than H, the tilt is reduced and the observational signatures are less striking. On the other hand, for $m \gtrsim H$, the field is exponentially damped, and stochastic effects are not efficient in displacing the field away from the minimum. Therefore, it is puzzling as to why the particle mass, a priori arbitrary, could be close to H in realistic scenarios. However, a situation with $m \approx H$ can naturally rise if the field is non-minimally coupled to gravity. That is, a coupling

$\mathcal{L} \supset cR\sigma^2$, where R is the Ricci scalar, can uplift the particle mass during inflation $m^2 = (c/12)H^2$, regardless of a smaller 'bare' mass. Here we have used $R = (1/12)H^2$ during inflation, and we notice for $c \sim \mathcal{O}(1)$, we can have a non-negligible blue-tilted spectrum.

The way the spectrum of σ affects the curvature perturbation depends on the cosmology, and in particular, the lifetime of σ. During inflation, the energy density stored in σ is of order H^4, as expected, since σ receives H-scale quantum fluctuations. This is subdominant compared to the energy stored in the inflaton field $\sim H^2 M_{\rm pl}^2$. This implies σ acts as a spectator field during inflation, and through the stochastic effects, σ obtains isocurvature fluctuations. After the end of inflation, σ dilutes as matter while the inflaton decay products dilute as radiation. Therefore, similar to the curvaton paradigm [209–212], the fractional energy density in σ increases with time. Eventually, σ decays into Standard Model radiation, and its isocurvature perturbations get imprinted onto the curvature perturbation. Different from the curvaton paradigm, in our scenario, σ does not dominate the energy density of the Universe, and also the fluctuations of the inflaton are not negligible. In particular, on large scales, observed via CMB and LSS, the fluctuations are red-tilted and sourced by the inflaton, as in ΛCDM cosmology. On the other hand, the blue-tilted σ fluctuations are subdominant on those scales, while dominant at smaller scales $\lesssim$ Mpc. These enhanced perturbations can source an SGWB, observable in future gravitational wave detectors, as we describe below.

The rest of the chapter is organized as follows. In section 4.2, we describe the evolution of the inflaton field and σ along with some general properties of curvature perturbation in our framework. In section 4.3, we compute the stochastic contributions to σ fluctuations to obtain its power spectrum. We then use these results in section 4.4 to determine the full shape of the curvature power spectrum, both on large and small scales. The small-scale enhancement of the curvature power spectrum leads to an observable SGWB

and we evaluate the detection prospects in section 4.5 in the context of μ-Hz to Hz-scale gravitational wave detectors. We conclude in section 4.6. We include some technical details relevant to the computation of SGWB in appendix C.1.

4.2 Cosmological History and Curvature Perturbation

We now describe in detail the cosmological evolution considered in this work. We assume that the inflaton field ϕ drives the expansion of the Universe during inflation and the quantum fluctuations of ϕ generate the density fluctuations that we observe in the CMB and LSS, as in standard cosmology. We also assume that there is a second real scalar field σ which behaves as a subdominant spectator field during inflation, as alluded to above. We parametrize its potential as,

$$V(\sigma) = \frac{1}{2}m^2\sigma^2 + \frac{1}{4}\lambda\sigma^4. \tag{4.1}$$

The σ field does not drive inflation but nonetheless obtains quantum fluctuations during inflation. In particular, σ obtains stochastic fluctuations around the minimum of its potential, as we compute in section 4.3. After the end of inflation, the inflaton is assumed to reheat into radiation with energy density ρ_r, which dominates the expansion of the Universe.

We have ignored the interaction between ϕ and σ. First, the two fields could be part of completely different sectors with no mediators that can couple the two sectors. Second, the inflaton can be modeled as a (pseudo)-Goldstone boson so that there is an approximate shift symmetry $\phi \to \phi + \text{constant}$. This shift symmetry is necessary to explain the lightness of the inflaton field. Furthermore, this approximate shift symmetry leads to an almost scale-invariant power spectra on large scales, observed in the CMB.

In the presence of this symmetry, the leading operator that couples the two fields is of the type $(\partial\phi)^2\sigma^2/\Lambda^2$ with some effective theory cutoff scale Λ. Here we are assuming a $\sigma \to -\sigma$ symmetry which is present in the potential $V(\sigma)$ (4.1). The cut-off scale can be high, for example, of the order of the Planck scale. This interaction can then be safely ignored.

The evolution of the σ field depends on its mass m, interaction λ, and its frozen (root mean squared) displacement σ_0 during inflation. As long as the 'effective' mass of σ: $m^2 + 3\lambda\sigma_0^2$, is smaller than the Hubble scale, σ remains approximately frozen at σ_0. However, after the Hubble scale falls below the effective mass, σ starts oscillating around its potential. The evolution of its energy density ρ_σ, during this oscillatory phase depends on the values of m and λ. If the quartic interactions dominate, with $\lambda\sigma^2 \gg m^2$, ρ_σ dilutes like radiation [213]. Eventually, the amplitude of σ decreases sufficiently, so that $\lambda\sigma^2 \lesssim m^2$, following which ρ_σ starts redshifting like matter. We illustrate these behaviors in Fig. 4.2.

Similar to the curvaton paradigm [209–212], during the epoch ρ_σ is diluting as matter, its fractional energy density, $f_\sigma(t) \equiv \rho_\sigma(t)/\rho_r(t)$, increases linearly with the scale factor $a(t)$. For our benchmark parameter choices, we assume σ to decay into SM radiation while $f_\sigma(t_d) \sim 1$, where t_d denotes the time of σ decay. After t_d, the evolution of the Universe coincides with standard cosmology.

With this cosmology in mind, we can track the evolution of various cosmological perturbations using the gauge invariant quantity ζ, the curvature perturbation on uniform-density hypersufaces [214],

$$\zeta = -\psi - H\frac{\delta\rho}{\dot{\rho}}. \tag{4.2}$$

Here ψ is a fluctuation appearing in the spatial part of the metric as, $\delta g_{ij} = -2a^2\psi\delta_{ij}$ (ignoring vector and tensor perturbations), $\delta\rho$ denotes a fluctuation around a homogeneous

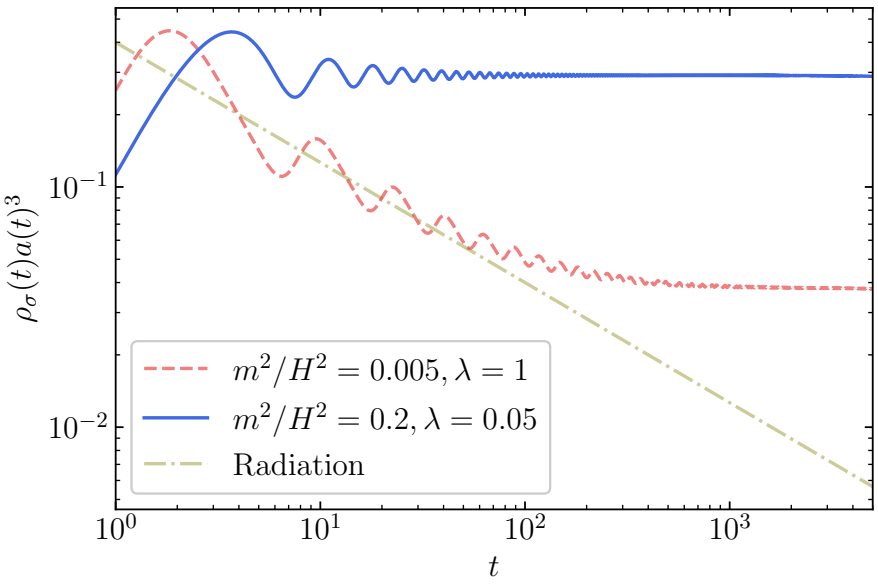

Figure 4.2: Time evolution of scalar field energy density $\rho_\sigma(t)$. In scenarios where the quartic term dominates the initial evolution (dashed red), the field dilutes as radiation (dot-dashed olive), $\rho_\sigma(t) \propto 1/a(t)^4$. Eventually, the mass term becomes important, and the behavior becomes $\rho_\sigma(t) \propto 1/a(t)^3$. The benchmark choices in this work will mimic the blue curve where the evolution of $\rho_\sigma(t)$ is always dominated by the mass term with a matter-like dilution. For both the blue and the red curves, $t = 1$ corresponds to the moment when the Hubble scale is approximately equal to the effective mass and the field starts oscillating.

density ρ, and an overdot denotes a derivative with respect to physical time t. We assume that the decay products of ϕ do not interact with σ during their cosmological evolution. Since there is no energy transfer between the two sectors, their energy densities evolve as,

$$\dot{\rho}_r = -4H\rho_r\,, \quad \dot{\rho}_\sigma = -3H\rho_\sigma, \tag{4.3}$$

where we have focused on the epoch where σ dilutes like matter. For the benchmark parameter choices discussed below, the matter-like dilution for σ onsets soon after inflation. Similar to eq. (4.2), we can parametrize gauge invariant fluctuations in radiation and σ with the variables,

$$\zeta_r = -\psi + \frac{1}{4}\frac{\delta\rho_r}{\rho_r}, \;\; \zeta_\sigma = -\psi + \frac{1}{3}\frac{\delta\rho_\sigma}{\rho_\sigma}. \tag{4.4}$$

In terms of the above variables, we can express eq. (4.2) as,

$$\zeta = \frac{4}{4+3f_\sigma}\zeta_r + \frac{3f_\sigma}{4+3f_\sigma}\zeta_\sigma = \zeta_r + \frac{f_\sigma}{4+3f_\sigma}S_\sigma. \tag{4.5}$$

Here $S_\sigma \equiv 3(\zeta_\sigma - \zeta_r)$ is the isocurvature perturbation between radiation and σ perturbations. In the absence of any energy transfer, ζ_r and ζ_σ are each conserved at super-horizon scales [215]. As a result, the evolution of ζ is entirely determined by the time-dependent relative energy density of between radiation and σ, $f_\sigma = \rho_\sigma/\rho_r$. Since ζ_r and S_σ are uncorrelated, the power spectrum for curvature perturbation $\langle\zeta(\mathbf{k})\zeta(\mathbf{k}')\rangle \equiv (2\pi)^3\delta(\mathbf{k}+\mathbf{k}')P_\zeta(k)$ is determined by,

$$P_\zeta(k) = P_{\zeta_r}(k) + \left(\frac{f_\sigma}{4+3f_\sigma}\right)^2 P_{S_\sigma}(k)\,, \tag{4.6}$$

or equivalently,

$$\Delta^2_\zeta(k) = \Delta^2_{\zeta_r}(k) + \left(\frac{f_\sigma}{4+3f_\sigma}\right)^2 \Delta^2_{S_\sigma}(k)\,, \tag{4.7}$$

where $\Delta^2_\zeta(k) = k^3P_\zeta(k)/(2\pi^2)$, with $\Delta^2_{\zeta_r}(k)$ and $\Delta^2_{S_\sigma}(k)$ defined analogously.

To compute the spectral tilt, we denote the comoving momentum of the mode that enters the horizon at t_d, the time of σ decay, as k_d which satisfies $k_d = a(t_d)H(t_d)$.

For $t > t_d$, ζ remains conserved with time on superhorizon scales. Correspondingly, for $k < k_d$, the spectral tilt is given by,

$$n_s - 1 \equiv \frac{\mathrm{d}\ln\Delta^2_\zeta(k)}{\mathrm{d}\ln k} = \frac{\Delta^2_{\zeta_r}(k)}{\Delta^2_\zeta(k)}\frac{\mathrm{d}\ln\Delta^2_{\zeta_r}(k)}{\mathrm{d}\ln k} + \left(\frac{f_\sigma}{4+3f_\sigma}\right)^2\frac{\Delta^2_{S_\sigma}(k)}{\Delta^2_\zeta(k)}\frac{\mathrm{d}\ln\Delta^2_{S_\sigma}(k)}{\mathrm{d}\ln k}. \tag{4.8}$$

We will consider scenarios where the radiation energy density ρ_r originates from the inflaton, and therefore, $\mathrm{d}\ln\Delta^2_{\zeta_r}(k)/\mathrm{d}\ln k \approx -0.04$ determines the spectral tilt observed on CMB scales [186]. On the other hand, σ acquires stochastic fluctuations to give rise to a blue-tilted power spectrum with $\mathrm{d}\ln\Delta^2_{S_\sigma}(k)/\mathrm{d}\ln k \sim 0.3$, as discussed next in section 4.3. Since we will be interested in scenarios with $f_\sigma \lesssim 1$, i.e., $(f_\sigma/(4+3f_\sigma))^2 \lesssim 0.02$, we require $\Delta^2_{S_\sigma}(k)/\Delta^2_\zeta(k) \lesssim 1$ on CMB-scales to be compatible with CMB measurements of n_s. We can also compute the running of the tilt,

$$\frac{\mathrm{d}n_s}{\mathrm{d}\ln k} \approx \left(\frac{f_\sigma}{4+3f_\sigma}\right)^2\frac{\Delta^2_{S_\sigma}(k)}{\Delta^2_\zeta(k)}\left(\frac{\mathrm{d}\ln\Delta^2_{S_\sigma}(k)}{\mathrm{d}\ln k}\right)^2. \tag{4.9}$$

Our benchmark parameter choices, discussed above, thus also satisfy the CMB constraints on $\mathrm{d}n_s/\mathrm{d}\ln k$ [186].

4.3 Review of the Stochastic Formalism

A perturbative treatment of self-interacting light scalar fields in de Sitter (dS) spacetime is subtle due to infrared divergences. A stochastic approach [207, 208] can be used to capture the nontrivial behavior of such fields in dS. In this formalism, the superhorizon components of the fields are considered classical stochastic fields that satisfy a Langevin equation, which includes a random noise originating from the sub-horizon physics. This gives rise to a Fokker-Planck equation for the probability distribution function (PDF) of the stochastic field, which can be used to calculate correlation func-

tions of physical observables. We now review these ideas briefly while referring the reader to refs. [207, 208, 216–219] for more details.

4.3.1 Langevin and Fokker-Planck Equations

The stochastic approach provides an effective description for the long-wavelength, superhorizon sector of the field theory by decomposing the fields into long-wavelength classical components and short-wavelength quantum operators. For instance, a light scalar field can be decomposed as

$$\begin{aligned}\sigma_{\rm tot.}(\mathbf{x},t) &= \sigma(\mathbf{x},t) \\ &+ \int \frac{\mathrm{d}^3k}{(2\pi)^3}\theta(k-\epsilon a(t)H)e^{-i\mathbf{k}\cdot\mathbf{x}}(a_{\mathbf{k}}u_k + a^\dagger_{-\mathbf{k}}u^*_k),\end{aligned} \tag{4.10}$$

where $\theta(\cdots)$ is the Heaviside step function, a is the scale factor, H is the Hubble scale, and $\epsilon \lesssim 1$ is a constant number (not to be confused with the slow-roll parameter) which defines the boundary between long ($k < \epsilon a(t)H$) and short ($k > \epsilon a(t)H$) modes. We have also denoted the classical part of the field as $\sigma(\mathbf{x},t)$. The quantum description of the short modes is characterized by the creation and annihilation operators $a_{\mathbf{k}}, a^\dagger_{\mathbf{k}}$ along with the mode functions $u_k(t), u^*_k(t)$.

For a light field with $|V''(\sigma)| \ll H^2$, it can be shown [207, 208, 216, 217] that the classical part of the field, $\sigma(\mathbf{x},t)$, follows a Langevin equation

$$\dot\sigma(\mathbf{x},t) = -\frac{1}{3H}V'(\sigma) + \xi(\mathbf{x},t). \tag{4.11}$$

Here an overdot and a prime denote derivative with respect to time and the field, respectively. The noise ξ arises from short-scale modes,

$$\xi(\mathbf{x},t) = \epsilon a H^2 \int \frac{\mathrm{d}^3k}{(2\pi)^3}\delta(k-\epsilon aH)e^{-i\mathbf{k}\cdot\mathbf{x}}(a_{\mathbf{k}}u_k + a^\dagger_{-\mathbf{k}}u^*_k), \tag{4.12}$$

with a correlation

$$\langle \xi(\mathbf{x}_1,t_1)\xi(\mathbf{x}_2,t_2)\rangle = \frac{H^3}{4\pi^2}\delta(t_1-t_2)j_0(\epsilon aH|\mathbf{x}_1-\mathbf{x}_2|), \tag{4.13}$$

where $j_0(x) = \sin x/x$ is the zeroth order spherical Bessel function. We see that the noise is uncorrelated in time (i.e., it is a white noise), but also it is uncorrelated over spatial separations larger than $(\epsilon a H)^{-1}$.

The Langevin equation (4.11) gives rise to a Fokker-Planck equation for the one-point PDF,

$$\frac{\partial P_{\rm FP}(t,\sigma(\mathbf{x},t))}{\partial t} = \left[\frac{V''(\sigma(\mathbf{x},t))}{3H} + \frac{V'(\sigma(\mathbf{x},t))}{3H}\frac{\partial}{\partial\sigma} + \frac{H^3}{8\pi^2}\frac{\partial^2}{\partial\sigma^2}\right] P_{\rm FP}(t,\sigma(\mathbf{x},t)). \tag{4.14}$$

Here $P_{\rm FP}(t,\sigma(\mathbf{x},t))$ is the PDF of the classical component to take the value $\sigma(\mathbf{x},t)$ at time t. Thus the Fokker-Planck equation describes how an ensemble of field configurations evolves as a function of time, according to the underlying Langevin equation. In this equation, the first and second terms on the right-hand side represent classical drift terms that depend on the potential $V(\sigma)$. The third term represents a diffusion contribution from the noise ξ. While the classical drift tries to move the central value of the field towards the minimum of the potential, the diffusion contribution pushes the field away from the minimum. An equilibrium is achieved when these two effects balance each other. This equilibrium solution can be obtained by setting $\partial P_{\rm FP}/\partial t = 0$ in (4.14), and is given by

$$P_{\rm FP,eq}(\sigma) = \frac{1}{\mathcal{N}}\exp\left(-\frac{8\pi^2}{3H^4}V(\sigma)\right), \tag{4.15}$$

where $\mathcal{N}$ is a normalization constant. Upon a variable change

$$\tilde{P}_{\rm FP}(t,\sigma) \equiv \exp\left(\frac{4\pi^2 V(\sigma)}{3H^4}\right) P_{\rm FP}(t,\sigma), \tag{4.16}$$

eq. (4.14) can written as

$$\frac{\partial \tilde{P}_{\rm FP}(t,\sigma)}{\partial t} = \frac{H^3}{4\pi^2}\underbrace{\left[-\frac{1}{2}\left(v'^2 - v''\right) + \frac{1}{2}\frac{\partial^2}{\partial\sigma^2}\right]}_{D_\sigma}\tilde{P}_{\rm FP}(t,\sigma), \tag{4.17}$$

with $v(\sigma) = 4\pi^2 V(\sigma)/(3H^4)$. We can recast the above as an eigenvalue equation. To that end, we write

$$\tilde{P}_{\rm FP}(t,\sigma) = \sum_n a_n e^{-\Lambda_n t}\psi_n(\sigma), \tag{4.18}$$

where $\psi_n(\sigma)$ satisfies the equation

$$D_\sigma \psi_n(\sigma) = -\frac{4\pi^2}{H^3}\Lambda_n \psi_n(\sigma). \tag{4.19}$$

The eigenfunctions $\psi_n(\sigma)$ form an orthonormal basis of functions and a_n's are some arbitrary coefficients.

This time-independent eigenvalue equation (4.19) can be solved numerically for a generic potential $V(\sigma)$, as we discuss below with an example. By definition, and independent of the form of the potential, the eigenfunction ψ_0 corresponding to the eigenvalue $\Lambda_0 = 0$, determines the equilibrium distribution. Solution of the eq. (4.19) for $\Lambda_0 = 0$ is given by

$$\psi_0(\sigma) = \frac{1}{\sqrt{\mathcal{N}}}\exp\left(-\frac{4\pi^2}{3H^4}V(\sigma)\right). \tag{4.20}$$

Thus comparing to eq. (4.15) we get,

$$P_{\rm FP,eq}(\sigma) = \psi_0(\sigma)^2\,. \tag{4.21}$$

4.3.2 Two-point Correlation Function and Power Spectrum

We are interested in calculating the two-point correlation functions of cosmological perturbations. Any such two-point correlation function depends only on the geodesic distance s between the two points. Given the coordinates of the two points $(\mathbf{x}_1, t_1)$ and $(\mathbf{x}_2, t_2)$, this distance can be parametrized by $z = 1 + H^2 s^2/2$ with

$$z = \cosh H(t_1 - t_2) - \frac{1}{2}e^{H(t_1+t_2)}\left(H|\mathbf{x}_1 - \mathbf{x}_2|\right)^2. \tag{4.22}$$

To understand the significance of the variable z, we first write the two-point correlation function for an arbitrary function of σ, $g(\sigma)$, as

$$G_g(\mathbf{x}_1, t_1; \mathbf{x}_2, t_2) = \langle g(\sigma(\mathbf{x}_1, t_1)) g(\sigma(\mathbf{x}_2, t_2)) \rangle. \tag{4.23}$$

To compute this, it is more convenient to calculate the temporal correlation first, and then use the fact that equal-time correlations over spatially separated points are related to the temporal correlation through the de Sitter-invariant variable z (4.22). In particular, for coincident points G_g is a function of $(t_1 - t_2)$ only, which can be expressed in terms of z for large $|z|$ as,

$$G_g(t_1 - t_2) \approx G_g(H^{-1} \ln |2z|). \tag{4.24}$$

However, for an equal time correlation function we can also write,

$$|2z| \approx (He^{Ht} |\mathbf{x}_1 - \mathbf{x}_2|)^2, \tag{4.25}$$

which gives,

$$G_g(t_1 - t_2) \simeq G_g\left(\frac{\ln |2z|}{H}\right) \simeq G_g\left(\frac{2}{H} \ln(aH|\mathbf{x}_1 - \mathbf{x}_2|)\right), \tag{4.26}$$

where the approximations hold as long as $|z| \gg 1$ and we used $a(t) = \exp(Ht)$.

Now we aim at formally calculating $G_g(t)$ in terms of solutions of the Fokker-Planck equation. The temporal correlation can be written as (see, e.g., [207, 208, 219])

$$G_g(t) = \int \mathrm{d}\sigma \int \mathrm{d}\sigma_0 P_{\mathrm{FP,eq}}(\sigma_0) g(\sigma_0) \Pi(t, \sigma; \sigma_0) g(\sigma), \tag{4.27}$$

where $\Pi(t, \sigma; \sigma_0)$ is the kernel function of the time evolution of the probability distribution function, i.e., if the probability distribution is $\delta(\sigma - \sigma_0)$ at $t = 0$ it would be $\Pi(t, \sigma; \sigma_0)$ at time t. In particular, it is defined by

$$P_{\mathrm{FP}}(t; \sigma) = \int \mathrm{d}\sigma_0 \Pi(t, \sigma; \sigma_0) P(0; \sigma_0). \tag{4.28}$$

In terms of re-scaled probabilities, we can rewrite the above as,

$$\tilde{P}_{\rm FP}(t;\sigma) = \int \mathrm{d}\sigma_0 \tilde{\Pi}(t,\sigma;\sigma_0)\tilde{P}_{\rm FP}(0;\sigma_0)\,, \tag{4.29}$$

$$\Pi(t,\sigma;\sigma_0) = e^{-v(\sigma)}\tilde{\Pi}(t,\sigma;\sigma_0)e^{v(\sigma_0)}. \tag{4.30}$$

It follows that $\tilde{\Pi}$ satisfies the same Fokker-Planck equation as $\tilde{P}_{\rm FP}$ (4.17). Therefore, the solutions can be written as

$$\tilde{\Pi}(t;\sigma,\sigma_0) = \sum_n \psi_n(\sigma_0)e^{-\Lambda_n t}\psi_n(\sigma), \tag{4.31}$$

which obeys the initial condition $\tilde{\Pi}(0;\sigma,\sigma_0) = \delta(\sigma-\sigma_0)$ is satisfied. Therefore, according to (4.27) we have[3]

$$\begin{aligned} G_g(t) = \sum_n \int \mathrm{d}\sigma_0 \psi_0(\sigma_0)g(\sigma_0)\psi_n(\sigma_0)e^{-\Lambda_n t} \\ \times \int \mathrm{d}\sigma \psi_n(\sigma)g(\sigma)\psi_0(\sigma) = \sum_n g_n^2 e^{-\Lambda_n t}, \end{aligned} \tag{4.32}$$

where

$$g_n \equiv \int \mathrm{d}\sigma \psi_n(\sigma)g(\sigma)\psi_0(\sigma). \tag{4.33}$$

We see that in late times the correlation is dominated by the smallest $\Lambda_n \neq 0$.

We can now present the equal-time correlation function by combining (4.26) and (4.32) [207, 208, 219]:

$$G_g(|\mathbf{x}_1-\mathbf{x}_2|) = \sum_n \frac{g_n^2}{(aH|\mathbf{x}_1-\mathbf{x}_2|)^{2\Lambda_n/H}}. \tag{4.34}$$

We note that this depends on the physical distance between the two points at time t, namely, $a|\mathbf{x}_1-\mathbf{x}_2|$. This correlation function has the following dimensionless power spectrum [219],

$$\begin{aligned} \Delta_g^2(k) &= \frac{k^3}{2\pi^2}P_g(k) = \frac{k^3}{2\pi^2}\int \mathrm{d}^3 r e^{-i\mathbf{k}\cdot\mathbf{r}}G_g(r) \\ &= \sum_n \frac{2g_n^2}{\pi}\Gamma\left(2-\frac{2\Lambda_n}{H}\right)\sin\left(\frac{\pi\Lambda_n}{H}\right)\left(\frac{k}{aH}\right)^{2\Lambda_n/H} \end{aligned} \tag{4.35}$$

[3]Note that $P_{\rm FP,eq}(\sigma_0) = \psi_0(\sigma_0)^2 = \psi_0(\sigma_0)\psi_0(\sigma)e^{4\pi^2 V(\sigma)/3H^4}e^{-4\pi^2 V(\sigma_0)/3H^4}$.

where Γ denotes the gamma function. This expression is valid in the limit $k \ll aH$. So far our discussion has been general and is valid for any potential under the slow-roll approximation and the assumption of a small effective mass, $|V''(\sigma)| \ll H^2$. In the next section, we discuss a concrete example with $V(\sigma)$ given in eq. (4.1).

4.4 Large Curvature Perturbation from Stochastic Fluctuations

We focus on the potential in eq. (4.1) to demonstrate how large curvature perturbation can arise from stochastic fluctuations. We first describe various equilibrium quantities and how to obtain the power spectra P_{S_σ}, and consequently evaluate P_ζ which determines the strength of the GW signal.

4.4.1 Equilibrium Configuration

The normalized PDF for the one-point function is given by eq. (4.15). For convenience, we reproduce it here

$$P_{\rm FP,eq}(\sigma) = \frac{1}{\mathcal{N}} \exp\left(-\frac{8\pi^2 V(\sigma)}{3H^4}\right), \tag{4.36}$$

with

$$\mathcal{N} = \frac{2\sqrt{2}\sqrt{\lambda}}{\exp\left(\frac{m^4\pi^2}{3H^4\lambda}\right) m K_{\frac{1}{4}}\left(\frac{m^4\pi^2}{3H^4\lambda}\right)}. \tag{4.37}$$

Here $K_n(x)$ is the modified Bessel function of the second kind. The mean displacement of the field can be computed as,

$$\langle\sigma^2\rangle = \int_0^\infty {\rm d}\sigma \sigma^2 P_{\rm FP,eq}(\sigma) = \frac{m^2}{2\lambda}\left(-1 + \frac{K_{\frac{3}{4}}\left(\frac{m^4\pi^2}{3H^4\lambda}\right)}{K_{\frac{1}{4}}\left(\frac{m^4\pi^2}{3H^4\lambda}\right)}\right). \tag{4.38}$$

In the appropriate limits, this can be simplified to,

$$\langle\sigma^2\rangle\Big|_{\lambda\to 0} = \frac{3H^4}{8\pi^2 m^2}, \tag{4.39}$$

$$\langle\sigma^2\rangle\Big|_{m\to 0} = \sqrt{\frac{3}{2\lambda}}\frac{\Gamma(3/4)}{\Gamma(1/4)\pi}H^2, \tag{4.40}$$

matching the standard results [208]. We can also compute the average energy density of the field as,

$$\begin{aligned}\langle V(\sigma)\rangle &= \int_0^\infty \mathrm{d}\sigma V(\sigma) P_{\mathrm{FP,eq}}(\sigma) \\ &= \frac{1}{32}\left(\frac{3H^4}{\pi^2} - \frac{4m^4}{\lambda} + \frac{4m^4}{\lambda}\frac{K_{\frac{3}{4}}\left(\frac{m^4\pi^2}{3H^4\lambda}\right)}{K_{\frac{1}{4}}\left(\frac{m^4\pi^2}{3H^4\lambda}\right)}\right),\end{aligned} \tag{4.41}$$

reducing to,

$$\langle V(\sigma)\rangle\Big|_{\lambda\to 0} = \frac{3H^4}{16\pi^2}, \tag{4.42}$$

$$\langle V(\sigma)\rangle\Big|_{m\to 0} = \frac{3H^4}{32\pi^2}. \tag{4.43}$$

To ensure that σ does not dominate energy density during inflation, we require

$$\langle V(\sigma)\rangle \ll 3H^2 M_{\mathrm{pl}}^2. \tag{4.44}$$

Finally, we compute $\langle V''(\sigma)\rangle$ to check the validity of slow-roll of the σ field,

$$\begin{aligned}\langle V''(\sigma)\rangle &= \int_0^\infty \mathrm{d}\sigma V''(\sigma) P_{\mathrm{FP,eq}}(\sigma) \\ &= \frac{1}{2}m^2\left(-1 + \frac{3K_{\frac{3}{4}}\left(\frac{m^4\pi^2}{3H^4\lambda}\right)}{K_{\frac{1}{4}}\left(\frac{m^4\pi^2}{3H^4\lambda}\right)}\right),\end{aligned} \tag{4.45}$$

which reduces to,

$$\langle V''(\sigma)\rangle\Big|_{\lambda\to 0} = m^2, \tag{4.46}$$

$$\langle V''(\sigma)\rangle\Big|_{m\to 0} = \frac{3\sqrt{3}\Gamma(3/4)}{\sqrt{2}\pi\Gamma(1/4)}\sqrt{\lambda}H^2 \approx 0.4\sqrt{\lambda}H^2. \tag{4.47}$$

To ensure slow-roll, we require

$$\langle V''(\sigma)\rangle \ll H^2. \tag{4.48}$$

4.4.2 Power Spectrum

To obtain isocurvature power spectrum, P_{S_σ}, we need to compute the two-point function of $\delta\rho_\sigma/\rho_\sigma$. We can write this more explicitly as,

$$\frac{\delta\rho_\sigma(\mathbf{x})}{\rho_\sigma} = \frac{\rho_\sigma(\mathbf{x}) - \langle\rho_\sigma(\mathbf{x})\rangle}{\langle\rho_\sigma(\mathbf{x})\rangle} = \frac{\rho_\sigma(\mathbf{x})}{\langle\rho_\sigma(\mathbf{x})\rangle} - 1. \tag{4.49}$$

where we can approximate $\rho_\sigma \approx V(\sigma)$, since $\langle V(\sigma)\rangle$ is approximately frozen, as long as eq. (4.48) is satisfied. Referring to eq. (4.33) and eq. (4.35), the relevant coefficient g_n for ρ_σ is determined by,

$$g_n = \frac{\int \mathrm{d}\sigma\psi_n(\sigma)\rho_\sigma\psi_0(\sigma)}{\int \mathrm{d}\sigma\psi_0(\sigma)\rho_\sigma\psi_0(\sigma)}. \tag{4.50}$$

For $n > 0$, the last term in eq. (4.49) does not contribute because of the orthogonality of the eigenfunctions.

The eigenfunctions ψ_n and the eigenvalues Λ_n relevant for eq. (4.35) can be obtained by solving the eigensystem for the potential eq. (4.1). In terms of variables, $z = \lambda^{1/4}\sigma/H$ and $\alpha = m^2/(\sqrt{\lambda}H^2)$, the eigenvalue eq. (4.19) can be written as [219],

$$\begin{aligned}\frac{\partial^2\psi_n}{\partial z^2} + \left(-\left(\frac{4\pi^2}{3}\right)^2(\alpha z + z^3)^2 + \frac{4\pi^2}{3}(\alpha + 3z^2)\right)\psi_n \\ = -\frac{8\pi^2}{\sqrt{\lambda}}\frac{\Lambda_n}{H}\psi_n.\end{aligned} \tag{4.51}$$

Given the potential in eq. (4.1), the eigenfunctions are odd (even) functions of σ for odd (even) values of n. Since ρ_σ is an even function of σ, eq. (4.50) implies $g_1 = 0$, and therefore, the leading coefficient is g_2 with the eigenvalue Λ_2 determining the first non-zero contribution to the spectral tilt. We show the numerical results for the eigenvalues for some benchmark parameter choices in Table 4.1.

The curvature power spectrum Δ^2_ζ depends on both $\Delta^2_{S_\sigma}$ and f_σ, as in eq. (4.7). With the values of g_n, Λ_n in Table 4.1, we can compute the dimensionless power spectrum $\Delta^2_{S_\sigma}$ using eq. (4.35), where we can evaluate the factor of aH at the end of inflation.

m^2/H^2	λ	Λ_2/H	g_2^2	Λ_4/H	g_4^2
0.2	0.05	0.16	1.99	0.37	0.03
0.2	0.07	0.17	1.98	0.40	0.05
0.2	0.1	0.18	1.98	0.44	0.07
0.25	0.05	0.19	1.99	0.42	0.02
0.25	0.07	0.20	1.99	0.45	0.03
0.25	0.1	0.21	1.98	0.49	0.05
0.3	0.05	0.22	1.99	0.48	0.01
0.3	0.07	0.23	1.99	0.51	0.02
0.3	0.1	0.24	1.99	0.54	0.03

Table 4.1: Eigenvalues for some benchmark parameter choices corresponding to the potential in eq. (4.1).

Furthermore, for our benchmark parameter choices, only the eigenvalue Λ_2 is relevant. Therefore, eq. (4.35) can be simplified as,

$$\Delta^2_{S_\sigma}(k) \approx \frac{2g_2^2}{\pi}\Gamma\left(2-\frac{2\Lambda_2}{H}\right)\sin\left(\frac{\pi\Lambda_2}{H}\right)\left(\frac{k}{k_{\rm end}}\right)^{2\Lambda_2/H}, \tag{4.52}$$

where $k_{\rm end} = a_{\rm end}H_{\rm end}$.

The precise value of $k_{\rm end}$ depends on the cosmological history after the CMB-observable modes exit the horizon. It is usually parametrized as the number of e-foldings $N(k) \equiv \ln(a_{\rm end}/a_k)$, where a_k is the scale factor when a k-mode exits the horizon during inflation, defined by $k = a_k H_k$. Assuming an equation of state parameter w between the end of inflation and the end of the reheating phase, we can derive the relation [220, 221],

$$\frac{k}{a_0H_0} = \left(\frac{\sqrt{\pi}}{90^{1/4}}\frac{T_0}{H_0}\right)e^{-N(k)}\left(\frac{V_k^{1/2}}{\rho_{\rm end}^{1/4}M_{\rm pl}}\right)\left(\frac{\rho_{\rm RH}}{\rho_{\rm end}}\right)^{\frac{1-3w}{12(1+w)}} \times \frac{g_{*,s,0}^{1/3}g_{*,\rm RH}^{1/4}}{g_{*,s,\rm RH}^{1/3}}. \tag{4.53}$$

Here $g_{*,\mathrm{RH}}$ and $g_{*,s,\mathrm{RH}}$ are the effective number of degrees of freedom in the energy density and entropy density, respectively, at the end of the reheating phase; V_k is the inflationary energy density when the k-mode exits the horizon; ρ_{end} and ρ_{RH} are the energy densities at the end of inflation and reheating, respectively. Plugging in the CMB temperature T_0 and the present-day Hubble parameter H_0, we arrive at

$$\begin{aligned} N(k) \approx 67 - \ln\left(\frac{k}{a_0 H_0}\right) + \ln\left(\frac{V_k^{1/2}}{\rho_{\mathrm{end}}^{1/4} M_{\mathrm{pl}}}\right) \\ + \frac{1-3w}{12(1+w)} \ln\left(\frac{\rho_{\mathrm{RH}}}{\rho_{\mathrm{end}}}\right) + \ln\left(\frac{g_{*,\mathrm{RH}}^{1/4}}{g_{*,s,\mathrm{RH}}^{1/3}}\right). \end{aligned} \tag{4.54}$$

Significant sources of uncertainty in $N(k)$ comes from V_k, ρ_{end}, ρ_{RH}, and w. Furthermore, eq. (4.54) assumes a standard cosmological history where following reheating, the Universe becomes radiation dominated until the epoch of matter-radiation equality. We now consider some benchmark choices with which we can evaluate $N(k)$. We set $k = a_0 H_0$, assume $V_k^{1/4} = 10^{16}$ GeV, close to the current upper bound [186], $\rho_{\mathrm{end}} \simeq V_k/100$, motivated by simple slow-roll inflation models, and $w \approx 0$ [222–224].[4] Then depending on the reheating temperature, we get

$$N(k) = \begin{cases} 62, & T_{\mathrm{RH}} = 6 \times 10^{15}\ \mathrm{GeV}, \\ 59, & T_{\mathrm{RH}} = 10^{11}\ \mathrm{GeV}. \end{cases} \tag{4.55}$$

For the first benchmark, we have assumed an instantaneous reheating after inflation, while for the second benchmark, the reheating process takes place for an extended period of time. For these two benchmarks, $k_{\mathrm{end}} \approx 4 \times 10^{23}$ Mpc^{-1} and 10^{22} Mpc^{-1}, respectively.

To determine $\Delta_\zeta^2(k)$, we also need to evaluate f_σ as a function of time. We can

[4]The precise value of w is model dependent, see, e.g., [225–230] and [231] for a review. However, this does not affect the superhorizon behavior of ζ_r and S_σ that we described above. Instead, w primarily affects the number of e-foldings $N(k)$ in (4.54). For example, using $w = 0.2(0.1)$ makes a 0.5%(0.2%) change in $N(k)$ for $T_{\mathrm{RH}} = 6 \times 10^{15}$ GeV in (4.55). For $T_{\mathrm{RH}} = 10^{11}$ GeV, using $w = 0.2(0.1)$ makes a 3%(2%) change in $N(k)$. Given these changes are less than 5%, we will use $w \approx 0$ in the rest of the analysis.

express the time dependence of f_σ in terms of k in the following way. A given k-mode re-enters the horizon when $k = a_k H_k$, and assuming radiation domination, we get $k/k_{\rm end} = a_{\rm end}/a_k$. Since f_σ increases with the scale factor before σ decay, we can express $f_\sigma(t) = f_\sigma(t_d)(k_d/k)$, for $t < t_d$, where k_d and k are the modes that re-enter the horizon at time t_d and t, respectively. Therefore, the final expression for the curvature power spectrum at the time of mode re-entry follows from eq. (4.7),

$$\Delta^2_\zeta(k) = \begin{cases} \Delta^2_{\zeta_r}(k) + \left(\frac{f_\sigma(t_d)}{4+3f_\sigma(t_d)}\right)^2 \Delta^2_{S_\sigma}(k), \;\; k < k_d, \\ \Delta^2_{\zeta_r}(k) + \left(\frac{f_\sigma(t_d)(k_d/k)}{4+3f_\sigma(t_d)(k_d/k)}\right)^2 \Delta^2_{S_\sigma}(k), \;\; k > k_d. \end{cases} \tag{4.56}$$

To determine the scale k_d, we consider the benchmarks discussed above, along with some additional choices for other parameters.

Benchmark 1. We focus on the first benchmark in eq. (4.55). For $m^2 = 0.2H^2$ and $\lambda \simeq 0.05 - 0.1$, we get $\langle V(\sigma)\rangle \approx 0.02H^4$ from eq. (4.41), implying $\langle V(\sigma)\rangle/V_k \approx 3\times 10^{-12}$ for $H = 5\times 10^{13}$ GeV. Assuming instantaneous reheating, and $\rho_{\rm end} \simeq V_k/100$, we see $f_\sigma \simeq 1$ for $a \simeq (1/3)\times 10^{10} a_{\rm end}$. As benchmarks, we assume σ decays when $f_\sigma = 1$ and $1/3$. Using $k_{\rm end} \approx 4\times 10^{23}$ Mpc^{-1}, we can then evaluate $k_d \approx 10^{14}$ Mpc^{-1} and $k_d \approx 3\times 10^{14}$ Mpc^{-1}, respectively. The result for the curvature power spectrum with these choices is shown in Fig. 4.3 (left).

Benchmark 2. We now discuss the second benchmark in eq. (4.55). We again choose $m^2 = 0.2H^2$ and $\lambda \simeq 0.05 - 0.1$, for which we get $\langle V(\sigma)\rangle \approx 0.02H^4$ from eq. (4.41). This implies $\langle V(\sigma)\rangle/V_k \approx 3\times 10^{-12}$ for $H = 5\times 10^{13}$ GeV, as before. The rest of the parameters can be derived in an analogous way, with one difference. During the reheating epoch, with our assumption $w \approx 0$, f_σ does not grow with the scale factor since the dominant energy density of the Universe is also diluting as matter. Accounting

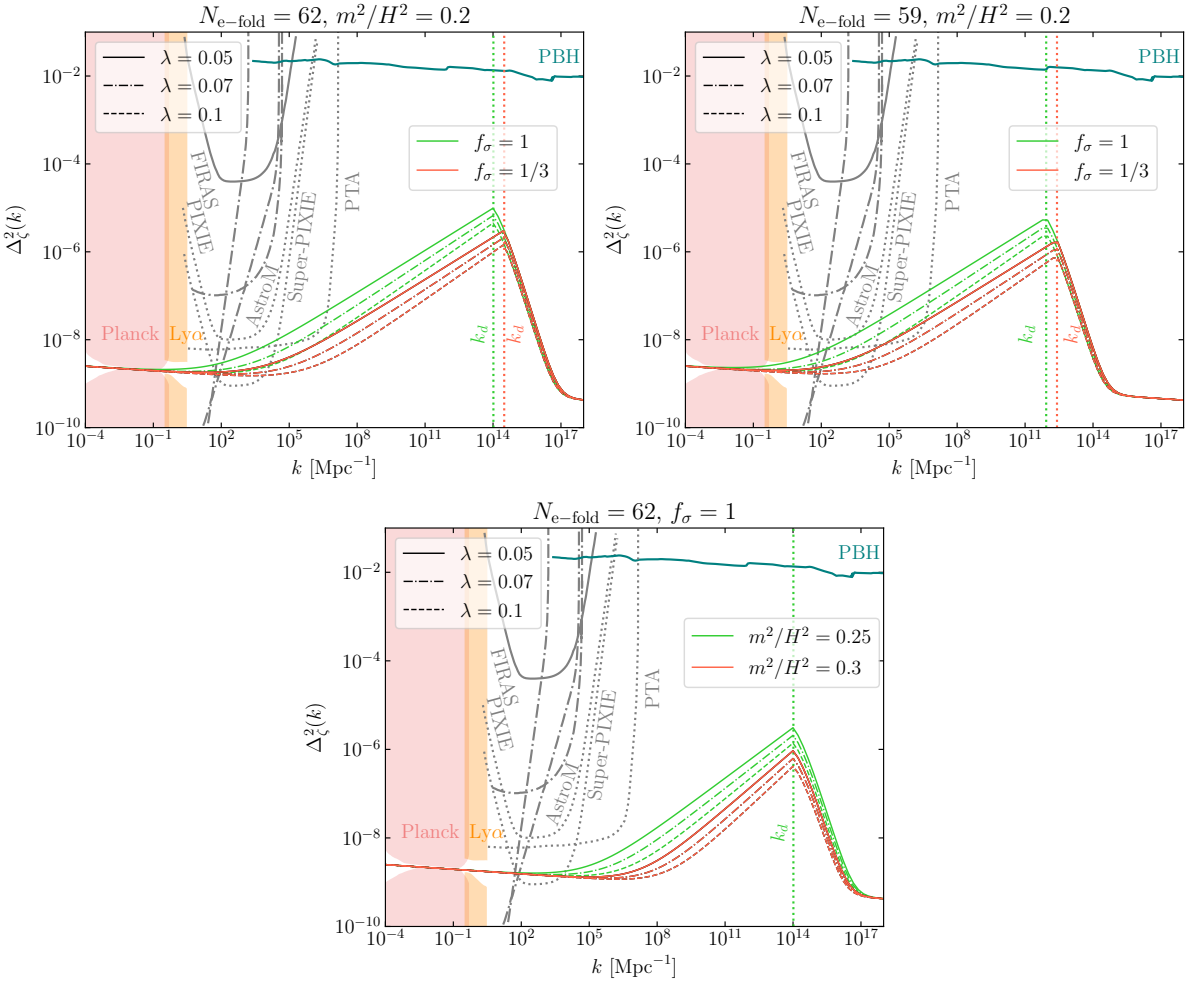

Figure 4.3: Power spectrum of curvature perturbations for the benchmarks discussed above. Stochastic effects lead to a blue-tilted spectrum of σ, with larger m and λ corresponding to larger tilts, leading to faster decay as k gets smaller. The blue-tilt is eventually cut off at k_d, the k-mode that reenters the horizon at the time of σ decay. For k larger than k_d, the fractional energy density in σ at the time of mode-reentry is smaller. Correspondingly, Δ^2_ζ gets suppressed. Eventually, for very large k, the effects of σ become negligible, and Δ^2_ζ reverts back to its standard, slightly red-tilted behavior. A smaller value of $f_\sigma(k_d)$, the fractional energy density at the time σ decay, suppresses the effect of σ to Δ^2_ζ, and hence leads to a suppressed peak. This mechanism predicts signatures in CMB spectral distortion measurements [232], especially in Super-PIXIE [233], along with Pulsar Timing Array (PTA) probes for enhanced DM substructure [234], and precision astrometry probes (AstroM) [235]. We also show constraints from FIRAS [236] and non-observation of primordial black holes (PBH) [189].

for this gives $k_d \approx 8 \times 10^{11}$ Mpc^{-1} and $k_d \approx 3 \times 10^{12}$ Mpc^{-1}, for $f_\sigma = 1$ and $1/3$, respectively, with the resulting curvature power spectrum shown in Fig. 4.3 (center).

Benchmark 3. This is same as the first benchmark discussed above, except we focus on $m^2 = 0.25H^2$ and $0.3H^2$ along with $f_\sigma = 1$. The result is shown in Fig. 4.3 (right).

We note that for all three cases, the power spectrum does not become as large as to give rise to PBH. It can also be checked that the correction to the large-scale power spectrum, relevant for the CMB, from the enhanced small-scale power spectrum, is small. In fact, repeating the argument of [237], we find

$$\delta P_\zeta(k_L) \sim \frac{\Delta_\zeta^4}{k_s^3} \sim \frac{\Delta_\zeta^4}{\Delta^2_{\zeta,\mathrm{CMB}}} \frac{k_L^3}{k_s^3} P_\zeta(k_L). \tag{4.57}$$

For $\Delta_\zeta^2 \sim 10^{-5}$ and $k_s \sim 10^{14}$ Mpc^{-1}, as in Fig. 4.3 (left), we have

$$\delta P_\zeta(k_L) \sim 10^{-46} P_\zeta(k_L), \tag{4.58}$$

for $k_L \sim 10^{-1}$ Mpc^{-1} and $\Delta^2_{\zeta,\mathrm{CMB}} \sim 10^{-9}$ (corresponding to a typical scale probed by the CMB).

4.5 Gravitational Wave Signature

4.5.1 Secondary Gravitational Waves from Scalar Curvature Perturbation

We now review how large primordial curvature perturbations can source GW at the second order in perturbation theory [238–241] (for a review see [188]). We then evaluate the GW spectrum sourced by Δ_ζ^2 computed in section 4.4. We start our discussion with a brief review of the essential relations, following [242], and expand the discussion further

in appendix C.1. For some recent work on scalar-induced gravitational waves, see, e.g., Refs. [243, 244].

We can write a tensor perturbation in Fourier space as,

$$h_{ij}(\tau,\mathbf{x}) = \sum_{\lambda=+,\times}\int \frac{\mathrm{d}^3k}{(2\pi)^3} e^{i\mathbf{k}\cdot\mathbf{x}}\epsilon_{ij}^{\lambda}(\mathbf{k})h_\lambda(\tau,\mathbf{k})\,, \tag{4.59}$$

where $\epsilon_{ij}^{\lambda=\{+,\times\}}(\mathbf{k})$ are polarization tensors:

$$\epsilon_{ij}^{+}(\mathbf{k}) = \frac{1}{\sqrt{2}}\left(\mathrm{e}_{1,i}(\mathbf{k})\mathrm{e}_{1,j}(\mathbf{k}) - \mathrm{e}_{2,i}(\mathbf{k})\mathrm{e}_{2,j}(\mathbf{k})\right), \tag{4.60}$$

$$\epsilon_{ij}^{\times}(\mathbf{k}) = \frac{1}{\sqrt{2}}\left(\mathrm{e}_{1,i}(\mathbf{k})\mathrm{e}_{2,j}(\mathbf{k}) + \mathrm{e}_{2,i}(\mathbf{k})\mathrm{e}_{1,j}(\mathbf{k})\right), \tag{4.61}$$

with $\mathrm{e}_{1,2}$ the orthonormal bases spanning the plane transverse to $\mathbf{k}$. The equation of motion determining the generation and evolution of GW is given by

$$h_\lambda''(\tau,\mathbf{k}) + 2\mathcal{H}h_\lambda'(\tau,\mathbf{k}) + k^2 h_\lambda(\tau,\mathbf{k}) = 4\mathcal{S}_\lambda(\tau,\mathbf{k}), \tag{4.62}$$

where $'$ denotes derivative with respect to the conformal time τ and $\mathcal{H} = a'/a$ is the conformal Hubble parameter. The second-order (in scalar metric perturbation Φ) source term is given by[5]

$$\begin{aligned}\mathcal{S}_\lambda(\tau,\mathbf{k}) = \int \frac{\mathrm{d}^3q}{(2\pi)^3}\frac{Q_\lambda(\mathbf{k},\mathbf{q})}{3(1+w)}\Big[&2(5+3w)\Phi_{\mathbf{p}}\,\Phi_{\mathbf{q}} \\ &+ \tau^2(1+3w)^2\Phi_{\mathbf{p}}'\,\Phi_{\mathbf{q}}' + 2\tau(1+3w)(\Phi_{\mathbf{p}}\,\Phi_{\mathbf{q}}' + \Phi_{\mathbf{p}}\,\Phi_{\mathbf{q}}')\Big].\end{aligned} \tag{4.64}$$

We have defined $\mathbf{p} \equiv \mathbf{k} - \mathbf{q}$, $\Phi_{\mathbf{k}} \equiv \Phi(\tau,\mathbf{k})$, and a projection operator $Q_\lambda(\mathbf{k},\mathbf{q})$:

$$Q_\lambda(\mathbf{k},\mathbf{q}) \equiv \epsilon_\lambda^{ij}(\mathbf{k})q_i q_j\,. \tag{4.65}$$

[5] We parametrize the scalar metric fluctuations, for vanishing anisotropic stress, as

$$\mathrm{d}s^2 = -\left(1+2\Phi\right)\mathrm{d}t^2 + a^2\left(1-2\Phi\right)\delta_{ij}\mathrm{d}x^i\mathrm{d}x^j \tag{4.63}$$

The metric perturbation $\Phi(\tau, \mathbf{k})$ can be written in terms of the primordial curvature perturbation $\zeta(\mathbf{k})$,

$$\Phi(\tau, \mathbf{k}) = \frac{3+3w}{5+3w} T_\Phi(k\tau)\zeta(\mathbf{k})\,, \tag{4.66}$$

via a transfer function $T_\Phi(k\tau)$ which depends on w. With the above quantities, one can now solve eq. (4.62) using the Green function method,[6]

$$h_\lambda(\tau, \mathbf{k}) = \frac{4}{a(\tau)} \int_{\tau_0}^{\tau} \mathrm{d}\bar{\tau} G_{\mathbf{k}}(\tau, \bar{\tau}) a(\bar{\tau}) \mathcal{S}_\lambda(\bar{\tau}, \mathbf{k})\,. \tag{4.67}$$

Using the solutions of eq. (4.62), the power spectrum $P_\lambda(\tau, k)$, defined via,

$$\langle h_{\lambda_1}(\tau, \mathbf{k}_1) h_{\lambda_2}(\tau, \mathbf{k}_2)\rangle \equiv (2\pi)^3 \delta_{\lambda_1\lambda_2} \delta^3(\mathbf{k}_1 + \mathbf{k}_2) P_{\lambda_1}(\tau, k_1)\,, \tag{4.68}$$

can be written as,

$$\begin{aligned} &\langle h_{\lambda_1}(\tau, \mathbf{k}_1) h_{\lambda_2}(\tau, \mathbf{k}_2)\rangle = \\ &16 \int \frac{\mathrm{d}^3 q_1}{(2\pi)^3} \frac{\mathrm{d}^3 q_2}{(2\pi)^3} Q_{\lambda_1}(\mathbf{k}_1, \mathbf{q}_1) Q_{\lambda_2}(\mathbf{k}_2, \mathbf{q}_2) I(|\mathbf{k}_1 - \mathbf{q}_1|, q_1, \tau_1) \\ &\times\ I(|\mathbf{k}_2 - \mathbf{q}_2|, q_2, \tau_2) \langle \zeta(\mathbf{q}_1)\zeta(\mathbf{k}_1 - \mathbf{q}_1)\zeta(\mathbf{q}_2)\zeta(\mathbf{k}_2 - \mathbf{q}_2)\rangle\,. \end{aligned} \tag{4.69}$$

Here

$$I(p, q, \tau) = \frac{1}{a(\tau)} \int_{\tau_0}^{\tau} \mathrm{d}\bar{\tau}\ G_{\mathbf{k}}(\tau, \bar{\tau}) a(\bar{\tau}) f(p, q, \bar{\tau})\,, \tag{4.70}$$

and

$$\begin{aligned} \frac{(5+3w)^2}{3(1+w)} f(p, q, \tau) &= 2(5+3w) T_\Phi(p\tau)\, T_\Phi(q\tau) \\ &+ \tau^2 (1+3w)^2 T'_\Phi(p\tau)\, T'_\Phi(q\tau) \\ &+ 2\tau(1+3w) \left[T_\Phi(p\tau)\, T'_\Phi(q\tau) + T'_\Phi(p\tau)\, T_\Phi(q\tau)\right]. \end{aligned} \tag{4.71}$$

where $T'_\Phi(p\tau) = \partial T_\Phi(p\tau)/\partial\tau$. We note that the power spectrum is sourced by the four-point correlation function of super-horizon curvature perturbations, and is further modified by the sub-horizon evolution as encapsulated in $I(p, q, \tau)$.

[6]Scale factors appearing in the I integral as $a(\bar{\tau})/a(\tau)$ are the artifact of $G_{\mathbf{k}}(\tau, \bar{\tau})$ being Green's function of the new variable $v(\tau, \mathbf{k}) = ah(\tau, \mathbf{k})$ and not h_λ itself; see Appendix C.1.2.

The four-point function in eq. (4.69) has both disconnected and connected contributions, from the scalar power spectrum and trispectrum, respectively. The connected contribution usually contributes in a subdominant way compared to the disconnected piece in determining total GW energy density; see [245] for a general argument.[7] Therefore, in the following, we focus only on the disconnected contribution, which can be written as

$$P_\lambda(\tau,k)\Big|_{\rm d} = 32\int\frac{{\rm d}^3q}{(2\pi)^3}Q_\lambda(\mathbf{k},\mathbf{q})^2 I(|\mathbf{k}-\mathbf{q}|,q,\tau)^2 \times P_\zeta(q)P_\zeta(|\mathbf{k}-\mathbf{q}|)\,. \tag{4.72}$$

For a derivation of this formula see appendix C.1.3.

GW signal strength can be characterized by SGWB energy density per unit logarithmic interval of frequency and normalized to the total energy density [248],

$$h^2\Omega_{\rm GW} = \frac{1}{\rho_{\rm tot}}\frac{{\rm d}\rho_{\rm GW}}{{\rm d}\log f} \tag{4.73}$$

where the present day Hubble parameter is given by $H_0 = 100h$ km/s/Mpc and $\rho_{\rm tot} = 3M_{\rm pl}^2H_0^2$ is the critical energy density in terms of the reduced Planck mass $M_{\rm pl} \approx 2.4\times 10^{18}$ GeV. The total energy density $\rho_{\rm GW}$ is given by,

$$\rho_{\rm GW} = \frac{M_{\rm pl}^2}{4}\int{\rm d}\ln k\frac{k^3}{16\pi^2}\times \sum_\lambda\left(\langle\dot h_\lambda(t,\mathbf{k})\dot h_\lambda(t,-\mathbf{k})\rangle' + \frac{k^2}{a^2}\langle h_\lambda(t,\mathbf{k})h_\lambda(t,-\mathbf{k})\rangle'\right), \tag{4.74}$$

with the primes denoting the fact that momentum-conserving delta functions are factored out, $\langle h_\lambda(t,\mathbf{k})h_\lambda(t,\mathbf{k}')\rangle = (2\pi)^3\delta^3(\mathbf{k}+\mathbf{k}')\langle h_\lambda(t,\mathbf{k})h_\lambda(t,-\mathbf{k})\rangle'$. Approximating $\dot h_\lambda(t,\mathbf{k}) \approx (k/a)h_\lambda(t,\mathbf{k})$, we can simplify to get,[8]

$$\Omega_{\rm GW} = \frac{1}{48}\left(\frac{k}{a(\tau)H(\tau)}\right)^2\sum_{\lambda=+,\times}\Delta_\lambda^2(\tau,k), \tag{4.75}$$

[7] See also [242, 246, 247] for examples where the connected contribution can be important.

[8] Note that we are using the convention at which the spatial part of the metric is given by $a^2(\delta_{ij}+h_{ij}/2){\rm d}x^i{\rm d}x^j$. If we were using an alternative convention $a^2(\delta_{ij}+h_{ij}){\rm d}x^i{\rm d}x^j$, then the factor of 1/48 would be replaced by 1/12 as in refs. [240, 248].

where $\Delta^2_\lambda(\tau,k) = (k^3/(2\pi^2))P_\lambda(\tau,k)$.

The above expression can be rewritten in form convenient for numerical evaluation (see appendix C.1.4 for a derivation),[9]

$$\Omega_{\rm GW}(k) = \frac{2}{48\alpha^2}\int_0^\infty {\rm d}t \int_{-1}^1 {\rm d}s\ \mathcal{K}_{\rm d}(u,v)\Delta^2_\zeta(uk)\Delta^2_\zeta(vk) \tag{4.76}$$

where $u = |\mathbf{k}-\mathbf{q}|/k = p/k, v = q/k, s = u-v, t = u+v-1$, and $\mathcal{K}_{\rm d}$ is the kernel function following from manipulating the integrand of eq. (4.72). This kernel function is illustrated in fig. 4.4a.

We now focus on the scenario where GW is generated during a radiation dominated epoch and set $w = 1/3$. We can then write (see Appendix C.1.1 for details),

$$T_\Phi(k\tau) = \frac{9\sqrt{3}}{(k\tau)^3}\left(\sin\frac{k\tau}{\sqrt{3}} - \frac{k\tau}{\sqrt{3}}\cos\frac{k\tau}{\sqrt{3}}\right), \tag{4.77}$$

and plot this function in fig. 4.4b. We note that after entering the horizon, modes start to oscillate and decay, and as a result, the sub-horizon modes do not significantly contribute to GW generation. In fig. 4.4c, we confirm that at any given time $f(p,q,\tau)$ is suppressed for shorter modes that have re-entered the horizon earlier. Finally, the green function is given by (see appendix C.1.2 for details)

$$G_{\mathbf{k}}(\tau,\bar\tau) = \frac{\sin[k(\tau-\bar\tau)]}{k}. \tag{4.78}$$

With these expressions, we can obtain a physical understanding of GW generation via eq. (4.72). The Green function, given in eq. (4.78), is an oscillatory function of time whose frequency is k. The quantity $f(p,q,\tau)$ is also an oscillatory and decaying function of time (see fig. 4.4c), inheriting these properties from the transfer function (4.77). Therefore, the dominant contribution to the integral (4.70) is a resonant contribution when the momentum of the produced GW is of the same order as the momentum of the scalar

[9]Note that the integration variable u and v are swapped with t and s since in the $t-s$ space, integration limits are independent of the integration variables.

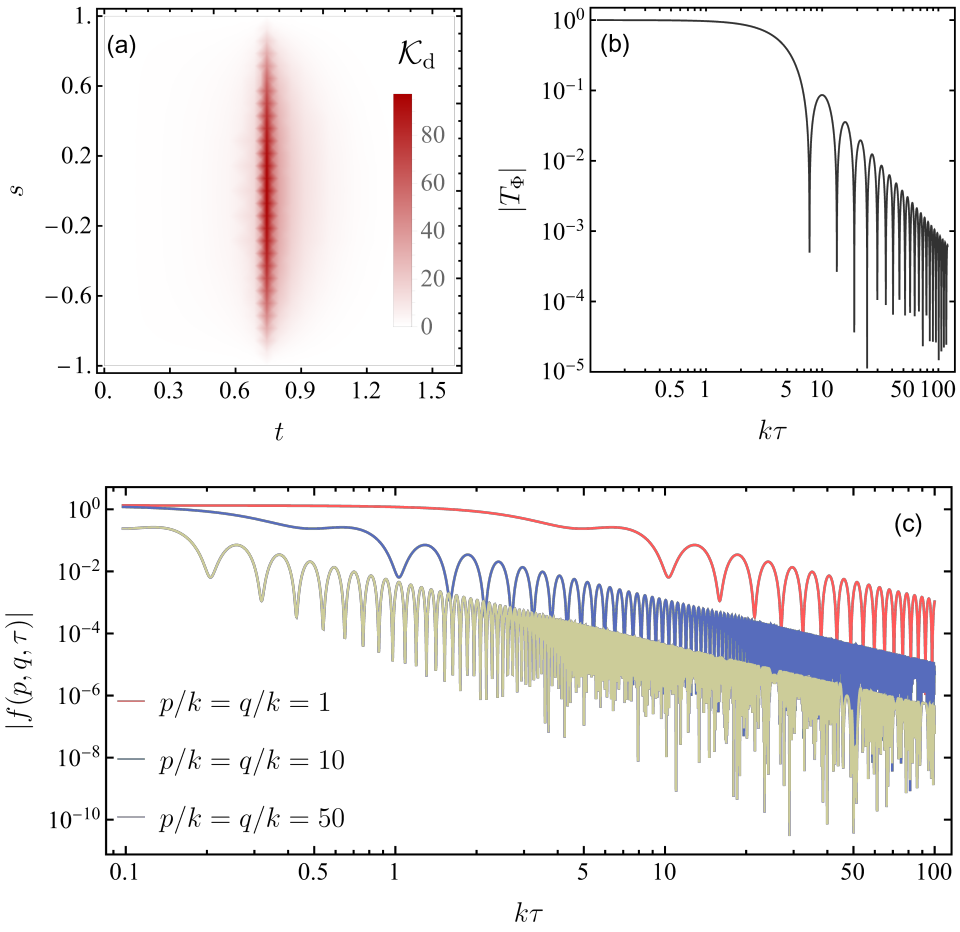

Figure 4.4: (a) The kernel function from eq. (4.76). We note a clear resonance contribution from $t \simeq 0.7$ corresponding to $u + v \simeq \sqrt{3}$. (b) The transfer function T_Φ. (c) Function $f(p, q, \tau)$ as in eq. (4.71). We see that for the scalar modes that enter the horizon earlier, with $p, q > k$, this function is more suppressed as expected from the behavior of the transfer function.

modes, i.e., $k \sim p \sim q$. In particular, the resonant point is at $u + v \simeq \sqrt{3}$ [245] as shown in fig. 4.4a. GW generation is suppressed in other parts of the phase space. For example, the source term, which contains gradients of the curvature perturbation [239], is suppressed by small derivatives if any of the wavenumbers p, q of ζ is much smaller than k. On the other hand, if p, q are much larger than k, then the scalar modes would have decayed significantly after entering the horizon by the time $k \sim H$, and thus the production of GW with momentum k gets suppressed.

To obtain the final result for $\Omega_{\rm GW}$, we note that the GW comoving wavenumber k is related to the present-day, redshifted frequency f of the generated GW via

$$f = f_* \left(\frac{a_*}{a_0}\right) = \frac{k}{2\pi} \simeq 1.5\,\text{mHz} \left(\frac{k}{10^{12}\,\text{Mpc}^{-1}}\right), \tag{4.79}$$

where f_* and a_* are respectively the frequency and the scale factor at the time of GW generation. Using these expressions, we arrive at our final result, shown in Fig. 4.5, for the same benchmark choices discussed in Fig. 4.3. We see that stochastic effects can naturally give rise to a large enough SGWB, within the sensitivity range of DECIGO, BBO, μ-Ares, and Ultimate DECIGO [249–251].

4.6 Conclusion

In this work, we have discussed an early Universe scenario containing a light spectator field, along with an inflaton field. The fluctuations of the inflaton are red-tilted and explain the observed fluctuations in the CMB and LSS. On the other hand, the spectator field σ naturally acquires a blue-tilted power spectrum. This blue-tilted power spectrum is eventually cut-off at very small scales since when such small-scale modes enter the horizon, the spectator field contributes subdominantly to the total energy density. As a consequence, primordial black holes are not produced in this scenario. Overall, this

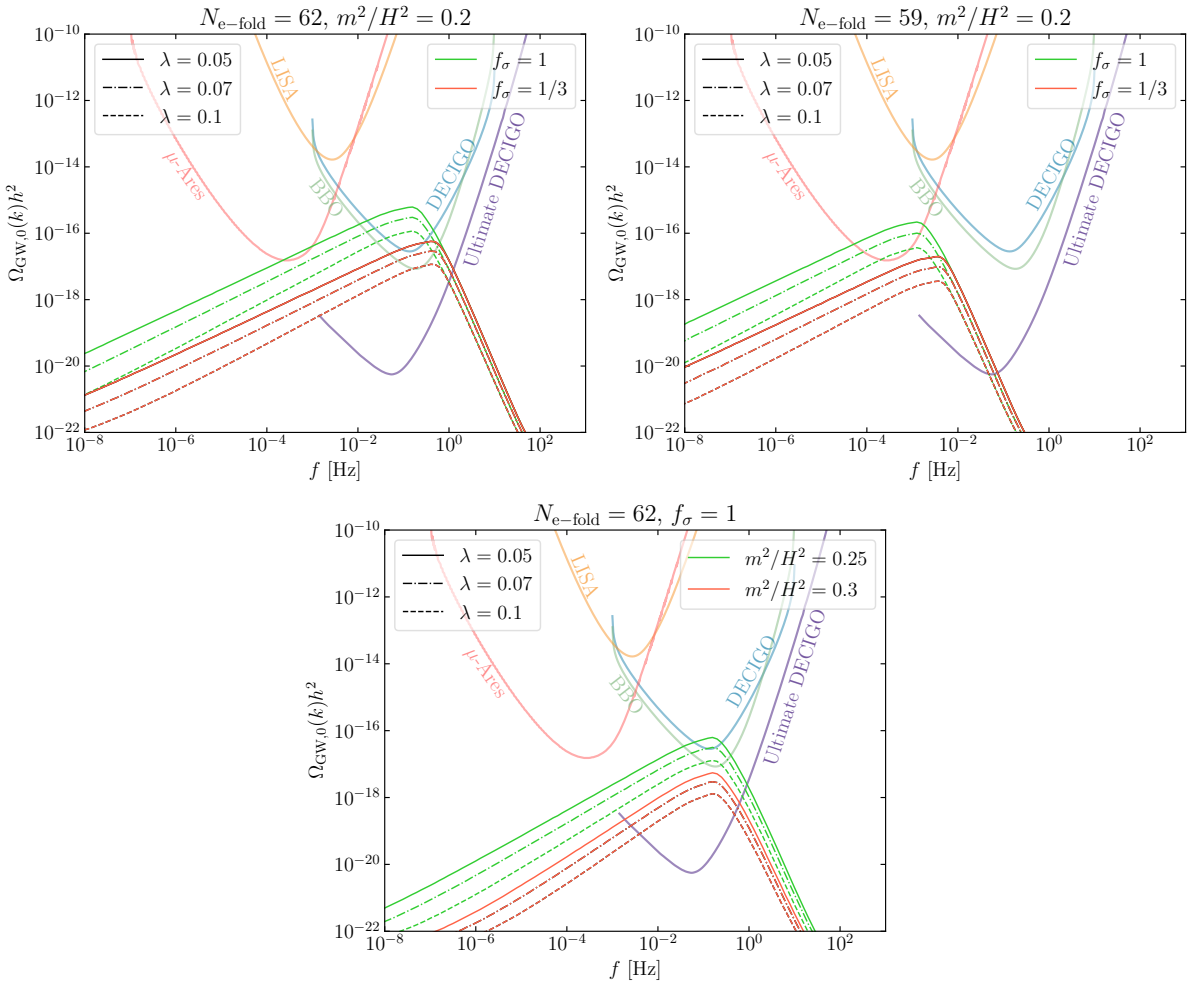

Figure 4.5: Gravitational wave spectrum for the benchmarks discussed in Fig. 4.3. We notice that the number of e-folds after CMB-observable modes exited the horizon determines the peak frequency of the spectrum, and correspondingly, different detectors can be sensitive to the signal. Although a similarly peaked spectrum would appear in the context of cosmological phase transitions (PT), the low-frequency tail of this GW spectrum is different from the usual f^3 tail. While in the context of PT the f^3 scaling originates due to causality and superhorizon behavior of fluctuations, in our scenario, the f-scaling is determined by σ mass. The differing frequency dependence can then be used to discriminate between the two classes of signals.

mechanism of generating a blue-tilted spectrum works for any generic inflaton potential and does not require any particular fine-tuning or structure such as an inflection point or a bump on the potential or an ultra slow-roll phase.

The blue-tilted spectrum gives rise to large curvature perturbations at small scales. These, in turn, source a stochastic gravitational wave background (SGWB) when the perturbations re-enter the horizon. Focusing on some benchmark choices for the number of e-foldings and spectator field potential, we have shown that this scenario predicts observable gravitational waves at future detectors operating in 10^{-5} Hz to 10 Hz range, with strengths $\Omega_{\rm GW}h^2 \simeq 10^{-20} - 10^{-15}$.

There are various interesting future directions. In particular, we have worked in a regime where σ does not dominate the energy density during the cosmological history. It would be interesting to explore the consequences of an early matter-dominated era caused by the σ field. We have also seen that the low-frequency scaling of the SGWB spectrum depends on the mass and coupling of σ and is generally different from the f^3-scaling expected in the context of cosmological PT, or $f^{2/3}$-scaling expected in the context of binary mergers. This different frequency dependence can be used to identify the origin of an SGWB, and distinguish between various cosmological or astrophysical contributions. Along these lines, it would be interesting to carry out a quantitative analysis to understand how well we can separate any two frequency dependencies, for example, by doing a Fisher analysis.

Note Added

While we were finishing this work, results from NANOGrav [252], EPTA, InPTA [253, 254], PPTA [255], CPTA [256] appeared. Secondary gravitational waves from the scalar perturbation can in principle give rise to the signal [257, 258]. Such scalar perturbations

could be generated in a model similar to the one considered in this paper. However, the frequency dependence of $\Omega_{\rm GW}h^2$ determined by the NANOGrav result is [252] 1.8 ± 0.6. We note that for a free field with mass m, the frequency dependence of $\Omega_{\rm GW}h^2$ is given by, $4m^2/(3H^2)$. So for the central value, one would naively infer $m^2/H^2 = 1.4$. Therefore to interpret it in terms of a free field, we require a mass bigger than the Hubble scale. However, since for larger than Hubble-scale masses, the stochastic effects are not efficient, one may have to go beyond the stochastic scenario to explain the NANOGrav observations. We could instead consider a regime in which the misalignment contribution is important [197, 198]. We will leave a detailed analysis of this scenario to future work.

Chapter 5

An Effective Cosmological Collider

5.1 Introduction

Local effective field theory (EFT) provides a powerful way to parametrize the effects of ultraviolet physics in a model-independent way. It has been employed with great success across a wide range of physical settings, from chiral perturbation theory and heavy quark effective theory to the EFT of inflation and the Standard Model (SM) EFT (see [259] and references therein for recent reviews). However, the simple recipe for constructing these EFTs — by assembling the set of local, higher-dimensional operators consistent with the symmetries and field content of the infrared theory — is often complicated by a vast redundancy of description associated with the insensitivity of observables to parameterizations of the fields.

Such redundancies are typically accommodated by identifying a 'basis' of operators that is minimal, independent, and non-redundant with respect to the observables of interest. Operators outside the chosen basis can be expressed in terms of operators within the basis (or in some cases be eliminated entirely) in a variety of ways. For example, integration-by-parts (IBP) can be used to write certain operators in terms of others,

dropping the boundary terms provided that fields vanish at spacetime infinity. Operators that differ by the lowest-order classical equations of motion (EOM) can be exchanged at linear order when on-shell quantities are concerned [260–263]. In flat space, the LSZ reduction formula renders S-matrix elements insensitive to general field redefinitions. Reducing the scope of allowed EFT operators to a minimal basis with the methods at hand greatly facilitates the calculation of observables.

The simplifications arising from a minimal operator basis are perhaps most apparent when computing S-matrix elements in flat space, where EOM and IBP relations lead to a dramatic reduction in the number of operators at a given order in power counting [263]. However, important differences arise in a cosmological context, particularly for an inflationary, quasi-de Sitter (dS) spacetime. In this case, we are interested not in the S-matrix, but rather in the correlation functions of density perturbations at the end of inflation. As a result, arbitrary field redefinitions are not allowed (or must be undone at some stage of the calculation) since such redefinitions would change the correlation functions. Consequently, only a limited set of transformations can be used to remove redundant couplings. Furthermore, the correlation functions are computed on a late-time spatial boundary. As a result, temporal boundary terms may be relevant when implementing IBP relations.

These subtleties become particularly relevant in the context of primordial non-Gaussianity (NG). The approximate scale invariance of primordial density perturbations, inferred through the cosmic microwave background (CMB), indicate that the interactions of the inflaton preserve an approximate shift symmetry, $\phi \to \phi + \text{constant}$. This implies both the self-interactions of the inflaton and interactions of the inflaton with other fields would involve operators with dimension 5 or higher, necessitating an EFT expansion. The question of how to construct a minimal operator basis by eliminating redundant operators arises immediately. Unsurprisingly, there is a long history of identifying non-redundant

self-interactions in the various EFTs of inflation [264, 265], e.g., [266–279].

In this work, we are interested in inflationary EFTs where additional heavy degrees of freedom are present. Such EFTs are especially relevant for the 'Cosmological Collider Physics' program [17, 280] which aims to study oscillatory NG induced by *on-shell* particle production during inflation. Particles with masses of order or larger than the inflationary Hubble scale $H_{\rm inf}$ can be produced as the Universe inflates. Following their production, the heavy particles can oscillate in time, eventually decaying into inflaton fluctuations. Those real-time oscillations of the heavy particles give rise to an oscillatory, scale-dependent NG. Intriguingly, from the frequency of the oscillations and the angular dependence of the NG, one can extract the mass and spin of the heavy particle, respectively. Since $H_{\rm inf} \lesssim 5 \times 10^{13}$ GeV [186], the prospect of doing on-shell mass-spin spectroscopy of such heavy particles through NG provides a unique opportunity to study fundamental physics at high energies. For various interesting work on this subject see Refs. [281–327].

Among the different types of theories that can give rise to a cosmological collider signal, gauge theories are particularly interesting. Of course, such theories play a central role in the SM, and (hidden) gauge theories are also ubiquitous in physics beyond the SM. With this motivation, in the present work we focus on the EFT of a gauge-Higgs sector coupled to the inflaton [295], where the presence of a Higgs field allows us to incorporate spontaneous symmetry breaking (SSB) and study NG mediated by gauge bosons.[1] Furthermore, we focus on a $U(1)$ gauge theory, as our primary goal will be to lay out the procedure of operator basis construction in dS spacetime where heavy fields are present. This can then be generalized to include non-Abelian gauge theories, which exhibit various interesting phenomena during inflation, such as thermalization and

[1] We clarify that by 'Higgs' we will mean a generic complex scalar field that is not necessarily the SM Higgs. It can however be the SM Higgs, for example, if the electroweak scale is uplifted to $H_{\rm inf}$ [295], or if loop corrections to SM are considered [292].

dissipation, as well as fermionic degrees of freedom, e.g., [328–333].

Establishing a minimal basis for the irrelevant interactions between the inflaton and an additional sector is essential to the determination of observable effects.[2] In doing so, one encounters subtleties analogous to those that arise in the study of inflationary self-interactions. Such a minimal basis relevant to cosmological collider signals was first developed in Ref. [285] for a heavy real scalar field coupled to the inflaton at lowest non-trivial order. In this work, we perform a systematic construction of a minimal operator basis for the more general gauge-Higgs-inflaton EFT by considering operators up to dimension 9. We impose an exact shift symmetry on the inflaton (discarding slow roll-suppressed corrections) and consider operators that describe the interactions of the inflaton with the gauge and the Higgs boson. To remove redundant operators, we primarily employ EOM and IBP relations. The EOM relations are closely analogous to those used in flat space [265]. On the other hand, IBP relations at dimension 5 do give rise to non-zero boundary terms that are *a priori* relevant. However, we show that such terms do not contribute to cosmological correlation functions involving heavy particles, and can be removed via appropriate field redefinitions. Beyond dimension 5, we find that IBP can indeed be used as in flat space for the operators of interest. One way to think about the irrelevance of these boundary terms is that they arise from interactions of the inflaton with heavy fields. The mode functions of the heavy fields decay at late times, corresponding to the physical effect that heavy particles get diluted as the Universe expands. Consequently, the temporal boundary terms resulting from IBP relations vanish for the operators and observables of interest.

To highlight some of our findings, we show that the only dimension 5 operator that

[2]Of course, there are also more direct routes to mapping the space of cosmological observables that are free of EFT operator redundancies, as in the cosmological bootstrap (see, e.g., [334] for a recent review). Nonetheless, exploring observables from the standpoint of EFT Lagrangians can be useful for interpreting the microscopic implications of data and estimating the observability of certain signatures, motivating the approach taken here.

is not redundant is an axionic coupling of the inflaton to the gauge field, namely $\phi F\tilde{F}$. In particular, by using EOM we find that an operator coupling the Higgs current to the inflaton is redundant, and the leading Higgs-inflaton coupling arises only at dimension 6. The same conclusion can be drawn using a field redefinition argument, as we discuss in an appendix. We also show that in the broken phase of the theory, a quadratic mixing between the inflaton and the longitudinal mode of the gauge boson, relevant for tree-level bispectrum signatures, first arises at dimension 9. Finally, we identify new operators at dimensions 7 and 8 involving the inflaton and gauge boson. These operators are expected to contribute with similar strengths for NG compared to some other operators that have been considered in the previous literature. While certain higher dimensional operators of the gauge-Higgs-inflaton theory have been considered in isolation in the previous literature, our systematic approach to enumerating an operator basis identifies additional operators that would be present in a generic EFT and could contribute to NG signals.

The rest of the chapter is organized as follows: We discuss aspects of the choice of operator basis on inflationary observables in Sec. 5.2, with a particular emphasis on the effect of boundary terms generated by integration-by-parts manipulations. In Sec. 5.3 we introduce an EFT of the inflaton coupled to an abelian gauge-Higgs theory, enumerating operators up to dimension 9 and reducing them to a minimal operator basis. Although boundary terms appear in certain cases, they do not affect the correlation functions of interest. With the minimal basis at hand, we enumerate interaction vertices and estimate the leading sizes of non-Gaussianity in Sec. 5.4. We conclude in Sec. 5.5. A number of general considerations and examples regarding IBP and boundary terms in inflationary spacetimes can be found in the appendices.

Notations and Conventions. We follow the 'mostly plus' metric signature: $(-,+,+,+)$. The operator ∇_μ denotes an ordinary covariant derivative, while the operator $D_\mu \equiv \nabla_\mu + ig_A A_\mu$ denotes a gauge covariant derivative. Unless explicitly stated, we will use units in which the Hubble scale during inflation, $H_{\rm inf} = 1$. Factors of $H_{\rm inf}$ can be restored using dimensional analysis.

5.2 Inflationary Observables and Operator Bases

The precise nature of cosmic inflation is still unknown. Different classes of mechanisms can explain the homogeneous cosmic expansion and the generation of primordial fluctuations during inflation. To capture certain model-independent features and signatures, it is therefore useful to construct an EFT consistent with the symmetries and the particle content of the theory. In this regard, there are two qualitatively different classes of EFTs relevant during inflation. The more UV-agnostic of the two treats the inflaton as a Goldstone boson arising from spontaneous breaking of time translation symmetry [264]. In this EFT, Lorentz invariance is (spontaneously) broken and as a result, one can allow qualitatively new sets of higher dimensional operators, in addition to the ones that follow from requiring Lorentz invariance.

Another class of EFT [265] is useful if we are to describe both the inflationary fluctuations *and* the homogeneous inflationary expansion, since the latter is not necessarily part of the Goldstone EFT [264]. The advantage of this second class of EFTs is that it could be valid up to a much higher energy scale and can more readily describe how reheating can happen at the end of inflation. While the Lorentz-breaking, Goldstone EFT would allow for the greatest generality, for concreteness, in this work we will focus on a Lorentz-invariant EFT description and assume that inflation is driven by a slowly rolling scalar field ϕ. Taking a bottom-up approach, we will also impose a strict shift symmetry on ϕ:

$\phi \to \phi +$ constant, motivated by the approximate scale invariance of primordial perturbations, and neglect subleading corrections from violation of this symmetry. Therefore, all the operators involving ϕ that we consider below will have (sometimes implicitly) derivative coupling $\nabla_\mu \phi$.

Our analysis will also encompass scenarios where the density fluctuations originate not from the inflaton ϕ, but a curvaton field σ, as in the curvaton paradigm [209–212]. In such scenarios, while the spacetime expansion is driven by ϕ, the density fluctuations in the late Universe originate from σ. However, to obtain (approximately) scale-invariant, superhorizon fluctuations, the mass of σ needs to be much smaller than $H_{\rm inf}$. Therefore, we can still impose a shift symmetry on σ. Consequently, our following analysis will exactly carry over to the curvaton scenario, with the replacement $\nabla_\mu \phi \to \nabla_\mu \sigma$. With this in mind, in the rest of the discussion we will focus on the standard inflationary slow-roll EFT where both the homogeneous expansion and fluctuations are sourced by ϕ.

5.2.1 Minimal Operator Bases

The approximate shift symmetry acting on the inflaton implies that the couplings between the inflaton and additional fields are necessarily irrelevant operators that may be organized systematically according to the relevant power-counting scheme. However, the full set of irrelevant operators consistent with the symmetries of the EFT is generally over-complete, leading to a redundancy of description whose severity depends on the observables of interest. In flat space where the observables are typically related to S-matrix elements, the redundancy of description corresponds to the freedom to perform nearly-arbitrary field redefinitions without altering the S-matrix elements. This can be used to arrive at minimal, non-redundant operator bases order-by-order in the power

counting, where the number of operators in a minimal basis is typically much smaller than the total number of operators consistent with symmetries. In practice, a non-redundant basis of operators can usually be obtained order-by-order in power counting by using the lowest-order equations of motion to eliminate operators.[3] Operators that differ by total derivatives can also be exchanged via integration-by-parts (IBP), as both spatial and temporal boundary terms are assumed to vanish in flat space.

The situation is somewhat different in cosmological contexts. In inflationary scenarios, we are interested in computing cosmological correlation functions at a fixed time slice towards the end of inflation, or when all the modes associated with the correlation function have exited the horizon. To be specific, we use the Poincare patch representation of dS spacetime

$$\mathrm{d}s^2 = \frac{-\mathrm{d}\eta^2 + \mathrm{d}\mathbf{x}^2}{\eta^2}, \tag{5.1}$$

and denote the conformal time on the time slice of interest by $\eta = \eta_0$. We then take $\eta_0 \to 0$ limit of the cosmological correlators to obtain the conserved correlation functions. Compared to the usual Minkowski spacetime, the presence of this boundary at η_0, where we evaluate the correlation functions, requires a reexamination of the standard operator basis manipulations, in particular those involving IBP. This is because the boundary terms on the space-like surface at η_0 may not vanish in general. The potential importance of boundary terms is already apparent for a massless free field in dS [266].

[3] There are various subtleties involved when using equations of motion to eliminate operators, enumerated in [335].

5.2.2 Massless Free Field in dS

Consider the Lagrangian of a free massless field and an equivalent expression obtained via IBP,

$$-\frac{1}{2}\int \mathrm{d}^4x\sqrt{-g}g^{\mu\nu}\nabla_\mu\phi\nabla_\nu\phi = -\frac{1}{2}\int \mathrm{d}^4x\sqrt{-g}g^{\mu\nu}\nabla_\mu(\phi\nabla_\nu\phi) + \frac{1}{2}\int \mathrm{d}^4x\sqrt{-g}\phi\Box\phi, \quad (5.2)$$

where $\Box\phi = g^{\mu\nu}\nabla_\mu\nabla_\nu\phi$. Using the EOM $\Box\phi \approx 0$ (neglecting $V(\phi)$) and Stokes' theorem we arrive at

$$-\frac{1}{2}\int \mathrm{d}^4x\sqrt{-g}g^{\mu\nu}\nabla_\mu\phi\nabla_\nu\phi = -\frac{1}{2}\int \mathrm{d}^3x\sqrt{\gamma}n_\mu(\phi\nabla^\mu\phi). \quad (5.3)$$

Here the spatial integration is over a space-like surface at η_0 and we have assumed the fields vanish at spatial infinity as well as at very early times when the fields are in their vacuum states. The vector $n_\mu = (1/\eta_0, 0, 0, 0)$ is normal to the space-like surface on which the induced metric is given by γ_{ij}, with a determinant $\gamma = 1/\eta_0^6$. Thus the (on-shell) action of a massless free field in dS can be written as a boundary term at η_0 which can be simplified as,

$$\frac{1}{2\eta_0^2}\int \mathrm{d}^3x\ \phi\partial_\eta\phi|_{\eta_0} = \frac{1}{2\eta_0^2}\int \frac{\mathrm{d}^3k}{(2\pi)^3}\ \phi_{\mathbf{k}}\partial_\eta\phi_{-\mathbf{k}}|_{\eta_0}\,. \quad (5.4)$$

To compute correlation functions, we can first derive the associated wavefunction which can be schematically written as (see [334] for a recent review),

$$\Psi[\varphi,\eta_0] = \int_{\phi(\eta_0)=\varphi,\phi(-\infty)=0}\mathcal{D}\phi e^{iS[\phi]} \propto e^{iS_{\mathrm{cl}}[\varphi,\eta_0]}, \quad (5.5)$$

where $\phi(\eta_0) = \varphi$ is the late time boundary condition while $\phi(-\infty) = 0$ ensures that the fields are in their vacuum state at early times. We have also evaluated the path integral using the saddle point approximation, up to a proportionality constant. With these choices, we can write

$$\phi_{\mathbf{k}} = \varphi_{\mathbf{k}}\frac{(1-ik\eta)e^{ik\eta}}{(1-ik\eta_0)e^{ik\eta_0}}. \quad (5.6)$$

The classical action is then given by (upon using $\varphi_{-\mathbf{k}} = \varphi^*_{\mathbf{k}}$),

$$S_{\rm cl} = \frac{1}{2\eta_0^2}\int \frac{{\rm d}^3k}{(2\pi)^3}\frac{k^2\eta_0}{(1-ik\eta_0)}|\varphi_{\mathbf{k}}|^2. \tag{5.7}$$

A correlation function at a late time, η_0 is given by

$$\langle\varphi(\mathbf{k}_1)\varphi(\mathbf{k}_2)\cdots\varphi(\mathbf{k}_n)\rangle = \frac{\int\mathcal{D}\varphi\varphi(\mathbf{k}_1)\varphi(\mathbf{k}_2)\cdots\varphi(\mathbf{k}_n)|\Psi[\varphi,\eta_0]|^2}{\int\mathcal{D}\varphi|\Psi[\varphi,\eta_0]|^2}. \tag{5.8}$$

Given the appearance of $|\Psi[\varphi,\eta_0]|^2$, only the imaginary terms in $S_{\rm cl}$ contribute to the determination of the correlation function. From (5.7), the part that diverges as $\eta_0 \to 0$ then does not contribute and the surviving contribution is given by

$$S_{\rm cl} \approx \int \frac{{\rm d}^3k}{(2\pi)^3}\frac{ik^3}{2}|\varphi_{\mathbf{k}}|^2. \tag{5.9}$$

We can now evaluate the two-point function using this wavefunction,

$$\begin{aligned}\langle\varphi(\mathbf{k}_1)\varphi(\mathbf{k}_2)\rangle &= \frac{\int\mathcal{D}\varphi\varphi(\mathbf{k}_1)\varphi(\mathbf{k}_2)|\Psi[\varphi,\eta_0]|^2}{\int\mathcal{D}\varphi|\Psi[\varphi,\eta_0]|^2},\\ &= \frac{\int\mathcal{D}\varphi\varphi(\mathbf{k}_1)\varphi(\mathbf{k}_2)\exp\left(-\int\frac{{\rm d}^3k}{(2\pi)^3}\frac{k^3}{2}|\varphi_{\mathbf{k}}|^2\right)}{\int\mathcal{D}\varphi\exp\left(-\int\frac{{\rm d}^3k}{(2\pi)^3}\frac{k^3}{2}|\varphi_{\mathbf{k}}|^2\right)},\\ &= \frac{1}{2k_1^3}(2\pi)^3\delta(\mathbf{k}_1+\mathbf{k}_2),\end{aligned} \tag{5.10}$$

as can also be derived using the standard 'in-in' computation (see [294] for a pedagogical review). This example illustrates that the imaginary part of $S_{\rm cl}$ is relevant for computing the correlation function, while the real part drops out from $|\Psi[\varphi,\eta_0]|^2$.

There is another way to reach the same conclusions as above, highlighting the role of the boundary. We can treat the fields in (5.4) as quantum operators, instead of classical functions. The two-point function can then be computed using the standard 'in-in' approach. Since the correlation functions are evaluated on the spatial surface and Eq. (5.4) is also evaluated on the same surface, we need to evaluate only equal-time propagators. We also note that the usual time evolution operator $T(\exp(-i\int{\rm d}t\,\mathbb{H}))$, with

$\mathbb{H}$ the Hamiltonian, can be schematically written as a boundary term $\sim \exp(-if(\eta_0))$. Thus the time ordering part gives a factor of $(-i)$, while the anti-time ordering part gives the complex conjugate factor $(+i)$, as in a standard bulk computation. With these ingredients, the result is given by

$$\begin{aligned}\langle\phi(\mathbf{k}_1)\phi(\mathbf{k}_2)\rangle' &= 2\times(-i)\frac{1}{2\eta_0^2}\frac{1}{4k_1^6}(1+ik_1\eta_0)^2(1-ik_1\eta_0)k_1^2\eta_0 + \text{c.c.}\\ &= \frac{1}{\eta_0}\frac{1}{4k_1^4}(-i+k_1\eta_0)(1+k_1^2\eta_0^2) + \text{c.c.}\\ &= \frac{1}{2k_1^3} + \mathcal{O}(\eta_0).\end{aligned} \tag{5.11}$$

Here we have used the standard notation $\langle\phi(\mathbf{k}_1)\phi(\mathbf{k}_2)\rangle \equiv (2\pi)^3\delta(\mathbf{k}_1+\mathbf{k}_2)\langle\phi(\mathbf{k}_1)\phi(\mathbf{k}_2)\rangle'$. Further examples of the relevance of boundary terms for massless and massive scalars, both free and with derivatively-coupled cubic interactions, are presented in Appendix D.2.

5.2.3 Operators Coupled to the Inflaton

In what follows, we will use arguments similar to those presented above to understand whether certain boundary terms contribute or not in determining cosmological correlators. While the boundary term was essential in the previous example, in many cases it can be neglected. In particular, we often encounter operators of the type

$$\int \mathrm{d}^4x\sqrt{|g|}\nabla_\mu\phi\nabla^\mu\mathcal{O}, \tag{5.12}$$

where $\mathcal{O}$ is any (composite) operator containing massive fields. Using an IBP we can write the above as

$$\int \mathrm{d}^4x\,\sqrt{|g|}\left(\nabla^\mu\left[\nabla_\mu\phi\cdot\mathcal{O}\right] - \Box\phi\cdot\mathcal{O}\right). \tag{5.13}$$

The second term does not contribute in any vertex for an 'in-in' diagram since $\Box f = 0$ where f is a mode function for the inflaton. The first term, on the other hand, can be

written as

$$\int \mathrm{d}^4x\, \partial_\mu \left(\sqrt{|g|} \nabla^\mu \phi \cdot \mathcal{O} \right). \tag{5.14}$$

For the spatial component, i.e., for $\mu = i$, the above does not contribute under the assumption that fields vanish at spatial infinity. Therefore, the only potentially non-trivial term is the one involving time derivatives,

$$-\int \mathrm{d}^4x\, \partial_\eta \left(\frac{1}{\eta^2} \partial_\eta \phi \cdot \mathcal{O} \right). \tag{5.15}$$

This determines the interacting Hamiltonian of interest,

$$\mathbb{H}_I = \int \mathrm{d}^3x\, \partial_\eta \left(\frac{1}{\eta^2} \partial_\eta \phi \cdot \mathcal{O} \right). \tag{5.16}$$

However, since this is a total time derivative, we can evaluate the time evolution operator as

$$T \exp\left(-i \int_{-\infty}^{\eta_0} \mathrm{d}\eta\, \mathbb{H}_I \right) = T \exp\left(-i \int \mathrm{d}^3x \left[\frac{1}{\eta^2} \partial_\eta \phi \cdot \mathcal{O} \right]_{\eta_0} \right). \tag{5.17}$$

The last term is evaluated at η_0 and it does not involve any time integrals. Therefore, the time ordering operator acts trivially. We can then conclude that if the term in the square brackets vanishes at η_0, the entire operator does not contribute to correlation functions involving the massive particle.[4]

5.3 The Abelian Gauge-Higgs-Inflaton EFT

We're now equipped to construct minimal operator bases in dS for EFTs where a shift-symmetric inflaton couples to additional fields. For concreteness, in this article we

[4] A detailed example of the action of the time ordering operator can be found in Appendix D.1, where we show that while boundary terms can be present in the intermediate stages of a computation, they do not give a (non-local) cosmological collider signal with the characteristic non-analytic momentum dependence. Rather, the boundary terms give a local contribution which can be accounted for by appropriate local field redefinitions.

will take the additional fields to comprise an abelian gauge-Higgs sector. This theory is of considerable interest as a benchmark for various cosmological collider signals, and captures most of the relevant features involved in constructing a minimal operator basis. A similar procedure can be followed for other effective theories containing different fields, such as fermions or non-abelian gauge bosons.

The Lagrangian containing the inflaton (ϕ), a Higgs ($\mathcal{H}$), and a $U(1)$ gauge field A_μ, up to dimension-4 is given by

$$\mathcal{L} \supset -\frac{1}{2}\nabla_\mu\phi\nabla^\mu\phi - V(\phi) - (D_\mu\mathcal{H})^\dagger D^\mu\mathcal{H} - V(|\mathcal{H}|^2) - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}. \tag{5.18}$$

Here $D_\mu\mathcal{H} = \nabla_\mu\mathcal{H} + ig_A A_\mu\mathcal{H}$ is the gauge covariant derivative and $V(|\mathcal{H}|^2)$, $V(\phi)$ are respectively the Higgs and inflaton potentials, whose detailed forms will not be important for our purposes. We will see in some cases that the surviving contributions from a given operator are slow roll-suppressed, in the sense of involving $\mathrm{d}V(\phi)/\mathrm{d}\phi$ or $\mathrm{d}^2V(\phi)/\mathrm{d}\phi^2$. We will not track such operators explicitly, under the assumption that their slow roll-suppressed contributions to observables are subdominant to other contributions. We assume the Higgs potential is such that it can acquire a vacuum expectation value $\langle\mathcal{H}\rangle = v/\sqrt{2}$, so that all states in the abelian gauge-Higgs sector are massive in the broken phase. Alternately, our results can also be applied to the theory of a shift-symmetric inflaton coupled to a complex scalar with a global $U(1)$ symmetry by taking the $g_A \to 0$ limit and assuming $\langle\mathcal{H}\rangle = 0$.

In reducing the operator basis, we will primarily use the following EOM and also implement IBP. The EOM for the inflaton is given by,

$$\Box\phi = V'(\phi), \quad \text{[Inflaton EOM]} \tag{5.19}$$

where we have denoted $\Box \equiv \nabla^\mu \nabla_\mu$. To obtain the EOM for the Higgs, we first expand

$$-(D_\mu \mathcal{H})^\dagger D^\mu \mathcal{H} = -\nabla_\mu \mathcal{H}^\dagger \nabla^\mu \mathcal{H} + i g_A A_\mu \mathcal{H}^\dagger \nabla^\mu \mathcal{H} - i g_A A^\mu \mathcal{H} \nabla_\mu \mathcal{H}^\dagger - g_A^2 A_\mu A^\mu \mathcal{H}^\dagger \mathcal{H}. \tag{5.20}$$

The EOM is then given by,

$$-\nabla_\mu \left(\nabla^\mu \mathcal{H} + i g_A A^\mu \mathcal{H}\right) = +i g_A A_\mu \nabla^\mu \mathcal{H} - g_A^2 A_\mu A^\mu \mathcal{H} - V'(|\mathcal{H}|^2)\mathcal{H}, \tag{5.21}$$

which can be written in terms of the gauge covariant derivative,

$$-\nabla_\mu D^\mu \mathcal{H} = i g_A A_\mu D^\mu \mathcal{H} - V'(|\mathcal{H}|^2)\mathcal{H}. \quad \text{[Higgs EOM]} \tag{5.22}$$

The EOM for the gauge field is given by

$$-\nabla_\mu F^{\mu\nu} = i g_A \left(\mathcal{H}^\dagger D^\nu \mathcal{H} - (D^\nu \mathcal{H})^\dagger \mathcal{H}\right). \quad \text{[Gauge Field EOM]} \tag{5.23}$$

The symmetries of the theory forbid relevant or marginal couplings between the inflaton and the abelian gauge-Higgs sector, so interactions are necessarily irrelevant. We assume the gauge-Higgs sector is weakly coupled and the Higgs VEV is parametrically smaller than the characteristic UV scale Λ suppressing the irrelevant operators, so that the natural power counting is in terms of the classical dimension of operators constructed out of the fields in the unbroken phase.

In what follows, we enumerate operators up to dimension 9, beginning with the complete set of operators at a given dimension allowed by symmetries, modulo some operators trivially related by EOM. We then reduce the operators to a minimal basis at a given order via EOM and IBP relations, taking care to verify that boundary terms arising from IBP do not contribute to the observables of interest. Since we are taking a bottom-up approach in which the Wilson coefficients of different operators are free parameters, we may use the lowest-order EOM to arrive at a minimal basis at a given

order in power-counting. Although these EOM manipulations do not correctly account for changes to Wilson coefficients at higher order in power counting, the values of these coefficients were already arbitrary. In this way, we can fix the operator basis using the lowest-order EOM by starting at dimension-5 and proceeding to successively higher dimensions. Note that more care would be required in manipulating operator bases when matching to a specific UV completion in which all Wilson coefficients take on specific values [335].

Needless to say, the number of possible operators grows rapidly with operator dimension. Although it is not too cumbersome to enumerate operators up to dimension-9 by hand, we have also cross-checked our results against the flat-space operator basis codes `DEFT` [336] and `Sym2Int` [337, 338].

5.3.1 Dimension 5

We start our analysis of irrelevant operators at dimension 5. While there are a number of operators consistent with the assumed symmetries, we show there is only one operator that contributes non-trivially to cosmological correlators. We first summarize the operators in Table 5.1. The Wilson coefficient for each operator is taken to be real; note that operators such as $\mathcal{O}_{5,1}$ and $\mathcal{O}_{5,2}$ can be interpreted as the real and imaginary parts of a single operator with a complex Wilson coefficient.

We note that $\mathcal{O}_{5,1}$ can be simplified as

$$\mathcal{O}_{5,1} = \nabla_\mu \phi \nabla^\mu (\mathcal{H}^\dagger \mathcal{H}). \tag{5.24}$$

We can use the EOM (5.23) for the gauge field and subsequently an IBP to write

$$\mathcal{O}_{5,2} = \frac{1}{g_A} \nabla_\nu \left(F^{\nu\mu} \nabla_\mu \phi \right). \tag{5.25}$$

Operator	Expression
$\mathcal{O}_{5,1}$	$\nabla_\mu\phi\left(\mathcal{H}^\dagger D^\mu\mathcal{H}+(D^\mu\mathcal{H})^\dagger\mathcal{H}\right)$
$\mathcal{O}_{5,2}$	$(-i)\nabla_\mu\phi\left(\mathcal{H}^\dagger D^\mu\mathcal{H}-(D^\mu\mathcal{H})^\dagger\mathcal{H}\right)$
$\mathcal{O}_{5,3}$	$\nabla_\mu\phi\nabla_\nu F^{\nu\mu}$
$\mathcal{O}_{5,4}$	$\phi F_{\mu\nu}\tilde{F}^{\mu\nu}$

Table 5.1: Allowed operators at dimension 5. Here and henceforth, $\tilde{F}^{\mu\nu}=(1/2)\epsilon^{\mu\nu\rho\sigma}F_{\rho\sigma}$.

In the process we have dropped a contribution of the type $\nabla_\nu\nabla_\mu\phi\cdot F^{\mu\nu}$ which vanishes identically for a torsion-free metric. This manipulation also shows that $\mathcal{O}_{5,2}$ is equivalent to $\mathcal{O}_{5,3}$. To comprehensively study the fate of $\mathcal{O}_{5,1}$ and $\mathcal{O}_{5,2}$, it is useful to consider both the unbroken and broken phases of the theory.

Unbroken Phase

We start with $\mathcal{O}_{5,1}$ which after an IBP gives,

$$\mathcal{O}_{5,1}=\int \mathrm{d}^4x\sqrt{-g}\left[\nabla^\mu(\nabla_\mu\phi\cdot\mathcal{H}^\dagger\mathcal{H})-\Box\phi\cdot\mathcal{H}^\dagger\mathcal{H}\right]. \tag{5.26}$$

The second term vanishes in the limit of vanishing inflaton potential and we will not consider it further. The first term is a boundary term that reduces to a spatial integral at η_0,

$$\mathcal{O}_{5,1}=\int \mathrm{d}^3x\sqrt{\gamma}n^\mu(\nabla_\mu\phi\cdot\mathcal{H}^\dagger\mathcal{H})=-\frac{1}{\eta_0^2}\int \mathrm{d}^3x\partial_\eta\phi\cdot\mathcal{H}^\dagger\mathcal{H}. \tag{5.27}$$

From (5.6), we note

$$\partial_\eta\phi_{\mathbf{k}}=\varphi_{\mathbf{k}}\frac{k^2\eta_0}{(1-ik\eta_0)}. \tag{5.28}$$

Given the late time scaling of $\mathcal{H}(\eta,\mathbf{k})\sim\eta_0^{3/2\pm i\mu}$ (as can be shown by considering mode functions of a massive particle, see, e.g., [320]), where $\mu\equiv\sqrt{m^2/H_{\rm inf}^2-9/4}$ is taken to

be positive, we have

$$\mathcal{O}_{5,1} \propto \eta_0^2 \to 0. \quad \text{[Unbroken Theory]} \tag{5.29}$$

Therefore, $\mathcal{O}_{5,1}$ does not contribute to late-time correlators.

We now consider $\mathcal{O}_{5,2}$ which is also a boundary term and can be rewritten as,

$$\mathcal{O}_{5,2} = \frac{1}{g_A}\int \mathrm{d}^3x\sqrt{\gamma}n_\nu\nabla_\mu\phi F^{\nu\mu} = -\frac{1}{g_A}\int \mathrm{d}^3x\partial_i\phi F_{\eta i}. \tag{5.30}$$

For an unbroken gauge theory, the physical degrees of freedom are the transverse components $A_i^\perp$ which satisfy $k_i A_i^\perp = 0$. Therefore, $\mathcal{O}_{5,2}$ does not contribute to late-time correlators. Note, that this argument does not rely on the vanishing of the mode functions at late times.

Broken Phase

We now repeat the above analysis for the case of a broken gauge theory starting with $\mathcal{O}_{5,1}$. We can still implement an IBP to write it as

$$\mathcal{O}_{5,1} = -\frac{1}{\eta_0^2}\int \mathrm{d}^3x\partial_\eta\phi\cdot\mathcal{H}^\dagger\mathcal{H}. \tag{5.31}$$

In the broken gauge theory, we can set one of the Higgs to its VEV to obtain a term quadratic in fluctuations. However, the result still scales as,

$$\mathcal{O}_{5,1} \propto \eta_0^{1/2} \to 0 \quad \text{[Broken Theory]}. \tag{5.32}$$

Hence $\mathcal{O}_{5,1}$ does not contribute in the broken gauge theory as well.

Next we turn to $\mathcal{O}_{5,2}$, which can be written as

$$\mathcal{O}_{5,2} = -\frac{1}{g_A}\int \mathrm{d}^3x\partial_i\phi F_{\eta i} = -\frac{1}{g_A}\int \mathrm{d}^3x\partial_i\phi(\partial_\eta A_i^{\parallel} - \partial_i A_\eta). \tag{5.33}$$

However, we need to take into account the longitudinal component of the gauge boson which is a combination of A_η and $A_i^{\parallel}$. The temporal component A_η falls as $\eta_0^{3/2\pm i\mu}$, with $\mu \equiv (m^2/H_{\rm inf}^2 - 1/4)^{1/2} > 0$, at late times, as can be seen from the massive gauge boson mode functions, e.g., [289]. Therefore the term proportional to $\partial_i A_\eta$ in (5.33) vanishes in the late time limit. After a spatial IBP, the surviving term can be written as

$$\mathcal{O}_{5,2} = \frac{1}{g_A}\int \mathrm{d}^3x \phi \partial_\eta \partial_i A_i^{\parallel}. \tag{5.34}$$

We can rewrite the above after using the constraint equation for the massive field,

$$\nabla_\mu A^\mu = 0 \Rightarrow -\eta^2 \partial_\eta A_\eta + 2\eta A_\eta + \eta^2 \partial_i A_i^{\parallel} = 0, \tag{5.35}$$

$$\mathcal{O}_{5,2} = \frac{1}{g_A}\int \mathrm{d}^3x\, \phi \partial_\eta \left(\partial_\eta A_\eta - \frac{2}{\eta} A_\eta\right). \tag{5.36}$$

This can be further simplified by using the EOM for A_η,

$$\partial_\eta^2 A_\eta - \frac{2}{\eta}\partial_\eta A_\eta - \partial_i^2 A_\eta + \frac{m^2+2}{\eta^2} A_\eta = 0, \tag{5.37}$$

$$\mathcal{O}_{5,2} = -\frac{m^2}{g_A \eta_0^2}\int \mathrm{d}^3x\, \phi A_\eta. \tag{5.38}$$

While this term does not vanish as $\eta_0 \to 0$, it does not contribute to a late-time correlation function. As an example, we can evaluate the contribution to the two-point inflaton correlation function from $\mathcal{O}_{5,2}$. That has a scaling:

$$\langle \phi(\mathbf{k}_1)\phi(\mathbf{k}_2)\rangle \propto \frac{1}{\eta_0^4} \times \eta_0^3 (1+k_1^2\eta_0^2)(1+k_1^2\eta_0^2)((-i)^2 + (+i)^2 + (+i)(-i) + (+i)(-i)) = 0, \tag{5.39}$$

where the last factor in the parenthesis comes from summing over the four 'in-in' subdiagrams, while the factor of η_0^3 comes from the massive A_η propagator at late times.

More generally, each factor of $\mathcal{O}_{5,2}$ appears in a correlator involving the inflaton field on the late time surface with a structure

$$\propto (-i)(1-ik\eta_0)\frac{|f_\eta(\eta_0,k)|}{\eta_0^2} + (+i)(1+ik\eta_0)\frac{|f_\eta(\eta_0,k)|}{\eta_0^2} = -2k\frac{|f_\eta(\eta_0,k)|}{\eta_0}, \tag{5.40}$$

where $A_\eta(\eta, \mathbf{k}) = f_\eta(\eta, k) b^\dagger_{\mathbf{k}} + f^*_\eta(\eta, k) b_{-\mathbf{k}}$, with $f_\eta(\eta, k)$ a mode function and $b^\dagger_{\mathbf{k}}$ a creation operator, along with their conjugates. The prefactors are $(-i)$ or $(+i)$ depending on whether the contribution originally came from a time ordering or anti-time ordering term.[5] We have also kept only the absolute value $|f_\eta(\eta_0, k)|$ since, in an inflaton correlator on the late time surface, we always have the combination $|f_\eta(\eta_0, k)|^2$ from the longitudinal mode propagator, so each factor of A_η effectively contributes a factor of $|f_\eta(\eta_0, k)|$. Noting that $|f_\eta(\eta_0, k)| \sim \eta_0^{3/2}$ for $\mu > 0$, we see $\mathcal{O}_{5,2}$ does not contribute to a correlation function from contractions on the late time surface. It can also be checked that bulk contractions with the operators in Table 5.5 vanish as $\eta_0 \to 0$. Therefore, $\mathcal{O}_{5,2}$ does not contribute to cosmological correlation functions overall.

Surviving Contribution at Dimension 5

The above analysis shows that the three operators $\mathcal{O}_{5,1}$, $\mathcal{O}_{5,2}$, and $\mathcal{O}_{5,3}$ are all redundant. The remaining operator is $\mathcal{O}_{5,4}$ which is also shift symmetric since $F_{\mu\nu}\tilde{F}^{\mu\nu}$ is a total derivative. This operator is non-trivial and has been discussed extensively in the context of gauge field production during inflation (see, e.g., [328]), and cosmological collider [339].

5.3.2 Dimension 6

At dimension 6, we only have one operator coupling the inflaton field to the gauge-Higgs sector,

$$\mathcal{O}_{6,1} \equiv (\nabla_\mu \phi)^2 \, \mathcal{H}^\dagger \mathcal{H}. \tag{5.41}$$

[5] Since here we are interested in evaluating just surface terms at η_0, the action of time ordering or anti-time ordering is trivial. However, the factors of $(-i)$ or $(+i)$ are still present depending on whether terms originally came from time ordering operator $T\exp(-i\int \mathrm{d}t\ \mathbb{H})$ or anti-time ordering operator $(T\exp(-i\int \mathrm{d}t\ \mathbb{H}))^\dagger$.

Operator	Expression
$\mathcal{O}_{7,1}$	$\lvert\mathcal{H}\rvert^2\nabla_\mu\phi\left(\mathcal{H}^\dagger D^\mu\mathcal{H}+(D^\mu\mathcal{H})^\dagger\mathcal{H}\right)$
$\mathcal{O}_{7,2}$	$\lvert\mathcal{H}\rvert^2\nabla_\mu\phi\left(\mathcal{H}^\dagger D^\mu\mathcal{H}-(D^\mu\mathcal{H})^\dagger\mathcal{H}\right)$
$\mathcal{O}_{7,3}$	$F^{\mu\nu}\nabla_\mu\phi\left(\mathcal{H}^\dagger D_\nu\mathcal{H}+(D_\nu\mathcal{H})^\dagger\mathcal{H}\right)$
$\mathcal{O}_{7,4}$	$(-i)F^{\mu\nu}\nabla_\mu\phi\left(\mathcal{H}^\dagger D_\nu\mathcal{H}-(D_\nu\mathcal{H})^\dagger\mathcal{H}\right)$

Table 5.2: Allowed operators at dimension 7.

This term is not reducible to another operator, and it contributes both in the broken and unbroken phase, via tree and loop-level diagrams, respectively. The associated signatures have been discussed in [291, 295].

5.3.3 Dimension 7

At this dimension, there are four possible operators to start with, as summarized in Table 5.2. We start the analysis with $\mathcal{O}_{7,1}$ which after IBP can be written as

$$\mathcal{O}_{7,1}=\nabla_\mu\phi\cdot\mathcal{H}^\dagger\mathcal{H}\nabla^\mu(\mathcal{H}^\dagger\mathcal{H})=\frac{1}{2}\nabla_\mu\phi\nabla^\mu|\mathcal{H}|^4=\frac{1}{2}\nabla^\mu\left(|\mathcal{H}|^4\nabla_\mu\phi\right), \tag{5.42}$$

where in the last line we have dropped a contribution proportional to $\Box\phi$ as that is slow roll-suppressed. The surviving term is a boundary term, which we can evaluate by following the procedure detailed in the previous section. On a late time surface at η_0, it scales as

$$\mathcal{O}_{7,1}\sim\sqrt{\gamma}n_\eta g^{\eta\eta}|\mathcal{H}|^4\partial_\eta\phi\sim\frac{1}{\eta_0^4}\times\eta_0^2\times|\mathcal{H}|^4\times\eta_0. \tag{5.43}$$

Now we need to set at least one of the Higgs to its fluctuation, otherwise we would just have a dimension 5 operator considered before. Since the $\mathcal{H}$ mode function has a scaling $\eta_0^{3/2}$, up to oscillatory parts, we see that the entire boundary term vanishes. Therefore $\mathcal{O}_{7,1}$ does not contribute.

Next we consider $\mathcal{O}_{7,2}$. We can use the EOM for the gauge field to write

$$\mathcal{O}_{7,2} = \frac{i}{g_A}|\mathcal{H}|^2\nabla_\mu\phi\nabla_\nu F^{\nu\mu}. \tag{5.44}$$

However, this is equivalent to $\mathcal{O}_{7,3}$. To see this, we write,

$$\mathcal{O}_{7,3} = F^{\mu\nu}\nabla_\mu\phi\nabla_\nu(\mathcal{H}^\dagger\mathcal{H}) = \nabla_\nu\left(F^{\mu\nu}\nabla_\mu\phi(\mathcal{H}^\dagger\mathcal{H})\right) - \nabla_\nu F^{\mu\nu}\cdot\nabla_\mu\phi\cdot\mathcal{H}^\dagger\mathcal{H}. \tag{5.45}$$

On the late-time surface, the total derivative term scales as

$$\frac{1}{\eta_0^4}\times\eta_0^4\partial_\eta A_i^{\parallel}|\mathcal{H}|^2. \tag{5.46}$$

We cannot set both the $\mathcal{H}$ to their VEVs, otherwise we just go back to a dimension 5 operator considered above. Therefore, since at least one factor of the $\mathcal{H}$ fluctuation has to be present, going as $\eta_0^{3/2}$, the entire term scales at least as η_0. Thus the total derivative does not contribute. Finally we consider $\mathcal{O}_{7,4}$, which upon using the gauge field EOM becomes

$$\mathcal{O}_{7,4} = \frac{1}{g_A}F_{\mu\nu}\nabla^\mu\phi\nabla_\rho F^{\rho\nu}, \tag{5.47}$$

and can contribute to correlation functions. In total there are two non-redundant operators at dimension 7, neither of which have yet been considered in the literature.

5.3.4 Dimension 8

Many more operators are allowed at dimension 8, as enumerated in Table 5.3. The operators $\mathcal{O}_{8,1}-\mathcal{O}_{8,6}$ form a non-redundant minimal basis. Of the remaining operators, $\mathcal{O}_{8,7}$ simply vanishes. To see this, we fix μ and ν such that they are not identical. If they are identical, then the associated contribution is already a part of $\mathcal{O}_{8,2}$. Then we can write,

$$F_{\mu\rho}\tilde{F}^{\nu\rho} = \frac{1}{2}F_{\mu\rho}\epsilon^{\nu\rho\alpha\beta}F_{\alpha\beta}. \tag{5.48}$$

Operator	Expression
$\mathcal{O}_{8,1}$	$F_{\mu\nu}F^{\mu\nu}(\nabla_\rho\phi)^2$
$\mathcal{O}_{8,2}$	$F_{\mu\nu}\tilde{F}^{\mu\nu}(\nabla_\rho\phi)^2$
$\mathcal{O}_{8,3}$	$\lvert\mathcal{H}\rvert^4(\nabla_\mu\phi)^2$
$\mathcal{O}_{8,4}$	$\lvert D_\mu\mathcal{H}\rvert^2(\nabla_\nu\phi)^2$
$\mathcal{O}_{8,5}$	$(D^\mu\mathcal{H})^\dagger D^\nu\mathcal{H}\nabla_\mu\phi\nabla_\nu\phi$
$\mathcal{O}_{8,6}$	$F_{\mu\rho}F^{\nu\rho}\nabla^\mu\phi\nabla_\nu\phi$
$\mathcal{O}_{8,7}$	$F_{\mu\rho}\tilde{F}^{\nu\rho}\nabla^\mu\phi\nabla_\nu\phi$
$\mathcal{O}_{8,8}$	$\tilde{F}_{\mu\rho}\tilde{F}^{\nu\rho}\nabla^\mu\phi\nabla_\nu\phi$
$\mathcal{O}_{8,9}$	$(\mathcal{H}^\dagger D^\mu\mathcal{H} + (D^\mu\mathcal{H})^\dagger\mathcal{H})\nabla_\mu\nabla_\nu\phi\nabla^\nu\phi$
$\mathcal{O}_{8,10}$	$(-i)(\mathcal{H}^\dagger D^\mu\mathcal{H} - (D^\mu\mathcal{H})^\dagger\mathcal{H})\nabla_\mu\nabla_\nu\phi\nabla^\nu\phi$

Table 5.3: Allowed operators at dimension 8.

The index ρ has to be different from both μ and ν. We denote the two possible values it can take by γ and δ where $\gamma \neq \delta$. Then we can rewrite the above,

$$F_{\mu\rho}\tilde{F}^{\nu\rho} = \frac{1}{2}F_{\mu\gamma}\epsilon^{\nu\gamma\alpha\beta}F_{\alpha\beta} + \frac{1}{2}F_{\mu\delta}\epsilon^{\nu\delta\alpha\beta}F_{\alpha\beta} \quad [\gamma \text{ and } \delta \text{ are not summed over}]. \tag{5.49}$$

The indices α and β can then take values between μ, δ, and γ:

$$\begin{aligned} F_{\mu\rho}\tilde{F}^{\nu\rho} &= F_{\mu\gamma}\epsilon^{\nu\gamma\mu\delta}F_{\mu\delta} + F_{\mu\delta}\epsilon^{\nu\delta\mu\gamma}F_{\mu\gamma}, \quad [\text{no summation over any index}] \\ &= F_{\mu\gamma}F_{\mu\delta}(\epsilon^{\nu\gamma\mu\delta} + \epsilon^{\nu\delta\mu\gamma}) = 0. \end{aligned} \tag{5.50}$$

Therefore, $\mathcal{O}_{8,7}$ does not contribute.

The operator $\mathcal{O}_{8,8}$ is reducible. To see this we can write,

$$\tilde{F}_{\mu\rho}\tilde{F}^{\nu\rho} = \frac{1}{4}\epsilon_{\mu\rho\alpha\beta}\epsilon^{\nu\rho\gamma\delta}F^{\alpha\beta}F_{\gamma\delta} = \frac{1}{2}\left(F^{\gamma\delta}F_{\gamma\delta}\delta^\nu_\mu - 2F^{\nu\delta}F_{\mu\delta}\right). \tag{5.51}$$

This implies $\mathcal{O}_{8,8}$ is reducible to $\mathcal{O}_{8,1}$ and $\mathcal{O}_{8,6}$.

Next we focus on $\mathcal{O}_{8,9}$ which can be written as

$$\mathcal{O}_{8,9} = \nabla^\mu(\mathcal{H}^\dagger\mathcal{H})\nabla_\mu\nabla_\nu\phi\nabla^\nu\phi = -\frac{1}{2}\Box(\mathcal{H}^\dagger\mathcal{H})\nabla_\nu\phi\nabla^\nu\phi + \text{boundary term}. \tag{5.52}$$

where we have used IBP in the last step. Using methods similar to those used earlier, we can show the boundary term scales as $\eta_0^{1/2} \to 0$, and hence is not relevant for cosmological correlators. The surviving term, however, is equivalent to $\mathcal{O}_{8,4}$. To see this, we can use the EOM (5.21) for $\mathcal{H}$ and $\nabla_\mu A^\mu = 0$ to write,

$$\Box(\mathcal{H}^\dagger\mathcal{H}) = 2(D_\mu\mathcal{H})^\dagger D^\mu\mathcal{H} + 2V'(|\mathcal{H}|^2)|\mathcal{H}|^2. \tag{5.53}$$

The last term determined by the Higgs potential contributes to $\mathcal{O}_{6,1}$ and $\mathcal{O}_{8,3}$.

Finally, $\mathcal{O}_{8,10}$ can be reduced using similar techniques. First, we can rewrite it using the EOM (5.23) and an IBP,

$$\frac{1}{2g_A}\nabla_\rho F^{\rho\mu}\nabla_\mu\left(\nabla_\nu\phi\nabla^\nu\phi\right) = -\frac{1}{2g_A}\nabla_\mu\nabla_\rho F^{\rho\mu}\left(\nabla_\nu\phi\nabla^\nu\phi\right) + \text{boundary term}. \tag{5.54}$$

Similar as above, one can check that the boundary term vanishes as $\eta_0^{3/2}$ at late times. Noting that $\nabla_\mu\nabla_\rho F^{\rho\mu} \propto R_{\mu\rho}F^{\rho\mu} = 0$, we conclude that $\mathcal{O}_{8,10}$ does not contribute. Thus we find that there are six non-redundant operators at dimension 8.

5.3.5 Dimension 9

At dimension 9 we start with a set of operators summarized in Table 5.4. Among these, $\mathcal{O}_{9,1}$ does not contribute, as can be seen by doing an IBP which gives a vanishing boundary term, along with a term involving $\Box\phi$ which vanishes in the slow-roll limit. Using IBP, we can also check that $\mathcal{O}_{9,2}$ and $\mathcal{O}_{9,3}$ are equivalent. $\mathcal{O}_{9,17}$ and $\mathcal{O}_{9,20}$ are both reducible in terms of other operators. There are a priori other permutations with three derivatives acting on ϕ. Those terms can, however, be reduced by using the fact that for a maximally symmetric spacetime such as dS, we can write the Riemann tensor

Operator	Expression
$\mathcal{O}_{9,1}$	$\nabla_\mu\phi(\mathcal{H}^\dagger D^\mu\mathcal{H}+(D^\mu\mathcal{H})^\dagger\mathcal{H})\lvert\mathcal{H}\rvert^4=\nabla_\mu\phi\nabla^\mu(\mathcal{H}^\dagger\mathcal{H})\lvert\mathcal{H}\rvert^4$
$\mathcal{O}_{9,2}$	$\nabla_\mu\phi(-i)(\mathcal{H}^\dagger D^\mu\mathcal{H}-(D^\mu\mathcal{H})^\dagger\mathcal{H})\lvert\mathcal{H}\rvert^4=\frac{1}{g}\nabla_\mu\phi\nabla_\nu F^{\nu\mu}\lvert\mathcal{H}\rvert^4$
$\mathcal{O}_{9,3}$	$\nabla^\nu\phi(\mathcal{H}^\dagger D^\mu\mathcal{H}+(D^\mu\mathcal{H})^\dagger\mathcal{H})\lvert\mathcal{H}\rvert^2F_{\mu\nu}=\nabla^\nu\phi\nabla^\mu(\mathcal{H}^\dagger\mathcal{H})\lvert\mathcal{H}\rvert^2F_{\mu\nu}$
$\mathcal{O}_{9,4}$	$\nabla^\nu\phi(-i)(\mathcal{H}^\dagger D^\mu\mathcal{H}-(D^\mu\mathcal{H})^\dagger\mathcal{H})\lvert\mathcal{H}\rvert^2F_{\mu\nu}=\frac{1}{g}\nabla^\nu\phi\nabla_\alpha F^{\alpha\mu}\lvert\mathcal{H}\rvert^2F_{\mu\nu}$
$\mathcal{O}_{9,5}$	$\nabla_\nu\phi(\mathcal{H}^\dagger D^\mu\mathcal{H}+(D^\mu\mathcal{H})^\dagger\mathcal{H})F_{\mu\alpha}F^{\nu\alpha}=\nabla_\nu\phi\nabla^\mu(\mathcal{H}^\dagger\mathcal{H})F_{\mu\alpha}F^{\nu\alpha}$
$\mathcal{O}_{9,6}$	$\nabla_\mu\phi(\mathcal{H}^\dagger D^\mu\mathcal{H}+(D^\mu\mathcal{H})^\dagger\mathcal{H})F_{\alpha\nu}F^{\alpha\nu}=\nabla_\mu\phi\nabla^\mu(\mathcal{H}^\dagger\mathcal{H})F_{\alpha\nu}F^{\alpha\nu}$
$\mathcal{O}_{9,7}$	$\nabla_\mu\phi(\mathcal{H}^\dagger D^\mu\mathcal{H}+(D^\mu\mathcal{H})^\dagger\mathcal{H})F_{\alpha\nu}\tilde{F}^{\alpha\nu}=\nabla_\mu\phi\nabla^\mu(\mathcal{H}^\dagger\mathcal{H})F_{\alpha\nu}\tilde{F}^{\alpha\nu}$
$\mathcal{O}_{9,8}$	$\nabla_\nu\phi(-i)(\mathcal{H}^\dagger D^\mu\mathcal{H}-(D^\mu\mathcal{H})^\dagger\mathcal{H})F_{\mu\alpha}F^{\nu\alpha}=\frac{1}{g}\nabla_\nu\phi\nabla_\beta F^{\beta\mu}F_{\mu\alpha}F^{\nu\alpha}$
$\mathcal{O}_{9,9}$	$\nabla_\mu\phi(-i)(\mathcal{H}^\dagger D^\mu\mathcal{H}-(D^\mu\mathcal{H})^\dagger\mathcal{H})F_{\alpha\nu}F^{\alpha\nu}=\frac{1}{g}\nabla_\mu\phi\nabla_\beta F^{\beta\mu}F_{\alpha\nu}F^{\alpha\nu}$
$\mathcal{O}_{9,10}$	$\nabla_\mu\phi(-i)(\mathcal{H}^\dagger D^\mu\mathcal{H}-(D^\mu\mathcal{H})^\dagger\mathcal{H})F_{\alpha\nu}\tilde{F}^{\alpha\nu}=\frac{1}{g}\nabla_\mu\phi\nabla_\beta F^{\beta\mu}F_{\alpha\nu}\tilde{F}^{\alpha\nu}$
$\mathcal{O}_{9,11}$	$\nabla_\mu\phi(\mathcal{H}^\dagger D^\mu\mathcal{H}+(D^\mu\mathcal{H})^\dagger\mathcal{H})(\nabla_\nu\phi)^2=\nabla_\mu\phi\nabla^\mu(\mathcal{H}^\dagger\mathcal{H})(\nabla_\nu\phi)^2$
$\mathcal{O}_{9,12}$	$\nabla_\mu\phi(-i)(\mathcal{H}^\dagger D^\mu\mathcal{H}-(D^\mu\mathcal{H})^\dagger\mathcal{H})(\nabla_\nu\phi)^2=\frac{1}{g}\nabla_\mu\phi\nabla_\alpha F^{\alpha\mu}(\nabla_\nu\phi)^2$
$\mathcal{O}_{9,13}$	$\nabla^\nu\phi\nabla_\nu(\mathcal{H}^\dagger\mathcal{H})\lvert D_\mu\mathcal{H}\rvert^2$
$\mathcal{O}_{9,14}$	$\nabla_\mu\phi\nabla^\nu(\mathcal{H}^\dagger\mathcal{H})(D^\mu\mathcal{H})^\dagger D_\nu\mathcal{H}$
$\mathcal{O}_{9,15}$	$\nabla_\nu\phi\nabla_\alpha F^{\alpha\nu}\lvert D_\mu\mathcal{H}\rvert^2$
$\mathcal{O}_{9,16}$	$\nabla_\nu\phi\nabla_\alpha F^{\alpha\mu}(D^\nu\mathcal{H})^\dagger D_\mu\mathcal{H}$
$\mathcal{O}_{9,17}$	$\nabla_\nu\nabla_\mu\phi\nabla^\mu(\mathcal{H}^\dagger\mathcal{H})\nabla^\nu(\mathcal{H}^\dagger\mathcal{H})$
$\mathcal{O}_{9,18}$	$\nabla_\nu\nabla_\mu\phi\nabla_\alpha F^{\alpha\mu}\nabla_\beta F^{\beta\nu}$
$\mathcal{O}_{9,19}$	$\nabla_\nu\nabla_\mu\phi\nabla_\alpha F^{\alpha\mu}\nabla^\nu(\mathcal{H}^\dagger\mathcal{H})$
$\mathcal{O}_{9,20}$	$\nabla_\nu\nabla_\mu\phi\nabla^\mu F^{\rho\nu}\nabla_\rho(\mathcal{H}^\dagger\mathcal{H})$

Table 5.4: Allowed operators at dimension 9.

as $R_{\mu\nu\rho\sigma} \propto (g_{\mu\rho}g_{\nu\sigma} - g_{\mu\sigma}g_{\nu\rho})$. The Bianchi identity for $F_{\mu\nu}$ is also useful in reducing certain terms. All the other terms would contribute to cosmological correlators, albeit with suppressed contributions compared to operators at lower dimensions, as we will see in the next section.

However, the operator $\mathcal{O}_{9,12}$ is special since it gives rise to a quadratic mixing between the inflaton and the longitudinal gauge boson. Such an operator is observationally relevant since it would mediate tree-level NG, and we have seen that no other operator up to dimension 8 could give rise to such a mixing. To elaborate on this further, we can rewrite $\mathcal{O}_{9,12}$ after an IBP as, (dropping the gauge coupling)

$$\mathcal{O}_{9,12} = \nabla_\nu \left[(\nabla_\rho\phi\nabla^\rho\phi)\nabla_\mu\phi F^{\nu\mu}\right] - \nabla_\nu(\nabla_\rho\phi\nabla^\rho\phi)\nabla_\mu\phi F^{\nu\mu}. \tag{5.55}$$

Here we have used the fact that $F^{\nu\mu}\nabla_\mu\nabla_\nu\phi = 0$.

The Boundary Term. We first consider the boundary term, following an analysis similar to the above. We can write the boundary term as,

$$\frac{i}{g_A}\int \mathrm{d}^3x\sqrt{|\gamma|}n_\nu \left[(\nabla_\rho\phi\nabla^\rho\phi)\nabla_\mu\phi F^{\nu\mu}\right] \tag{5.56}$$

Therefore, at late times this term scales as,

$$\sim \frac{1}{\eta_0^3}\times\frac{1}{\eta_0}\times\eta_0^2\partial_i\phi\times\eta_0^4 F_{\eta i} \to 0. \tag{5.57}$$

Thus this term can be dropped.

The Bulk Term. To obtain the inflaton-gauge boson mixing, we can focus on the $F^{\eta i}$ component from (5.55). Since our main purpose to illustrate the form of the quadratic mixing, we will not track the overall numerical and η factors. If we set $\mu = i$ and $\nu = 0$, then the term would have a $\ddot{\phi}_0$, and hence it would be slow-roll suppressed. However, for

$\mu = 0$ and $\nu = i$, we would have a contribution which is quadratic in fluctuations,

$$\mathcal{O}_{9,12} \supset \dot{\phi}_0^2 \partial_i \dot{\phi} F^{i\eta}. \tag{5.58}$$

To simplify this further we can do a spatial IBP to write,

$$\mathcal{O}_{9,12} \supset \dot{\phi}_0^2 \dot{\phi} \partial_i F^{i\eta}. \tag{5.59}$$

and use (5.35) and (5.37) to write

$$\mathcal{O}_{9,12} \supset \dot{\phi}_0^2 m^2 \dot{\xi} A_\eta. \tag{5.60}$$

Here we have written the inhomogeneous part of the inflaton field: $\phi(t, \mathbf{x}) = \phi_0(t) + \xi(t, \mathbf{x})$. This matches with the conclusion in [295], however only this particular dimension 9 operator was considered in isolation.

5.3.6 Summary and Classification

The non-redundant operators up to dimension-9 are summarized in Table 5.5. In the third column, we indicate whether the operator contributes to the bispectrum at tree or loop level. This differs depending on the phase of the gauge theory. For example, $\mathcal{O}_{6,1}$ can clearly contribute to the bispectrum at tree level in the broken phase, when we can decompose the Higgs field in unitary gauge as $\mathcal{H} = (h + v)/\sqrt{2}$, with h an interacting degree of freedom and v the VEV. In the unbroken phase, this is not possible, and the operator may only contribute via a loop diagram.

On top of the tree/loop classification of each operator's contribution to observables, in weakly coupled UV completions it may also be possible to assign a tree/loop classification to the operator's Wilson coefficients [340]. At present the tree/loop classification of Wilson coefficients has only been extended through dimension-8 [341], and we do not explicitly classify Wilson coefficients here.

Dimension	Operator	Observables
5	$\mathcal{O}_{5,4} = \phi F_{\mu\nu}\tilde{F}^{\mu\nu}$	Loop [339]
6	$\mathcal{O}_{6,1} = (\nabla_\mu\phi)^2\mathcal{H}^\dagger\mathcal{H}$	Tree [295] and Loop [292]
7	$\mathcal{O}_{7,2} = \lvert\mathcal{H}\rvert^2\nabla_\mu\phi\nabla_\nu F^{\nu\mu}$	Loop
	$\mathcal{O}_{7,4} = F_{\mu\nu}\nabla^\mu\phi\nabla_\rho F^{\rho\nu}$	Loop
8	$\mathcal{O}_{8,1} = F_{\mu\nu}F^{\mu\nu}(\nabla_\rho\phi)^2$	Loop [292]
	$\mathcal{O}_{8,2} = F_{\mu\nu}\tilde{F}^{\mu\nu}(\nabla_\rho\phi)^2$	Loop
	$\mathcal{O}_{8,3} = \lvert\mathcal{H}\rvert^4(\nabla_\mu\phi)^2$	Tree and Loop
	$\mathcal{O}_{8,4} = \lvert D_\mu\mathcal{H}\rvert^2(\nabla_\nu\phi)^2$	Loop [292]
	$\mathcal{O}_{8,5} = (D^\mu\mathcal{H})^\dagger D^\nu\mathcal{H}\nabla_\mu\phi\nabla_\nu\phi$	Loop
	$\mathcal{O}_{8,6} = F_{\mu\rho}F^{\nu\rho}\nabla^\mu\phi\nabla_\nu\phi$	Loop
9	$\mathcal{O}_{9,2} = \lvert\mathcal{H}\rvert^2\mathcal{O}_{7,2}$	Loop
	$\mathcal{O}_{9,4} = \lvert\mathcal{H}\rvert^2\mathcal{O}_{7,4}$	Loop
	$\mathcal{O}_{9,5} = \nabla_\nu\phi\nabla^\mu(\mathcal{H}^\dagger\mathcal{H})F_{\mu\alpha}F^{\nu\alpha}$	Loop
	$\mathcal{O}_{9,6} = \mathcal{O}_{5,1}F_{\alpha\nu}F^{\alpha\nu}$	Loop
	$\mathcal{O}_{9,7} = \mathcal{O}_{5,1}F_{\alpha\nu}\tilde{F}^{\alpha\nu}$	Loop
	$\mathcal{O}_{9,8} = \nabla_\nu\phi\nabla_\beta F^{\beta\mu}F_{\mu\alpha}F^{\nu\alpha}$	Loop
	$\mathcal{O}_{9,9} = \mathcal{O}_{5,3}F_{\alpha\nu}F^{\alpha\nu}$	Loop
	$\mathcal{O}_{9,10} = \mathcal{O}_{5,3}F_{\alpha\nu}\tilde{F}^{\alpha\nu}$	Loop
	$\mathcal{O}_{9,11} = \mathcal{O}_{5,1}(\nabla_\mu\phi)^2$	Tree and Loop
	$\mathcal{O}_{9,12} = \mathcal{O}_{5,3}(\nabla_\mu\phi)^2$	Tree [295] and Loop
	$\mathcal{O}_{9,13} = \mathcal{O}_{5,1}\lvert D_\mu\mathcal{H}\rvert^2$	Loop
	$\mathcal{O}_{9,14} = \nabla_\mu\phi\nabla^\nu(\mathcal{H}^\dagger\mathcal{H})(D^\mu\mathcal{H})^\dagger D_\nu\mathcal{H}$	Loop
	$\mathcal{O}_{9,15} = \mathcal{O}_{5,3}\lvert D_\mu\mathcal{H}\rvert^2$	Loop
	$\mathcal{O}_{9,16} = \nabla_\nu\phi\nabla_\alpha F^{\alpha\mu}(D^\nu\mathcal{H})^\dagger D_\mu\mathcal{H}$	Loop
	$\mathcal{O}_{9,18} = \nabla_\nu\nabla_\mu\phi\nabla_\alpha F^{\alpha\mu}\nabla_\beta F^{\beta\nu}$	Loop
	$\mathcal{O}_{9,19} = \nabla_\nu\nabla_\mu\phi\nabla_\alpha F^{\alpha\mu}\nabla^\nu(\mathcal{H}^\dagger\mathcal{H})$	Loop

Table 5.5: A minimal operator basis up to dimension-9. We have dropped some overall prefactors to write the operators more compactly in terms of dimension-5 and -7 operators. Operators with leading effects (as described in Sec. 5.4) are highlighted in red. The third column indicates whether these operators contribute to the bispectrum at tree or loop level. Observables highlighted in blue only arise in the broken phase.

5.4 Observational Implications

Having constructed a minimal basis, we now study the predictions for NG, especially in the context of the cosmological collider. We first briefly review some aspects that will also set up the notation. As discussed in the Introduction, particles with masses of order $H_{\rm inf}$ can be produced as the Universe inflates. After production these particles can propagate on-shell, oscillating in time, and eventually decay into inflaton fluctuations. Such processes then give rise to non-trivial correlations among different inflaton fluctuations, in particular, three- and higher-point correlation functions. In this work, we will focus on the three-point function, i.e., the bispectrum, characterized by three spatial momenta $\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3$:

$$\langle \mathcal{R}(\mathbf{k}_1)\mathcal{R}(\mathbf{k}_2)\mathcal{R}(\mathbf{k}_3)\rangle = (2\pi)^3\delta(\mathbf{k}_1+\mathbf{k}_2+\mathbf{k}_3)B(k_1,k_2,k_3). \tag{5.61}$$

The δ function above enforces spatial momentum conservation. We have denoted the gauge invariant comoving curvature perturbation by $\mathcal{R}$; for a detailed definition and review see, e.g., [214]. Conventionally, the function B is normalized with respect to the power spectrum so that there is no overall scale dependence,

$$F(k_1,k_2,k_3) = \frac{5}{6}\frac{B(k_1,k_2,k_3)}{P_{\mathcal{R}}(k_1)P_{\mathcal{R}}(k_2)+P_{\mathcal{R}}(k_2)P_{\mathcal{R}}(k_3)+P_{\mathcal{R}}(k_1)P_{\mathcal{R}}(k_3)}. \tag{5.62}$$

It is also a convention to characterize the 'strength' of NG, at the equilateral configuration where $k_1 = k_2 = k_3$, in terms of a single number $f_{\rm NL} \equiv F(k,k,k)$. In the case of the cosmological collider, the function F exhibits oscillations as a function of k_3/k_1, especially in the squeezed limit $k_3 \ll k_1 \approx k_2$. We will denote the associated strength of NG by a parameter $f_{\rm osc}$, defined via:

$$F_{\rm sq} \approx \frac{5}{12}\frac{B(k_1,k_2,k_3)}{P_{\mathcal{R}}(k_1)P_{\mathcal{R}}(k_3)} \equiv |f_{\rm osc}|\left[\left(\frac{k_3}{k_1}\right)^{\frac{3}{2}+i\mu} + \text{c.c.}\right]. \tag{5.63}$$

In the following, we will estimate the parametric dependence of $f_{\rm osc}$ on various Wilson coefficients and identify which operators are expected to give a leading signal in a generic EFT. To that end, we briefly revisit the power counting scheme.

5.4.1 Power Counting

As discussed in Section 5.3, the power counting scheme for the gauge-Higgs-inflaton EFT can be organized in terms of operator dimension. For simplicity we take operators to be suppressed by appropriate powers of a common UV scale Λ with Wilson coefficients $c_{n,a}$,

$$\mathcal{L} \supset \sum_{n=5,\cdots} \frac{c_{n,a}}{\Lambda^{n-4}} \mathcal{O}_{n,a}. \tag{5.64}$$

Here n determines the dimension of the operator while the index a runs over all the operators having the same dimension. The EFT scale has to satisfy certain restrictions. To control the inflaton derivative expansion in $(\partial\phi)^2/\Lambda^4$, we require $\Lambda > \sqrt{\dot{\phi}_0}$ [342]. We also require $\Lambda > v$ to control the expansion in v^2/Λ^2, i.e., for the EFT to be suitably organized in terms of the linearly-realized gauge symmetry.

5.4.2 'Monochromatic' Operators

The operators summarized in Table 5.5 contribute to several types of cosmological correlators. A given diagram could involve either the Higgs or the gauge boson or both. Diagrams in which both the Higgs and the gauge boson are present can give rise to interesting signatures. For a recent study outlining the techniques for computing such diagrams involving more than one massive field, see [343]. However, the 'monochromatic' signatures that could let us extract the mass and spin of the boson in the most imme-

diate manner would involve either the Higgs or the gauge boson, but not both.[6] With this in mind, we now highlight which operators would give rise to such monochromatic signatures.

Note that a given monochromatic signature typically accumulates contributions from multiple operators at successive orders in power counting, where higher-order contributions are suppressed by appropriate powers of v/Λ or $H_{\rm inf}/\Lambda$. In what follows, the *leading* monochromatic signatures are those that arise at the lowest order in power counting. Leading operators are indicated in red in Table 5.5.

Higgs Signature

The leading monochromatic operator involving the Higgs arises at dimension 6, namely $\mathcal{O}_{6,1}$. In the broken phase, this operator gives rise to tree-level NG, as was studied in detail in [295]. On the other hand, in the unbroken phase, the operator contributes to NG at the loop level, as studied in [292].

At dimension 7 there are no monochromatic Higgs operators, since all the non-redundant operators involve gauge bosons as well. At dimension 8, monochromatic Higgs signatures arise from $\mathcal{O}_{8,3}, \mathcal{O}_{8,4}, \mathcal{O}_{8,5}$. However, we typically expect these contributions to be suppressed by additional powers of $(H_{\rm inf}/\Lambda)$ or (v/Λ) compared to the dimension-6 contribution. At dimension 9, $\mathcal{O}_{9,11}$, $\mathcal{O}_{9,13}$, $\mathcal{O}_{9,14}$ would contribute, with additional suppression by powers of $(H_{\rm inf}/\Lambda)$ and/or (v/Λ). In particular, $\mathcal{O}_{9,13}$ and $\mathcal{O}_{9,14}$ would contribute to vertices having at least three Higgs fluctuations.

[6] If there is a hierarchy between the Higgs and the gauge boson mass, we can integrate out the heavier particle to effectively obtain a monochromatic contribution for the other particle. Here we instead focus on the case where both the Higgs mass and the gauge boson mass are comparable.

Gauge Boson Signature

At loop-level, the leading monochromatic gauge boson signature arises from $\mathcal{O}_{5,4}$. This operator has been studied extensively in the context of axion inflation [328], and in the context of the cosmological collider [339]. The operator $\mathcal{O}_{7,4}$ can also contribute, but potentially without the chemical potential-like structure which arises from $\mathcal{O}_{5,4}$. There are multiple possible contributions at dimension 8. For example, the effects of $\mathcal{O}_{8,1}$ were computed in [292], while $\mathcal{O}_{8,2}$ and $\mathcal{O}_{8,6}$ would also contribute at the same order in EFT power counting.

At dimension 9, $\mathcal{O}_{9,12}$ is special since it can give rise to a quadratic mixing between the inflaton and the longitudinal gauge boson, as discussed above. The operator $\mathcal{O}_{9,18}$ also contributes, albeit suppressed by powers of $(H_{\rm inf}/\Lambda)$. There are other operators at dimension 9 that involve one inflaton with three gauge bosons, and therefore do not contribute to a three-point function at the one-loop level.

5.4.3 Estimates

The vertices relevant for NG mediated at tree level or at one loop are summarized in Figs. 5.1 and 5.2. There are four types of vertices, many of which accumulate contributions from more than one operator in Table 5.5, as indicated by the corresponding Wilson coefficients $c_{n,a}$. For operators involving derivatives, we have estimated the size of the derivatives to be of order $H_{\rm inf}$, as appropriate for scenarios where all the mass scales are of the order $H_{\rm inf}$.

Using these vertices we can construct various tree and loop-level diagrams that can mediate non-gaussianities. For illustration, we only consider the leading Higgs and gauge-boson mediated NG.

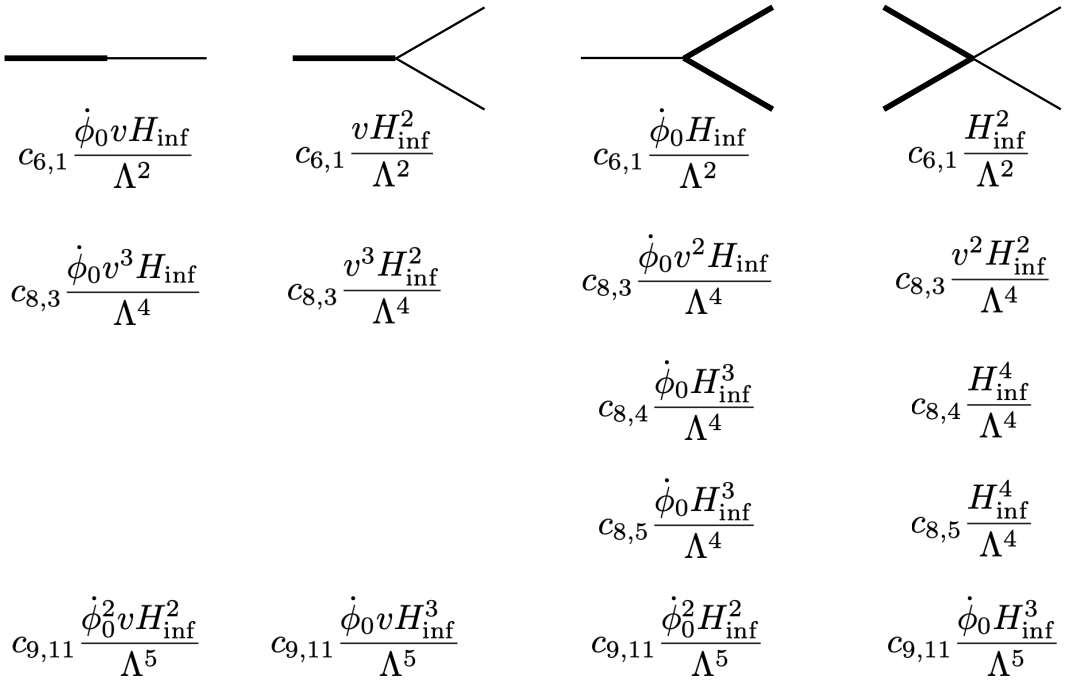

Figure 5.1: Higgs-inflaton vertices from various operators present in Table 5.5. The inflaton (Higgs) is denoted by a thin (solid) line. We have not included vertices with more than two Higgs fluctuations.

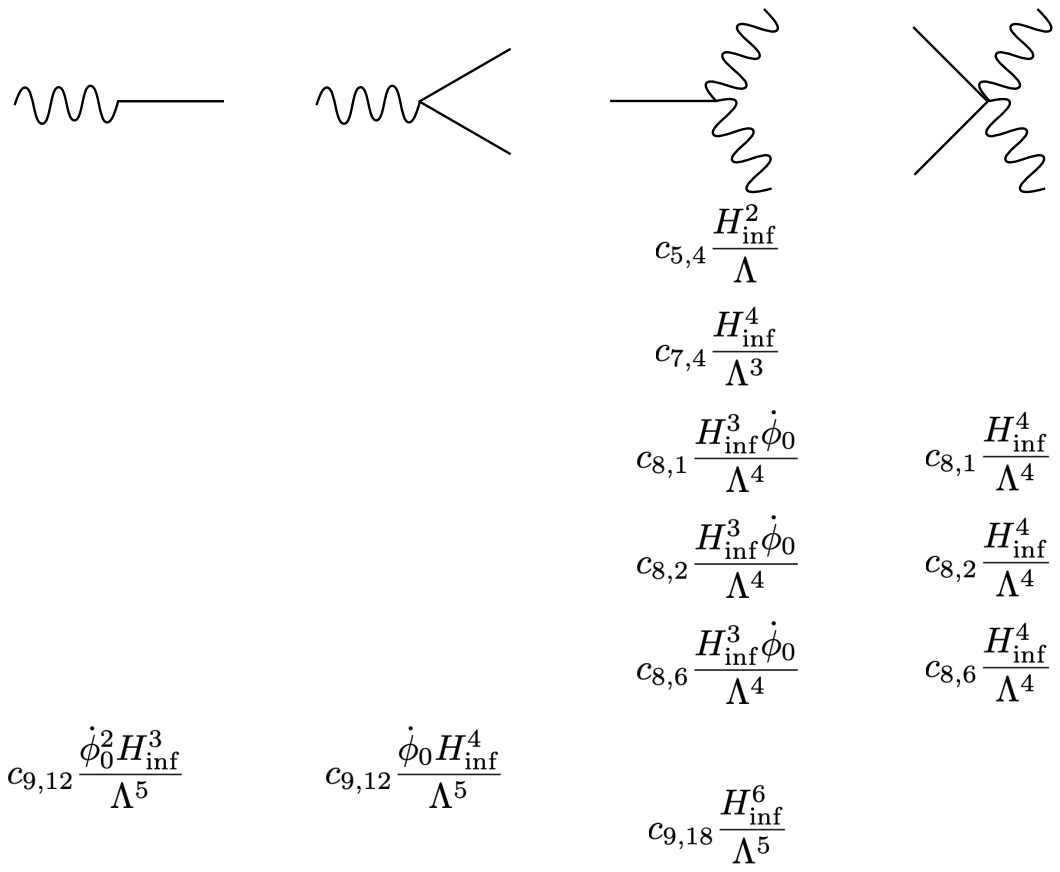

Figure 5.2: Gauge boson-inflaton vertices from various operators present in Table 5.5. The inflaton (gauge boson) is denoted by a straight (wavy) line. We have not included vertices with more than two gauge bosons.

Higgs. The signature in the broken phase was discussed in [295]; here we summarize the main conclusions. The dimension 6 operator $\mathcal{O}_{6,1}$ would mediate the leading tree-level NG through the vertices shown on the first row of Fig. 5.1. The parametric dependence of NG from the so-called 'single exchange' diagram can be estimated as,

$$f_{\rm osc,tree}^{\rm Higgs} \sim \frac{\mu^4 v^2}{\dot\phi_0^2 H_{\rm inf}^2}, \tag{5.65}$$

where $\mu^2 \sim c_{6,1}\dot\phi_0^2/\Lambda^2$ denotes a 'classical' correction to the Higgs mass from $\mathcal{O}_{6,1}$ when the inflaton is set to its VEV [295]. The total Higgs mass is then given by $m_{h,\rm full}^2 \sim \mu^2 + m_{h,\rm bare}^2$ where $m_{h,\rm bare}^2$ is the Higgs mass in the absence of any inflaton correction. Since cosmological collider signatures would be exponentially suppressed for $m_{h,\rm full} \gg H_{\rm inf}$, we require $m_{h,\rm full}^2 \sim H_{\rm inf}^2$. Barring any fine-tuning, this would mean $\mu^2 \sim H_{\rm inf}^2$ and $m_{h,\rm bare}^2 \sim H_{\rm inf}^2$. This implies the going rate for NG is $f_{\rm osc,tree}^{\rm Higgs} \sim (H_{\rm inf}^2 v^2)/\dot\phi_0^2$. For a more detailed numerical computation of the non-gaussianity, see Ref. [295] where the exponential fall off of the NG as a function of increasing Higgs mass is also computed.

In the unbroken phase, the NG signature arises at one loop, mediated by the right two vertices in the first row of Fig. 5.1. The corresponding estimate is

$$f_{\rm osc,loop}^{\rm Higgs} \sim \frac{1}{16\pi^2}\frac{\mu^4}{\dot\phi_0^2}. \tag{5.66}$$

For a detailed evaluation see Ref. [292].

Gauge Boson. We first discuss the tree-level signature in the broken phase, which was also discussed in [295]. For a tree-level signature, an essential ingredient is a quadratic mixing between an inflaton and the longitudinal mode of the gauge boson. Such a mixing arises at dimension 9, namely via $\mathcal{O}_{9,12}$, which also gives a cubic interaction between the gauge boson and the inflaton. These two vertices can contribute to NG via the so-called

single exchange diagram:

$$f_{\rm osc,tree}^{\rm gauge} \sim c_{9,12}^2 \frac{\dot{\phi}_0^4 H^2}{\Lambda^{10}}. \tag{5.67}$$

For $\Lambda \gtrsim \sqrt{\dot{\phi}_0}$, the above becomes $f_{\rm osc,tree}^{\rm gauge} \lesssim c_{9,12}^2 (H^2/\dot{\phi}_0)$. For a detailed evaluation see Ref. [295]. There is another diagram that can contribute to NG, involving the vertices with $c_{7,4}$ and $c_{9,12}$. The strength can be estimated as,

$$f_{\rm osc,tree}^{\rm gauge} \sim c_{7,4} c_{9,12}^2 \frac{\dot{\phi}_0^5}{\Lambda^{13}}. \tag{5.68}$$

Taking $\Lambda \gtrsim \sqrt{\dot{\phi}_0}$, the above estimate becomes $f_{\rm osc,tree}^{\rm gauge} \lesssim c_{7,4} c_{9,12}^2 (H/\dot{\phi}_0^{3/2})$, and therefore this contribution is expected to be subdominant compared to the previous process mediated purely via $c_{9,12}$. The dimension-5 operator determined by $c_{5,4}$ can give larger signals, both because it is a leading operator from an EFT perspective, and also because it can give a 'chemical potential' for gauge boson, potentially leading to exponential particle production. The cosmological collider signatures were computed in [339].

5.5 Conclusion

A systematic approach to constructing local EFTs entails not only fixing the power-counting and enumerating operators consistent with the infrared symmetries and fields, but also accommodating the resulting redundancy of description. This systematic approach is well-established in flat-space EFTs, where the irrelevance of boundary terms and invariance of S-matrix elements under field redefinitions make operator redundancies transparent. The situation is more complicated in cosmological EFTs, where boundary terms are not always negligible and the observables of interest are sensitive to field redefinitions. While minimal operator bases for inflaton self-interactions have been enumerated in various inflationary EFTs, much less progress has been made for EFTs of heavy particles coupled to the inflaton beyond the simplest cases at lowest order.

In this paper, we have developed a minimal operator basis for an abelian gauge-Higgs-inflaton EFT up to dimension 9, an archetypal example of a sector of heavy fields coupled to the inflaton relevant for cosmological collider physics. We have identified low-dimensional operators that are entirely redundant, as well as new non-redundant operators with potentially interesting observational signatures. Along the way, we have identified a number of useful methods for checking boundary terms arising from IBP relations, which can readily be applied to other EFTs of heavy particles coupled to the inflaton. The systematic enumeration of minimal operator bases in these EFTs is invaluable in light of the considerable interest in their cosmological signatures.

There are, of course, a number of interesting future directions. The methods presented in this paper may readily be generalized to other sectors coupled to the inflaton, including fermions and non-abelian gauge bosons. While we have focused on a Lorentz-preserving EFT of inflation with a shift-symmetric inflaton, similar methods may be applied in the more general Goldstone EFT of inflation [264]. Further, we have enumerated a number of operators with observational effects at loop-order. The precise computation of these effects remains important and will be the subject of future work. Finally, it would be very interesting to extend the general methods developed for operator bases in flat space [263] to cosmological contexts by accounting for the possible role of boundary terms. More broadly, we hope that the extensive attention devoted to operator bases in flat space EFTs may be equally applied to the plethora of EFTs arising in cosmological settings.

Chapter 6

Conclusions

This thesis has covered a wide range of topics in particle phenomenology and cosmology, and presented new results in building and analyzing parity solutions to solve the strong CP problem as well as progress in understanding the cosmological consequences of early universe particle models.

There is much further ground to pursue, both in extending these works and venturing further into the rich realm of particle theory and phenomenology. It's an exciting time to be in particle physics: While recent decades have seen a notable shift in morale following strings of null experimental results, we have no shortage of ideas to pursue BSM physics from both a theoretical and experimental lens. On the theory side, advancements in EFTs, scattering amplitudes, generalized symmetries, and cosmological correlators are poised to advance our knowledge of what QFT fundamentally is, while building new structures to simplify existing models and guide the formation of new ones. On the experimental side, we are still immersed in a golden age of cosmology, with upcoming gravitational wave experiments and probes of large-scale structure poised to return results on the tensor-to-scalar ratio, either a detection of primordial non-Gaussianities or new bound on f_{NL}, and signals of gravitational waves, which may or not be primordial in

origin. Whether these experiments detect new signatures or not, we will learn something about how our universe works. Perhaps this will be aligned with some of the ideas presented in this thesis, or perhaps the field requires a true paradigm shift to make progress. Either way, there is no shortage of ideas to ponder and work to be done. To conclude with one of my favorite quotes: "It's a magical world, Hobbes 'ol buddy...let's go exploring!"

Appendix A

Supplementary Material for Chapter 2

A.1 Mass eigenstates

A.1.1 Gauge and Higgs sectors

With the gauge group of Eq.(2.3), and the Higgs sector specified in table 2.1, spontaneous symmetry breaking takes place in two steps, as follows

$$SU(2)_L \times SU(2)_R \times U(1)_{\hat{Y}} \xrightarrow{v' \neq 0} SU(2)_L \times U(1)_Y \xrightarrow{v \neq 0} U(1)_{EM}. \tag{A.1}$$

The physical spectrum contains SM-like Z, $W^{\pm}$, and γ gauge bosons, as well as exotic Z' and $W'^{\pm}$ excitations. At tree-level, no mixing occurs in the charged gauge boson sector, and the mass eigenstates are given in terms of the gauge eigenbasis by the usual expression:

$$W^{\pm} = \frac{1}{\sqrt{2}}(W^1 \mp iW^2) \ , \tag{A.2}$$

and similarly in the W' sector. Tree-level masses are of the form $m_W = gv/2$ and $m_{W'} = gv'/2$, where we have assumed that $g' = g$, as mandated by generalized parity.

By contrast, in the neutral gauge boson sector mixing between SM and mirror fields takes place already at tree-level. At zeroth order in a v/v' expansion, the gauge eigenstates can be written in the mass eigenbasis as follows

$$\begin{pmatrix} W'^3_\mu \\ W^3_\mu \\ \hat{B}_\mu \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{\cos 2\theta_w}}{\cos\theta_w} & -\sin\theta_w \tan\theta_w & \sin\theta_w \\ 0 & \cos\theta_w & \sin\theta_w \\ -\tan\theta_w & -\tan\theta_w\sqrt{\cos 2\theta_w} & \sqrt{\cos 2\theta_w} \end{pmatrix} \begin{pmatrix} Z'_\mu \\ Z_\mu \\ A_\mu \end{pmatrix}, \tag{A.3}$$

where $\sin^2\theta_w \simeq 0.231$ as usual. Corrections to the above expression arise at $\mathcal{O}(v^2/v'^2)$. Masses for the SM-like Z and mirror Z' are given by

$$m_Z = \frac{gv}{2\cos\theta_w} + \mathcal{O}\left(\frac{v^2}{v'^2}\right), \qquad \text{and} \qquad m_{Z'} = \frac{gv'\cos\theta_w}{2\sqrt{\cos 2\theta_w}} + \mathcal{O}\left(\frac{v^2}{v'^2}\right). \tag{A.4}$$

After electroweak symmetry breaking, the Higgs sector consists of two real scalar fields, h and h', with masses given by $m_h \simeq 2\sqrt{\kappa}v$ and $m_{h'} \simeq \sqrt{2\lambda}v'$. Rotating from the gauge to the mass eigenbasis can be performed as follows

$$\begin{pmatrix} h \\ h' \end{pmatrix} \to \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} h \\ h' \end{pmatrix}, \tag{A.5}$$

with mixing angle $\alpha \sim v/v'$.

A.1.2 Fermion sector

Rotating from the flavor to the mass eigenbasis in the fermion sector requires solving the eigenvalue problem for the 6×6 matrices $\mathbb{M}_f^\dagger \mathbb{M}_f$, and $\mathbb{M}_f \mathbb{M}_f^\dagger$, with $\mathbb{M}_f$ as given in Eq.(2.12). This can be conveniently done as a perturbation expansion in $v/M, v'/M \ll 1$. In this section, we summarize the relevant results of this procedure. We focus first on the down-quark and lepton sectors (although we will use notation appropriate to the down-quark sector, we emphasize that identical results apply for leptons). The singularities of the up sector as related to the top quark merit a separate discussion that we present later.

Down-type quarks and leptons

The mass eigenvalues in the down-quark sector can be found by diagonalizing the two 3×3 matrices

$$\frac{v'v}{2} y_d'^* \mathcal{M}_d^{-1} y_d^T, \qquad \text{and} \qquad \mathcal{M}_d. \tag{A.6}$$

In full generality, i.e. without yet imposing generalized parity, the above matrices are not necessarily hermitian, and two unitary matrices are needed in order to bring them into real diagonal form. This corresponds to the unitary transformations

$$d \to \mathcal{O}_d^\dagger d, \; d' \to \mathcal{O}_{d'}^\dagger d', \qquad \text{and} \qquad D \to \mathcal{O}_D^\dagger D, \; D' \to \mathcal{O}_{D'}^\dagger D'. \tag{A.7}$$

By definition, the rotation matrices are such that

$$\mathcal{Y}_d \equiv \mathcal{O}_{d'}^* \left(\frac{v'v}{2} y_d'^* \mathcal{M}_d^{-1} y_d^T \right) \mathcal{O}_d^\dagger = \text{diag}(m_{d_i}) \; , \tag{A.8}$$

and

$$\mathcal{Y}_D \equiv \mathcal{O}_D^* \mathcal{M}_d \mathcal{O}_{D'}^\dagger = \text{diag}(m_{D_i}) \; , \tag{A.9}$$

where m_{d_i} and m_{D_i} are the masses of the SM and exotic heavy quarks respectively. As advertised in section 2.2.3, we will make the simplifying assumption that all three mirror quarks appear at a common scale $m_{D_i} \sim M \gg v, v'$. Imposing generalized parity makes both matrices in Eq.(A.6) hermitian. In this case, a single unitary matrix suffices to make them diagonal, and we have $\mathcal{O}_{d'} = \mathcal{O}_d^*$ and $\mathcal{O}_D = \mathcal{O}_{D'}^*$.

It is convenient to define two new matrices corresponding to the Yukawa couplings in this new basis

$$\tilde{y}_d \equiv \mathcal{O}_d^* y_d \mathcal{O}_D^\dagger, \qquad \text{and} \qquad \tilde{y}_d' \equiv \mathcal{O}_{d'} y_d' \mathcal{O}_{D'}^T \; . \tag{A.10}$$

With this definition, the tree-level masses of the SM-like fermions read

$$m_{d_i} = \frac{vv'}{2} \sum_j \frac{(\tilde{y}_d'^*)_{ij} (\tilde{y}_d)_{ij}}{m_{D_j}} = \frac{vv'}{2} \sum_j \frac{|(\tilde{y}_d)_{ij}|^2}{m_{D_j}} \; , \tag{A.11}$$

where the last step holds provided we impose generalized parity. From this expression, we can find an upper bound on the individual entries in the Yukawa matrix, of the form

$$|(\tilde{y}_d)_{ij}| \lesssim \left(\frac{2m_{d_i}m_{D_j}}{vv'}\right)^{1/2} \sim \left(\frac{m_{d_i}M}{vv'}\right)^{1/2} . \tag{A.12}$$

As advertised in section 2.2.3, bringing the full 6×6 matrix of Eq.(2.12) into diagonal form requires a further transformation that mixes the $SU(2)$-doublet and singlet fields, as specified in Eq.(2.20). In terms of the $\tilde{y}_d$ and $\tilde{y}'_d$ couplings defined earlier, the 3×3 blocks appearing in Eq.(2.20) can be written as

$$\epsilon_d = \frac{v}{\sqrt{2}}\hat{\gtrsim}_D^{-1}\tilde{y}_d^T, \qquad \text{and} \qquad \epsilon'_d = \frac{v'}{\sqrt{2}}\hat{\gtrsim}_D^{-1}\tilde{y}_d'^{\dagger}, \tag{A.13}$$

whose entries are of $\mathcal{O}(v/M)$ and $\mathcal{O}(v'/M)$ respectively. When generalized parity is only broken by the different vev's in the SM and mirror sectors, we have $\epsilon'_d = (v'/v)\epsilon_d^*$.

Up-type quarks

The diagonalization procedure in the up-quark sector is analogous to that for down-type quarks and leptons, although this time accommodating for the singularities of the third generation for which the see-saw mechanism cannot be implemented.

As before, at zeroth order in $v^{(\prime)}/M$, we perform transformations of the form

$$u \to \mathcal{O}_u^\dagger u,\ u' \to \mathcal{O}_{u'}^\dagger u', \qquad \text{and} \qquad U \to \mathcal{O}_U^\dagger U,\ U' \to \mathcal{O}_{U'}^\dagger U'. \tag{A.14}$$

On the one hand, the matrices $\mathcal{O}_U$ and $\mathcal{O}_{U'}$ must be chosen such that the vector-like mass matrix $\mathcal{M}_u$ is brought into diagonal form. In this case, we make the assumption that two of the eigenvalues of $\mathcal{M}_u$ are $m_{U_1}, m_{U_2} \sim M$, whereas the third one is much smaller, and for simplicity we will take it to vanish in what follows. On the other hand, the matrices $\mathcal{O}_u$ and $\mathcal{O}_{u'}$ must now be such that

$$(\tilde{y}_u'^*\hat{\gtrsim}_U^{-1}\tilde{y}_u^T)_{ij} = \delta_{ij}(\tilde{y}_u'^*\hat{\gtrsim}_U^{-1}\tilde{y}_u^T)_{ii}, \qquad \text{and} \qquad (\tilde{y}_u)_{i3} = (\tilde{y}'_u)_{i3} = 0 \tag{A.15}$$

for $i,j=1,2$, and where $\hat{\mathbb{m}}_U^{-1} \equiv \mathrm{diag}(m_{U_1}^{-1}, m_{U_2}^{-1}, 0)$, and the $\tilde{y}_u$ and $\tilde{y}'_u$ matrices are defined as in Eq.(A.10). Moreover, we define $y_t \equiv \tilde{y}_{33}$, and $y_{t'} \equiv \tilde{y}'^*_{33}$, which we may choose to be real and positive.

With this preliminaries, the tree-level mass eigenvalues in the top sector read

$$m_t = \frac{y_t}{\sqrt{2}} v \ , \qquad \text{and} \qquad m_{t'} = \frac{y_{t'}}{\sqrt{2}} v', \tag{A.16}$$

with $y_{t'} = y_t$ if we impose generalized parity. For the first and second generation, we have instead

$$m_{u_i} = \frac{vv'}{2} \sum_{j=1}^{2} \frac{(\tilde{y}'^*_u)_{ij}(\tilde{y}_u)_{ij}}{m_{U_j}} = \frac{vv'}{2} \sum_{j=1}^{2} \frac{|(\tilde{y}_u)_{ij}|^2}{m_{U_j}}, \tag{A.17}$$

where the last step holds provided we impose generalized parity. As before, we can now obtain an upper bound on the individual Yukawa entries, of the form

$$|(\tilde{y}_u)_{ij}| \lesssim \left(\frac{2 m_{u_i} m_{U_j}}{vv'} \right)^{1/2} \sim \left(\frac{m_{u_i} M}{vv'} \right)^{1/2} \qquad \text{for} \qquad i,j=1,2. \tag{A.18}$$

A further transformation mixing the $SU(2)$-singlet and doublet components is again necessary in order to diagonalize the full 6×6 mass matrix, which can be written as in Eq.(2.20). The corresponding ϵ_u and ϵ'_u blocks are now given by

$$\begin{aligned} (\epsilon_u)_{ij} &= \frac{v}{\sqrt{2}} (\hat{\mathbb{m}}_U^{-1} \tilde{y}^T)_{ij} \ , & (\epsilon_u)_{3j} &= -\frac{vv'}{2} \frac{m_{t'} (\tilde{y}'^* \hat{\mathbb{m}}_U^{-1} \tilde{y}^T)_{3j}}{m_{t'}^2 - \delta_{j3} m_t^2} \ , \\ (\epsilon'_u)_{ij} &= \frac{v'}{\sqrt{2}} (\hat{\mathbb{m}}_U^{-1} \tilde{y}'^T)^*_{ij} \ , & (\epsilon'_u)_{3j} &= -\frac{vv'}{2} \frac{m_t (\tilde{y}^* \hat{\mathbb{m}}_U^{-1} \tilde{y}'^T)^*_{3j}}{m_t^2 - \delta_{j3} m_{t'}^2} \ . \end{aligned} \tag{A.19}$$

for $i=1,2$ and $j=1,2,3$. Just as in the down-quark sector, if generalized parity is only broken by the difference between v and v', we have $\epsilon'_u = (v'/v)\epsilon^*_u$.

A.2 Radiatively induced EDM

A.2.1 One-loop EDM

We will now present a calculation of the one-loop correction to the EDM of elementary charged fermions that arises under the assumption that parity is only broken softly, both

in the scalar potential and through the presence of non-hermitian vector-like masses for the $SU(2)$-singlets. The relevant diagrams are those featured in figure 2.3. We will concentrate first on diagrams where either h or h' propagate inside the loop.

In full generality, a Dirac fermion f that interacts with another fermion ψ, and a neutral scalar ϕ through Yukawa couplings of the form

$$\mathcal{L} \supset L(\bar{f}_R \psi_L)\phi + R(\bar{f}_L \psi_R)\phi + \text{h.c.}, \tag{A.20}$$

will receive a one-loop EDM given by

$$\frac{d_f}{e} = \frac{Q_\psi}{16\pi^2}\frac{m_\psi}{m_\phi^2} A(r)\,\text{Im}\left(LR^*\right), \tag{A.21}$$

where $r = m_\psi^2/m_\phi^2$, and the loop function A is given by

$$A(r) = \frac{1}{2(1-r)^2}\left(3 - r + \frac{2\log r}{1-r}\right). \tag{A.22}$$

In the model we are considering, the Yukawa interactions involving both light and heavy fermions can be written as

$$\mathcal{L} \supset -\sum_{s=h,h'} s\left(\bar{d}_R \hat{\eta}^s d_L + \bar{d}_R \hat{\beta}^s D_L + \bar{D}_R \hat{\gamma}^s d_L + \bar{D}_R \hat{\delta}^s D_L\right) + \text{h.c.}, \tag{A.23}$$

where we are using notation specific to the down-quark sector, but analogous expressions apply for up-quarks and leptons (although the specific form of the Yukawa matrices will differ). It will be convenient to write the above matrices as $\hat{\omega}^s = R_{1s}\omega^h + R_{2s}\omega^{h'}$ (for $\omega = \eta, \beta, \gamma, \delta$), where $R_{11} = R_{22} = \cos\alpha$ and $R_{12} = -R_{21} = \sin\alpha$, with $\alpha \sim v/v'$ the mixing angle the Higgs sector. For the down-quark sector, the ω matrices can be conveniently written as follows

$$\eta^h = -\frac{\mathbb{Y}_d}{v}, \qquad \beta^h = -\frac{1}{\sqrt{2}}\left(\mathbb{Y}_d \tilde{y}^* \mathbb{Y}_D^{-1}\right), \qquad \gamma^h = \frac{\tilde{y}^T}{\sqrt{2}}, \qquad \delta^h = \frac{v}{2}\left(\tilde{y}^T \tilde{y}^* \mathbb{Y}_D^{-1}\right), \tag{A.24}$$

and

$$\eta^{h'} = -\frac{\mathbb{Y}_d}{v'}, \qquad \beta^{h'} = \frac{\tilde{y}'^*}{\sqrt{2}}, \qquad \gamma^{h'} = -\frac{1}{\sqrt{2}}\mathbb{Y}_D^{-1}\tilde{y}'^T\mathbb{Y}_d, \qquad \delta^{h'} = \frac{v'}{2}\left(\mathbb{Y}_D^{-1}\tilde{y}'^T\tilde{y}'^*\right). \tag{A.25}$$

The one-loop correction to the EDM of one of the SM-like quarks, d_i, is dominated by diagrams where the heavy mirror quarks propagate inside the loop. Since $m_{D_j} \sim M \gg m_h, m_{h'}$, we can expand the loop function $A(r)$ in the limit $r \gg 1$. Keeping the first two terms, we find

$$\frac{d_{d_i}}{e} \simeq -\sum_{j,s} \frac{Q_d}{32\pi^2 m_{D_j}} \left(1 - \frac{m_s^2}{m_{D_j}^2}\right) \mathrm{Im}\left(\hat{\beta}_{ij}^s \hat{\gamma}_{ji}^s\right). \tag{A.26}$$

The last factor in the previous expression can be written as

$$\mathrm{Im}(\hat{\beta}_{ij}^s \hat{\gamma}_{ji}^s) = R_{1s} R_{2s}\, \mathrm{Im}(\beta_{ij}^h \gamma_{ji}^{h'} + \beta_{ij}^{h'} \gamma_{ji}^h), \tag{A.27}$$

where we have taken into account that $\beta_{ij}^h \gamma_{ji}^h, \beta_{ij}^{h'} \gamma_{ji}^{h'} \in \mathbb{R}$, so those combinations don't appear on the right-hand-side. When summing over s in Eq.(A.26), the contribution from the leading term in the $m_s^2/m_{D_j}^2 \ll 1$ expansion vanishes since $\sum_s R_{1s} R_{2s} = 0$. The leading contribution to d_{d_i} then reads

$$\begin{aligned} \frac{d_{d_i}}{e} &\simeq \sum_{j,s} \frac{Q_d}{32\pi^2} \frac{m_s^2}{m_{D_j}^3} R_{1s} R_{2s}\, \mathrm{Im}\left(\beta_{ij}^h \gamma_{ji}^{h'} + \beta_{ij}^{h'} \gamma_{ji}^h\right) \\ &\simeq \sum_j \frac{Q_d}{32\pi^2} \frac{m_{h'}^2}{m_{D_j}^3} \sin\alpha\, \mathrm{Im}\left(\beta_{ij}^{h'} \gamma_{ji}^h\right), \end{aligned} \tag{A.28}$$

where in the last step we have neglected the first term in parenthesis since it is suppressed by a factor of $\mathcal{O}((vv'/M^2)^2)$ with respect to the second, and we have only kept the contribution from h', since the contribution from h is suppressed by an additional factor of $m_h^2/m_{h'}^2$.

From Eq.(A.24) and (A.25), we find

$$\beta_{ij}^{h'} \gamma_{ji}^h = \frac{1}{2} \tilde{y}_{ij}'^* \tilde{y}_{ij}. \tag{A.29}$$

When generalized parity is a good symmetry, we have $\tilde{y}' = \tilde{y}$, and therefore the above term is real, in turn leading to a vanishing EDM. In the presence of soft breaking through non-hermitian vector-like masses, the relationship $\tilde{y}' = \tilde{y}$ no longer holds, even if the

Yukawa couplings in the flavor basis remain identical since the breaking is soft. To see that the equality of the Yukawa couplings in the SM and mirror sectors no longer holds in the mass basis, it is useful to remind ourselves of the fermion mass diagonalization procedure discussed in appendix A.1.2. When generalized parity remains unbroken, the unitary matrices of Eq.(A.7) are such that $\mathcal{O}_F = \mathcal{O}^*_{F'}$ and $\mathcal{O}_{f'} = \mathcal{O}^*_f$. However, the non-hermiticity of the vector-like mass matrix means the unitary matrices needed to bring the 3×3 matrices of Eq.(A.6) into real diagonal form will no longer satisfy this simple relation. Instead, writing the new vector-like mass matrices as $\mathcal{M}_f + i\Delta\mathcal{M}_f$, with both $\mathcal{M}_f$ and $\Delta\mathcal{M}_f$ hermitian, the new unitary matrices are modified as follows

$$\begin{aligned} \mathcal{O}_{F'} &\to \tilde{\mathcal{O}}_{F'} = \mathcal{O}_{F'} + \Delta_{F'}, & \mathcal{O}_F &\to \tilde{\mathcal{O}}_F = \mathcal{O}^*_{F'} - \Delta^*_{F'}, \\ \mathcal{O}_f &\to \tilde{\mathcal{O}}_f = \mathcal{O}_f + \Delta_f, & \mathcal{O}_{f'} &\to \tilde{\mathcal{O}}_{f'} = \mathcal{O}^*_f - \Delta^*_f. \end{aligned} \tag{A.30}$$

The Δ matrices arise at $\mathcal{O}\left(|\Delta\mathcal{M}|/M\right)$, and are given by

$$(\Delta_{F'})_{ij} = i \sum_{k \neq i} \frac{[\Delta\tilde{\mathcal{M}}, \mathcal{Y}_F]_{ik}}{m^2_{F_i} - m^2_{F_k}} \left(\mathcal{O}_{F'}\right)_{kj}, \quad \text{and} \quad (\Delta_f)_{ij} = i \sum_{k \neq i} \frac{[\Delta\tilde{m}, \mathcal{Y}_f]_{ik}}{m^2_{f_i} - m^2_{f_k}} \left(\mathcal{O}_f\right)_{kj}, \tag{A.31}$$

where

$$\Delta\tilde{\mathcal{M}} \equiv \mathcal{O}_{F'} \Delta\mathcal{M} \mathcal{O}^\dagger_{F'} \qquad \text{and} \qquad \Delta\tilde{m} \equiv \frac{vv'}{2} \tilde{y}^* \mathcal{Y}_F^{-1} \Delta\tilde{\mathcal{M}} \mathcal{Y}_F^{-1} \tilde{y}^T. \tag{A.32}$$

Using the above expressions in the definition of $\tilde{y}$, it is possible to write

$$\tilde{y}'^* = \tilde{y}^* (\mathbb{1} + \xi), \tag{A.33}$$

where ξ is a matrix with entries of $\mathcal{O}(|\Delta\mathcal{M}|/M)$. Explicitly, after some massaging,

$$\tilde{y}'^*_{ij} = \sum_l \tilde{y}^*_{il} \left(\delta_{lj} + i \frac{\Delta\tilde{\mathcal{M}}_{lj}}{m_{F_l} + m_{F_j}} (1 - \delta_{lj}) + i \frac{vv'}{2} \sum_{n,k \neq i} \frac{\tilde{y}^*_{kj} \tilde{y}_{kn}}{m_{f_i} + m_{f_k}} \frac{\Delta\tilde{\mathcal{M}}_{ln}}{m_{F_l} m_{F_n}} \right). \tag{A.34}$$

In total:

$$\mathrm{Im}(\tilde{y}'^*_{ij} \tilde{y}_{ij}) = |\tilde{y}_{ij}|^2 \times \mathcal{O}\left(\frac{\Delta\mathcal{M}}{M}\right). \tag{A.35}$$

Plugging this back into Eq.(A.28), we have

$$\frac{d_{d_i}}{e} \simeq \sum_j \frac{Q_d}{32\pi^2}\frac{m_{h'}^2}{m_{D_j}^3}\sin\alpha\frac{1}{2}\,\mathrm{Im}(\tilde{y}'^*_{ij}\tilde{y}_{ij}) \simeq \frac{n_d Q_d}{32\pi^2}\frac{m_{d_i}}{M^2}\times\mathcal{O}\left(\frac{\Delta\mathcal{M}}{M}\right), \tag{A.36}$$

with $n_d = 3$ the number of mirror fermions appearing at the see-saw scale in the down-quark sector. The above expression also applies to the lepton sector, after making the obvious substitutions. In the up-quark sector, the expressions for the Yukawa couplings are somewhat different to those in Eq.(A.24) and (A.25), but can be similarly found by following the flavor-to-mass-basis rotation procedure outlined in section A.1.2. In the end, diagrams where h' and the mirror partners of the u and c quarks propagate inside the loop give the leading contribution to the one-loop EDM. Thus, the above expression also applies for the up-quark sector, this time with $n_u = 2$ instead.

Additional contributions arise from diagrams where Z and Z' propagate inside the loop (see figure 2.3). In this case, the leading contribution arises from diagrams involving Z' as well as heavy mirror fermions. In total, the final result is parametrically the same as that in Eq.(A.36), except for an additional suppression by a factor of $g^2\sin^2\theta_w$.

Although the potential one-loop correction to $\bar{\theta}$ that could arise as a result of the soft breaking through non-hermitian vector-like masses was already shown to vanish in [22], this can also be seen from the calculation we have just performed. The relevant diagrams contributing to the quark mass matrix, and therefore to $\bar{\theta}$, are those of figure 2.3, minus the external photon line. So although the appropriate loop function will be different, the overall correction will be similarly proportional to Eq.(A.29). Using Eq.(A.8) to rewrite $\tilde{y}'$ in terms of $\tilde{y}$, and the diagonal mass matrices, we find

$$\mathrm{Im}(\delta m_{d_i}) \propto \mathrm{Im}(\tilde{y}'^*_{ij}\tilde{y}_{ij}) \propto m_{d_i}\,\mathrm{Im}((\tilde{y}^{-1})_{ji}\tilde{y}_{ij}). \tag{A.37}$$

As a result, the corresponding contribution to $\bar{\theta}$ from the down-quark sector reads

$$\sum_i \frac{\mathrm{Im}(\delta m_{d_i})}{m_{d_i}} \propto \mathrm{Im}\left(\sum_i (\tilde{y}^{-1})_{ji}\tilde{y}_{ij}\right) = 0. \tag{A.38}$$

Notice the sum over quark flavors is crucial in the above cancellation.

A.2.2 One-loop $\bar{\theta}$

The calculation of the one-loop correction to the quark mass matrix, and, in turn, to $\bar{\theta}$, proceeds along similar lines to the EDM calculation we have just discussed. The leading contribution to $\bar{\theta}$ comes from corrections to the light quark masses, and it is due to diagrams where either h' or ϕ propagate inside the loop.

Rotating from the gauge to the mass basis in the scalar sector requires performing a transformation $s_i \to R_{ij} s_j$, with $s_i = \{h, h', \phi\}$, and R is a 3×3 orthogonal matrix that we parametrize in terms of the various mixing angles as

$$R = \begin{pmatrix} \mathrm{c}_\alpha \mathrm{c}_\beta & \mathrm{c}_\alpha \mathrm{s}_\beta \mathrm{s}_\gamma + \mathrm{s}_\alpha \mathrm{c}_\gamma & -\mathrm{c}_\alpha \mathrm{s}_\beta \mathrm{c}_\gamma + \mathrm{s}_\alpha \mathrm{s}_\gamma \\ -\mathrm{s}_\alpha \mathrm{c}_\beta & -\mathrm{s}_\alpha \mathrm{s}_\beta \mathrm{s}_\gamma + \mathrm{c}_\alpha \mathrm{c}_\gamma & \mathrm{s}_\alpha \mathrm{s}_\beta \mathrm{c}_\gamma + \mathrm{c}_\alpha \mathrm{s}_\gamma \\ \mathrm{s}_\beta & -\mathrm{c}_\beta \mathrm{s}_\gamma & \mathrm{c}_\beta \mathrm{c}_\gamma \end{pmatrix} \tag{A.39}$$

where $\mathrm{c}_\alpha = \cos\alpha$, $\mathrm{s}_\alpha = \sin\alpha$, etc. Parametrically, we expect $\mathrm{c}_\alpha \sim \mathrm{c}_\beta \sim \mathrm{c}_\gamma = \mathcal{O}(1)$, whereas $\mathrm{s}_\alpha \sim v/v'$, $\mathrm{s}_\beta \sim v/v_\phi$, and $\mathrm{s}_\gamma \sim v_\phi/v'$. In the down-quark sector, the Yukawa interactions of Eq.(A.23) need to be extended to include ϕ in the sum, and the $\hat{\omega}^s$ matrices are now given by

$$\hat{\omega}^s = R_{1s}\omega^h + R_{2s}\omega^{h'} + R_{3s}\omega^\phi \qquad \text{for} \qquad \omega = \eta, \beta, \gamma, \delta. \tag{A.40}$$

The expressions for ω^h and $\omega^{h'}$ are as in Eq.(A.24) and (A.25), whereas for ϕ we have

$$\begin{aligned} \eta^\phi &= -i\frac{vv'}{2}\left(\tilde{y}'^* \tilde{\succ}_D^{-1} \tilde{\tilde{y}} \tilde{\succ}_D^{-1} \tilde{y}^T\right), & \beta^\phi &= \frac{iv'}{\sqrt{2}}\left(\tilde{y}'^* \tilde{\succ}_D^{-1} \tilde{\tilde{y}}\right), \\ \gamma^\phi &= \frac{iv}{\sqrt{2}}\left(\tilde{\tilde{y}} \tilde{\succ}_D^{-1} \tilde{y}^T\right), & \delta_\phi &= -i\tilde{\tilde{y}}, \end{aligned} \tag{A.41}$$

where $\tilde{\tilde{y}} \equiv \mathcal{O}_{F'} \bar{y} \mathcal{O}_{F'}^\dagger$, as usual.

In the notation of Eq.(A.20), the one-loop correction to $\mathrm{Im}(\delta m_f)$ is given by

$$\mathrm{Im}(\delta m_f) = \frac{m_\psi}{16\pi^2} \mathcal{F}(m_\psi, m_\phi)\, \mathrm{Im}(LR^*), \tag{A.42}$$

where the loop function $\mathcal{F}$ now reads

$$\mathcal{F}(m_\psi, m_\phi) = \frac{1}{m_\psi^2 - m_\phi^2}\left[m_\psi^2\left(\log\frac{m_\psi^2}{\mu^2} - 1\right) - m_\phi^2\left(\log\frac{m_\phi^2}{\mu^2} - 1\right)\right]. \tag{A.43}$$

For the case at hand, the leading one-loop correction to the mass of the SM-like fermions involves diagrams where the heavy mirror partners appearing at scale M propagate inside the loop. Specifically, in the down-quark sector, we have

$$\mathrm{Im}(\delta m_{d_i}) = \sum_{s,j}\frac{m_{D_j}}{16\pi^2}\mathcal{F}(m_{D_j}, m_s)\,\mathrm{Im}(\hat{\beta}^s_{ij}\hat{\gamma}^s_{ji}). \tag{A.44}$$

In analogy to the discussion in the previous section, the leading term in $\mathcal{F}$ in the limit $m_{D_j} \gg m_s$ is independent of m_s, and its contribution to $\mathrm{Im}(\delta m_f)$ vanishes as a result of the orthogonality of the mixing matrix in the scalar sector. The leading correction to $\mathrm{Im}(\delta m_{d_i})$ then reads

$$\begin{aligned}\mathrm{Im}(\delta m_{d_i}) &\simeq \sum_{j,s}\frac{m_s^2}{16\pi^2 m_{D_j}}\log\frac{m_{D_j}^2}{m_s^2}\,\mathrm{Im}(\hat{\beta}^s_{ij}\hat{\gamma}^s_{ji}) \\ &\simeq \sum_{j,s=h',\phi}\frac{m_s^2}{16\pi^2 m_{D_j}}\log\frac{m_{D_j}^2}{m_s^2}\,\mathrm{Im}(\mathrm{s}_\gamma\beta^{h'}_{ij}\gamma^\phi_{ji} + \mathrm{s}_\gamma\mathrm{s}_\alpha\beta^\phi_{ij}\gamma^h_{ji}),\end{aligned} \tag{A.45}$$

where in the last step we have neglected the contribution from h, which is suppressed by a factor of m_h^2/m_s^2 compared to that from h' and ϕ. The two terms inside the parenthesis are given by

$$\mathrm{s}_\gamma\,\mathrm{Im}(\beta^{h'}_{ij}\gamma^\phi_{ji}) = \mathrm{s}_\gamma\frac{v}{2}\frac{\tilde{y}'^*_{ij}\tilde{\bar{y}}_{jk}\tilde{y}_{ik}}{m_{D_k}} \sim \mathrm{s}_\gamma\frac{m_{d_i}\bar{y}}{v'} \sim \frac{v_\phi m_{d_i}\bar{y}}{v'^2}, \tag{A.46}$$

where in the last step we have substituted $\mathrm{s}_\gamma \sim v_\phi/v'$, as we expect when $v_\phi \lesssim v'$, and

$$\mathrm{s}_\gamma\mathrm{s}_\alpha\,\mathrm{Im}(\beta^\phi_{ij}\gamma^h_{ji}) = \mathrm{s}_\gamma\mathrm{s}_\alpha\frac{v'}{2}\frac{\tilde{y}'^*_{ik}\tilde{\bar{y}}_{kj}\tilde{y}_{ij}}{m_{D_k}} \sim \mathrm{s}_\gamma\mathrm{s}_\alpha\frac{m_{d_i}\bar{y}}{v} \sim \frac{v_\phi m_{d_i}\bar{y}}{v'^2}. \tag{A.47}$$

Both terms are therefore of the same order. When $v_\phi \lesssim v'$, the contribution from ϕ to $\mathrm{Im}(\delta m_{d_i})$ is subleading to that from h'. Setting $m_{h'} \simeq \sqrt{2\lambda}v' \sim v'$, we then have

$$\mathrm{Im}(\delta m_{d_i}) \sim \frac{m_{d_i}}{16\pi^2}\frac{\bar{y}v_\phi}{M}\log\frac{M^2}{m_{h'}^2}, \tag{A.48}$$

and the contribution to $\bar{\theta}$ from the down-quark sector reads

$$\bar{\theta} \simeq \sum_i \frac{\mathrm{Im}(\delta m_{d_i})}{m_{d_i}} \sim \frac{1}{16\pi^2}\frac{\bar{y}v_\phi}{M}\log\frac{M^2}{m_{h'}^2}. \tag{A.49}$$

The above expression agrees with the parametric estimate presented in section 2.4.2, except for the log factor that is not captured in our spurion analysis.

A.3 Kaon mixing

The Δm_K and $|\epsilon_K|$ parameters characterizing the kaon sector can be written as

$$\Delta m_K = 2\mathrm{Re}(m_{12}^K), \qquad \text{and} \qquad |\epsilon_K| = \frac{\kappa_\epsilon |\,\mathrm{Im}(m_{12}^K)|}{\sqrt{2}\Delta m_K}, \tag{A.50}$$

where $m_{12}^K \equiv \frac{1}{2m_K}\langle K^0|\mathcal{H}_{\rm eff}|\bar{K}^0\rangle$, and $\mathcal{H}_{\rm eff}$ refers to the effective hamiltonian appropriate to describe kaon mixing.

In the SM, $\mathcal{H}_{\rm eff}$ is generated at one-loop through box diagrams involving two W gauge bosons. The corresponding contribution reads

$$\mathcal{H}_{\rm eff} \supset -\frac{G_F^2 m_W^2}{4\pi^2}(\bar{d}_L\gamma_\mu s_L)(\bar{d}_L\gamma^\mu s_L)\sum_{\alpha,\beta}\lambda_\alpha\lambda_\beta F(x_\alpha,x_\beta) + \text{h.c.}, \tag{A.51}$$

where the loop function F is given by [344]

$$\begin{aligned} F(x_\alpha,x_\beta) = {} & \frac{x_\alpha^2\log x_\alpha}{(x_\beta-x_\alpha)(1-x_\alpha)^2}\left(1-2x_\beta+\frac{x_\alpha x_\beta}{4}\right) + \{x_\alpha \leftrightarrow x_\beta\} \\ & + \frac{1}{(1-x_\alpha)(1-x_\beta)}\left(\frac{7x_\alpha x_\beta}{4}-1\right), \end{aligned} \tag{A.52}$$

and $\lambda_\alpha = V_{\alpha d}^* V_{\alpha s}$ for $\alpha = u,c,t$. In the present model, the sum over α and β in Eq.(A.51) must be extended to include the additional members of the up-quark sector. The corresponding couplings can be read off from Eq.(2.39), and are given by

$$\lambda_\alpha = \Delta V_{\alpha d}^* \Delta V_{\alpha s} \qquad \text{for} \qquad \alpha = U,C,T. \tag{A.53}$$

An additional contribution to $\mathcal{H}_{\text{eff}}$ arises from diagrams involving one W and one W'. In this case:

$$\mathcal{H}_{\text{eff}} \supset -\frac{G_F^2 m_W^2}{4\pi^2}\beta(\bar{d}_R s_L)(\bar{d}_L s_R)\sum_{\alpha,\beta}\lambda_\alpha^{LR}\lambda_\beta^{RL}\tilde{F}(\beta, x_\alpha, x_\beta) + \text{h.c.}, \tag{A.54}$$

where $\beta \equiv m_W^2/m_{W'}^2 = v^2/v'^2$, and the loop function now reads [345]

$$\tilde{F}(\beta, x_\alpha, x_\beta) = \sqrt{x_\alpha x_\beta}\left\{(1+\beta)I_2(x_\alpha, x_\beta, \beta) - (4+\beta x_\alpha x_\beta)I_1(x_\alpha, x_\beta, \beta)\right\}, \tag{A.55}$$

with

$$\begin{aligned} I_1(\beta, x_\alpha, x_\beta) &= \frac{x_\alpha \log x_\alpha}{(1-x_\alpha)(1-\beta x_\alpha)(x_\alpha - x_\beta)} + \{x_\alpha \leftrightarrow x_\beta\} - \frac{\beta \log \beta}{(1-\beta)(1-\beta x_\alpha)(1-\beta x_\beta)}, \\ I_2(\beta, x_\alpha, x_\beta) &= \frac{x_\alpha^2 \log x_\alpha}{(1-x_\alpha)(1-\beta x_\alpha)(x_\alpha - x_\beta)} + \{x_\alpha \leftrightarrow x_\beta\} - \frac{\log \beta}{(1-\beta)(1-\beta x_\alpha)(1-\beta x_\beta)}. \end{aligned} \tag{A.56}$$

The relevant couplings follow from the interactions in Eq.(2.39) and (2.40). For the u and c quarks, we have $\lambda_\alpha^{LR} = \lambda_\alpha^{RL} = \lambda_\alpha$, whereas for their heavy partners

$$\lambda_\alpha^{LR} = \Delta V_{\alpha d}^* \Delta V'_{\alpha s}, \qquad \text{and} \qquad \lambda_\alpha^{RL} = \Delta V'^*_{\alpha d}\Delta V_{\alpha s}, \qquad \text{for} \qquad \alpha = U, C. \tag{A.57}$$

In the top sector, on the other hand, we have

$$\lambda_t^{LR} = V_{td}^*\Delta V'_{3s},\ \lambda_t^{RL} = \Delta V'^*_{3d}V_{ts}, \qquad \text{and} \qquad \lambda_T^{LR} = \Delta V_{3d}^* V_{ts},\ \lambda_T^{RL} = V_{td}^*\Delta V_{3s}. \tag{A.58}$$

Appendix B

Supplementary Material for Chapter 3

B.1 Freeze-in of γ'

As discussed in Sec. 3.5.3 and Sec. 3.5.4, a thermal population of mirror photon is not consistent with the mirror electron being DM. If the latter is produced via a freeze-out mechanism, it leads to an overabundance of mirror up-quarks, forbidden by direct searches. If the e' is produced via freeze-in, it leads to a yield larger than the observed one for DM. Therefore, we instead considered a frozen-in population of γ' that re-scatters into e'.

Computing the yield of e' through this mechanism then requires tracking the distribution $f_{\gamma'}(t, E)$ of mirror photons as they are produced from gluons thorugh the processes of Fig. 3.6. Neglecting all terms proportional to $f_{\gamma'}$ at leading order, the Boltzmann equation reads

$$\frac{\partial f_{\gamma'}}{\partial t} - HE\frac{\partial f_{\gamma'}}{\partial E} = \frac{1}{E}\int d\Pi_1 d\Pi_2 d\Pi_3 \, (2\pi)^4 \, \delta^{(4)}(\mathbf{p}_1+\mathbf{p}_2-\mathbf{p}_3-\mathbf{p}_{\gamma'})|\mathcal{M}_{12\to 3\gamma'}|^2 f_1(t,p_1) f_2(t,p_2) \tag{B.1}$$

with $d\Pi_i \equiv \frac{d^3\mathbf{p}_i}{(2\pi)^3 2E_i}$ and $p \equiv |\mathbf{p}|$. In our case, the particles labeled by $1, 2$ are SM gluons in thermal equilibrium, and 3 is a mirror photon or SM gluon in the final state. Anticipating that the mirror photons will create heavy e', we focus on high energy mirror photons, and hence on high energy gluons. In addition, since the gluons are in thermal equilibrium, we can approximate $f_{1,2}(t,p) \sim e^{-|p|/T}$ in the comoving frame. Then, via momentum conservation, we find

$$f_1(t,p_1)f_2(t,p_2) = e^{-(p_3+E)/T} \ . \tag{B.2}$$

Consequently, we can first compute

$$\int d\Pi_1 d\Pi_2 \, (2\pi)^4 \, \delta^{(4)}(\mathbf{p}_1 + \mathbf{p}_2 - \mathbf{p}_3 - \mathbf{p}_{\gamma'})|\mathcal{M}_{12\to 3\gamma'}|^2, \tag{B.3}$$

which is Lorentz invariant. In the center of mass frame, it reads

$$\int d\cos\theta \frac{1}{16\pi} \left| \mathcal{M}\left(s, t = -\frac{s}{2}\left(1+\cos\theta\right)\right) \right|^2 \ . \tag{B.4}$$

In the present case, we have computed the cross sections to be,

$$\begin{aligned} |\mathcal{M}_{gg\to\gamma'\gamma'}|^2 &= \frac{e'^4 g_s^4 \left(145586s^4 + 516433s^3t + 725757s^2t^2 + 450254st^3 + 141256t^4\right)}{114307200\pi^4 m_{u'}^8} \\ |\mathcal{M}_{gg\to g\gamma'}|^2 &= \frac{e'^2 g_s^6 \left(145586s^4 + 516433s^3t + 725757s^2t^2 + 450254st^3 + 141256t^4\right)}{121927680\pi^4 m_{u'}^8} \ . \end{aligned} \tag{B.5}$$

We integrate over $\cos\theta$ and express the result in terms of s in order to have the expression in any Lorentz frame, including the comoving frame where we know the gluon distribution:

$$gg \to \gamma'\gamma' \ : \frac{224881 e'^4 g_s^4 s^4}{4572288000\pi^5 m_{u'}^8} \ , \qquad gg \to g\gamma' \ : \quad \frac{224881 e'^2 g_s^6 s^4}{4877107200\pi^5 m_{u'}^8} \ . \tag{B.6}$$

Performing the last integration over $\mathbf{p}_3$, we obtain

$$\int d\Pi_3 s^4 e^{-(p_3+E)/T} = \frac{2E^4 e^{-E/T}}{(2\pi)^2} \int dp_3 d\cos\theta \, p_3^5 \left(1-\cos\theta\right)^4 e^{-(p_3+E)/T} = \frac{1536 E^4 T^6 e^{-E/T}}{(2\pi)^2} \ , \tag{B.7}$$

where we have used the rotational invariance of the system to align $\mathbf{p}_{\gamma'}$ with the z-axis. Summing over the two production channels, we find

$$\frac{\partial f_{\gamma'}}{\partial t} - HE\frac{\partial f_{\gamma'}}{\partial E} = \frac{224881 e'^2 g_s^4 \left(16e'^2 + 15g_s^2\right) T^6 E^3 e^{-\frac{E}{T}}}{190512000\pi^7 m_{u'}^8}. \tag{B.8}$$

We then wish to manipulate this equation into a numerically-friendly form. First, we convert the time derivative into a temperature derivative using $s = \frac{2\pi^2}{45} h_{\rm eff} T^3$,

$$\frac{\partial f_{\gamma'}}{\partial t} = \frac{\partial f_{\gamma'}}{\partial T}\frac{\partial T}{\partial s}\frac{\partial s}{\partial t} = -\left(\frac{8\pi^3}{90}\right)^{1/2} \frac{T^3}{M_{\rm pl}} \frac{h_{\rm eff}(T)}{g_*^{1/2}(T)} \frac{\partial f_{\gamma'}}{\partial T}\,, \tag{B.9}$$

where $g_*^{1/2} \equiv \frac{h_{\rm eff}}{g_{\rm eff}^{1/2}}\left(1 + \frac{T}{3h_{\rm eff}}\frac{dh_{\rm eff}}{dT}\right)$, and we write the Hubble parameter in terms of temperature,

$$H(T) = \left(\frac{\pi^3 g_{\rm eff}(T)}{90}\right)^{1/2} \frac{T^2}{M_P}. \tag{B.10}$$

The equation becomes

$$T\frac{\partial f_{\gamma'}}{\partial T} + \alpha_1(T) E \frac{\partial f_{\gamma'}}{\partial E} = -\alpha_2(T) T^4 E^3\, e^{-E/T}, \tag{B.11}$$

where we define $\alpha_1(T) = \left(\frac{g_{\rm eff}(T) g_*(T)}{8 h_{\rm eff}^2(T)}\right)^{1/2}$ and $\alpha_2(T) = \frac{224881 e'^2 g_s^4 \left(16e'^2+15g_s^2\right)}{190512000\pi^7 m_{u'}^8} \frac{g_*^{1/2}(T)}{h_{\rm eff}(T)} \sqrt{\frac{90}{8\pi^3}} M_P$ for simplicity. Because we know that the freeze-in is UV-dominated, we can take α_1 and α_2 to be constant, giving us the analytical solution,

$$f(T,E) = \frac{\alpha_2 E^3 T^{-3\alpha_1}}{\alpha_1 - 1}\left[T^{3\alpha_1+4}\left(\frac{E}{T}\right)^{\frac{3\alpha_1+4}{1-\alpha_1}} \Gamma\left(\frac{3\alpha_1+4}{\alpha_1-1}, \frac{E}{T}\right)\right.$$
$$\left. - T_{\rm RH}^{3\alpha_1+4}\left(E\, T_{\rm RH}^{\alpha_1-1} T^{-\alpha_1}\right)^{\frac{3\alpha_1+4}{1-\alpha_1}} \Gamma\left(\frac{3\alpha_1+4}{\alpha_1-1}, T^{-\alpha_1} E\, T_{\rm RH}^{\alpha_1-1}\right)\right]. \tag{B.12}$$

where Γ is the so-called incomplete gamma function. In our Boltzmann codes, we are using the above formula even for energies larger than $m_{u'}$, although Eqs. (B.5) only hold for $s \leq m_{u'}^2$. However, as in footnote 7, energies larger than $m_{u'}$ essentially do not affect the γ' distribution and the resulting e' relic abundance.

Appendix C

Supplementary Material for Chapter 4

C.1 Scalar-induced gravitational waves: technical details

In this appendix, we review the formalism relevant to computing GW energy density for the sake of completeness, following the notation and analysis of Ref. [242].

C.1.1 Transfer functions

The equation of motion for the scalar perturbation Φ in the absence of isocurvature perturbations is,

$$\Phi''(\tau, \mathbf{k}) + 3(1 + c_s^2)\mathcal{H}\Phi'(\tau, \mathbf{k}) + c_s^2 k^2 \Phi(\tau, \mathbf{k}) = 0\,, \tag{C.1}$$

where $c_s^2 \simeq w$ is the sound speed of the fluid. Defining dimensionless parameter $y = \sqrt{w}k\tau$, we rewrite this equation as

$$\frac{\mathrm{d}^2\Phi(y,\mathbf{k})}{\mathrm{d}y^2} + \frac{6(1+w)}{1+3w}\frac{1}{y}\frac{\mathrm{d}\Phi(y,\mathbf{k})}{\mathrm{d}y} + \Phi(y,\mathbf{k}) = 0\,. \tag{C.2}$$

A general solution is given by,

$$\Phi(y,\mathbf{k}) = y^{-\gamma}\left[C_1(\mathbf{k})J_\gamma(y) + C_2(\mathbf{k})Y_\gamma(y)\right]\,, \tag{C.3}$$

where J_γ and Y_γ are spherical Bessel functions of the first and second kind, respectively, of order γ

$$\gamma = \frac{3(1+w)}{1+3w} - 1\,. \tag{C.4}$$

In the radiation dominated era, in which $w = 1/3 \to \gamma = 1$, we have

$$\Phi(y,\mathbf{k}) = \frac{1}{y^2}\left[C_1(\mathbf{k})\left(\frac{\sin y}{y} - \cos y\right) + C_2(\mathbf{k})\left(\frac{\cos y}{y} + \sin y\right)\right]. \tag{C.5}$$

We can deduce the initial conditions of this solution by considering the early-time limit $k\tau \ll 1$,

$$\frac{\sin y}{y} - \cos y \simeq \frac{y^2}{3} \quad \text{and} \quad \frac{\cos y}{y} + \sin y \simeq \frac{1}{y}\,. \tag{C.6}$$

The first term ($\propto C_1$) is then constant in this limit, while the second term ($\propto C_2$) decays as $1/y^3 \sim 1/a^3$. We can therefore assume the initial conditions,

$$C_1(\mathbf{k}) = 2\zeta(\mathbf{k}), \quad C_2(\mathbf{k}) = 0\,, \tag{C.7}$$

which gives a particular solution,

$$\Phi(\tau,\mathbf{k}) = \frac{2}{3}\zeta(\mathbf{k})\frac{3}{y^2}\left(\frac{\sin y}{y} - \cos y\right)\,, \tag{C.8}$$

resulting in the transfer function, via (4.66),

$$T_\Phi(k\tau) = \frac{3}{(k\tau/\sqrt{3})^3}\left(\sin\frac{k\tau}{\sqrt{3}} - \frac{k\tau}{\sqrt{3}}\cos\frac{k\tau}{\sqrt{3}}\right)\,. \tag{C.9}$$

We can now see the distinct behavior of super-horizon ($k\tau \ll 1$) and sub-horizon ($k\tau \gg 1$) modes in the radiation dominated era. While the super-horizon modes freeze via our analysis above, the sub-horizon modes oscillate and damp as $\sim \cos k\tau/(k\tau)^2$.

In the matter dominated era, $w = 0$ and the equation of motion for Φ becomes,

$$\Phi''(\tau, \mathbf{k}) + 3\mathcal{H}\Phi'(\tau, \mathbf{k}) = 0\,, \tag{C.10}$$

leading to a constant transfer function.

C.1.2 Green's function and GW solution

In this subsection, we discuss in detail the solutions to eq. (4.62), which is derived using the second-order Einstein equation, $G^{(2)}_{ij} = 8\pi G T^{(2)}_{ij}$, for second-order tensor and first-order scalar contributions. We neglect scalar anisotropic stress, and second-order vector and scalar perturbations. In other words, we use the following perturbed FLRW metric in the Newtonian gauge,

$$\mathrm{d}s^2 = -\left(1 + 2\Phi\right)\mathrm{d}t^2 + a^2\left(\left(1 - 2\Phi\right)\delta_{ij} + \frac{1}{2}h_{ij}\right)\mathrm{d}x^i\mathrm{d}x^j, \tag{C.11}$$

assuming a perfect fluid energy-momentum tensor with equation of state w. Using lower order solutions and projecting out spatial indices using polarization tensors, i.e. $\epsilon^{ij}_{\lambda} T_{ij} = T_{\lambda}$ for any tensor T, we recover (4.62). For simplicity, we define a new variable $v(\tau, \mathbf{k}) = a h_{\lambda}(\tau, \mathbf{k})$, which gives the equation of motion for $v(\tau, \mathbf{k})$,

$$v''(\tau, \mathbf{k}) + \left[k^2 - \frac{a''(\tau)}{a(\tau)}\right] v(\tau, \mathbf{k}) = 4a(\tau)\mathcal{S}_{\lambda}(\tau, \mathbf{k})\,. \tag{C.12}$$

We need the two homogeneous solutions of this equation $v_1(\tau)$ and $v_2(\tau)$ to construct the Green's function,

$$G_{\mathbf{k}}(\tau, \bar{\tau}) = \frac{v_1(\tau)v_2(\bar{\tau}) - v_1(\bar{\tau})v_2(\tau)}{v_1'(\bar{\tau})v_2(\bar{\tau}) - v_1(\bar{\tau})v_2'(\bar{\tau})}\,. \tag{C.13}$$

For each $\mathbf{k}$ we have

$$v''_{1,2}(\tau) + \left[k^2 - \frac{a''(\tau)}{a(\tau)}\right] v_{1,2}(\tau) = 0 \tag{C.14}$$

which, using $a \propto \tau^\alpha$ and $x = k\tau$, leads to

$$\frac{\mathrm{d}^2 v_{1,2}(x)}{\mathrm{d}x^2} + \left[1 - \frac{\alpha(\alpha-1)}{x^2}\right] v_{1,2}(x) = 0\,, \tag{C.15}$$

where $\alpha = 2/(1+3w)$. The solutions are

$$v_1(x) = \sqrt{x} J_{\alpha-1/2}(x) \tag{C.16}$$

$$v_2(x) = \sqrt{x} Y_{\alpha-1/2}(x) \tag{C.17}$$

where $J_{\alpha-1/2}$ and $Y_{\alpha-1/2}$ are again spherical Bessel functions of first and second kind, respectively. We note that

$$\frac{\mathrm{d}v_1}{\mathrm{d}x} = \frac{\alpha}{\sqrt{x}} J_{\alpha-1/2}(x) - \sqrt{x} J_{\alpha+1/2} \tag{C.18}$$

$$\frac{\mathrm{d}v_2}{\mathrm{d}x} = \frac{\alpha}{\sqrt{x}} Y_{\alpha-1/2}(x) - \sqrt{x} Y_{\alpha+1/2}\,. \tag{C.19}$$

Now, we can calculate the expression in the denominator of the Green's function,

$$\begin{aligned} v'_1(x)v_2(x) - v_1(x)v'_2(x) &= kx\Big[J_{\alpha-1/2}(x)Y_{\alpha+1/2}(x) - \\ &\qquad J_{\alpha+1/2}(x)Y_{\alpha-1/2}(x)\Big] \\ &= -\frac{2}{\pi}\,. \end{aligned} \tag{C.20}$$

The second equality can be checked explicitly via `Mathematica`. Thus, (C.13) simplifies to

$$\begin{aligned} G_{\mathbf{k}}(\tau,\bar{\tau}) = \frac{\pi}{2}\sqrt{\tau\bar{\tau}}\Big[&J_{\alpha-1/2}(k\bar{\tau})Y_{\alpha-1/2}(k\tau) - \\ &J_{\alpha-1/2}(k\tau)Y_{\alpha-1/2}(k\bar{\tau})\Big]\,. \end{aligned} \tag{C.21}$$

In the radiation dominated era, $\alpha = 1$, and so,

$$G_{\mathbf{k}}(\tau, \bar{\tau}) = \frac{\sin k(\tau - \bar{\tau})}{k}\,, \tag{C.22}$$

where we have used (C.54) to replace Bessel functions of order 1/2. In the matter dominated era we have $\alpha = 2$, and so,

$$G_{\mathbf{k}}(\tau, \bar{\tau}) = \frac{1}{k}\Bigg[\left(\frac{\bar{\tau} - \tau}{\tau\bar{\tau}}\right)\cos k(\tau - \bar{\tau}) + \left(\frac{1/k^2 - \tau\bar{\tau}}{\tau\bar{\tau}}\right)\sin k(\tau - \bar{\tau})\Bigg]\,. \tag{C.23}$$

where we have again used (C.54) to replace Bessel functions of order 3/2.

Having calculated the Green's functions, we can now write the solution for $h_\lambda(\tau, \mathbf{k})$ in the form of (4.67).

C.1.3 Connected and disconnected 4-point correlation function

The primordial 4-point correlation function of ζ can be written in terms of disconnected and connected pieces

$$\begin{aligned}\langle\zeta(\mathbf{k}_1)\zeta(\mathbf{k}_2)\zeta(\mathbf{k}_3)\zeta(\mathbf{k}_4)\rangle =& \langle\zeta(\mathbf{k}_1)\zeta(\mathbf{k}_2)\zeta(\mathbf{k}_3)\zeta(\mathbf{k}_4)\rangle_{\rm d} \\ &+ \langle\zeta(\mathbf{k}_1)\zeta(\mathbf{k}_2)\zeta(\mathbf{k}_3)\zeta(\mathbf{k}_4)\rangle_{\rm c},\end{aligned} \tag{C.24}$$

where

$$\begin{aligned}&\langle\zeta(\mathbf{k}_1)\zeta(\mathbf{k}_2)\zeta(\mathbf{k}_3)\zeta(\mathbf{k}_4)\rangle_{\rm d} = \\ &(2\pi)^6\delta^3(\mathbf{k}_1 + \mathbf{k}_2)\delta^3(\mathbf{k}_3 + \mathbf{k}_4)P_\zeta(k_1)P_\zeta(k_3) \\ &+ (2\pi)^6\delta^3(\mathbf{k}_1 + \mathbf{k}_3)\delta^3(\mathbf{k}_2 + \mathbf{k}_4)P_\zeta(k_1)P_\zeta(k_2) \\ &+ (2\pi)^6\delta^3(\mathbf{k}_1 + \mathbf{k}_4)\delta^3(\mathbf{k}_2 + \mathbf{k}_4)P_\zeta(k_1)P_\zeta(k_4)\,,\end{aligned} \tag{C.25}$$

and

$$\langle\zeta(\mathbf{k}_1)\zeta(\mathbf{k}_2)\zeta(\mathbf{k}_3)\zeta(\mathbf{k}_4)\rangle_{\rm c} = (2\pi)^3\delta^3(\mathbf{k}_1+\mathbf{k}_2+\mathbf{k}_3+\mathbf{k}_4)\mathcal{T}(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3,\mathbf{k}_4)\,. \tag{C.26}$$

Here, $P_\zeta(k)$ and $\mathcal{T}(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3,\mathbf{k}_4)$ are the scalar power spectrum and trispectrum, respectively. We focus on the disconnected contribution below. The relevant 4-point correlation function for the GW power spectrum (4.69) is

$$\begin{aligned}&\langle\zeta(\mathbf{q}_1)\zeta(\mathbf{k}_1-\mathbf{q}_1)\zeta(\mathbf{q}_2)\zeta(\mathbf{k}_2-\mathbf{q}_2)\rangle_{\rm d} = \\ &(2\pi)^6\delta^3(\mathbf{k}_1+\mathbf{k}_2)\left[\delta^3(\mathbf{q}_1+\mathbf{q}_2)+\delta^3(\mathbf{k}_1+\mathbf{q}_2-\mathbf{q}_1)\right] \\ &\times P_\zeta(q_1)P_\zeta(|\mathbf{k}_1-\mathbf{q}_1|).\end{aligned} \tag{C.27}$$

The two terms in the above expressions are equivalent when substituted in the integrand of (4.69). The second term can be manipulated as

$$\begin{aligned}&\delta^3(\mathbf{k}_1+\mathbf{k}_2)\delta^3(\mathbf{k}_1+\mathbf{q}_2-\mathbf{q}_1)Q_{\lambda_1}(\mathbf{k}_1,\mathbf{q}_1)Q_{\lambda_2}(\mathbf{k}_2,\mathbf{q}_2) \\ &\qquad\times I(|\mathbf{k}_1-\mathbf{q}_1|,q_1,\tau)I(|\mathbf{k}_2-\mathbf{q}_2|,q_2,\tau) \\ &= Q_{\lambda_1}(\mathbf{k}_1,\mathbf{q}_1)Q_{\lambda_2}(-\mathbf{k}_1,\mathbf{q}_1-\mathbf{k}_1)I(|\mathbf{k}_1-\mathbf{q}_1|,q_1,\tau) \\ &\qquad\times I(q_1,|\mathbf{k}_1-\mathbf{q}_1|,\tau) \\ &= Q_{\lambda_1}(\mathbf{k}_1,\mathbf{q}_1)^2 I(|\mathbf{k}_1-\mathbf{q}_1|,q_1,\tau)^2\end{aligned} \tag{C.28}$$

which is the same result we get from the first term. Here we have used identities given in eqs. (C.51)-(C.53). Thus, the disconnected GW power spectrum is given by (4.72).

C.1.4 Recasting integrals for numerical computation

Here we provide steps to recast (4.72) into a form suitable for numerical integration.

Change of variables. We perform two successive changes of variables to recast the integrals. First, we perform the transformation $\{q, \cos\theta\} \to \{u, v\}$, where

$$u \equiv \frac{|\mathbf{k}-\mathbf{q}|}{k}\,, \quad v \equiv \frac{q}{k}, \tag{C.29}$$

and the inverse transformation is

$$q = vk\,, \quad \cos\theta = \frac{1+v^2-u^2}{2v}\,. \tag{C.30}$$

The determinant of the Jacobian for this transformation is,

$$\det(J_{\{q,\cos\theta\}\to\{u,v\}}) = -\partial_v q \partial_u \cos\theta = -\frac{ku}{v}\,. \tag{C.31}$$

which implies

$$\begin{aligned}\int \mathrm{d}^3q &= \int_0^\infty q^2\mathrm{d}q \int_{-1}^1 \mathrm{d}\cos\theta \int_0^{2\pi} \mathrm{d}\phi \\ &= k^3 \int_0^\infty \mathrm{d}v\, v \int_{|1-v|}^{1+v} \mathrm{d}u\, u \int_0^{2\pi} \mathrm{d}\phi\,.\end{aligned} \tag{C.32}$$

Second, we perform $\{u, v\} \to \{s, t\}$ where

$$s \equiv u - v\,, \quad t \equiv u + v - 1\,, \tag{C.33}$$

and the inverse transformation is

$$u = \frac{s+t+1}{2}\,, \quad v = \frac{t-s+1}{2}\,. \tag{C.34}$$

The determinant of the Jacobian for the second transformation is then

$$\det(J_{\{u,v\}\to\{s,t\}}) = \frac{1}{2}\,. \tag{C.35}$$

Hence, we have[1]

$$\int_0^\infty \mathrm{d}v \int_{|1-v|}^{1+v} \mathrm{d}u = \frac{1}{2}\int_0^\infty \mathrm{d}t \int_{-1}^1 \mathrm{d}s. \tag{C.36}$$

[1]For $v < 1$, the lower limit of integration over s is $1 - 2v$. However, in this case we already have $1 - 2v > -1$.

The final result is

$$\int \mathrm{d}^3 q = \frac{k^3}{2} \int_0^\infty \mathrm{d}t \int_{-1}^1 \mathrm{d}s \; uv \int_0^{2\pi} \mathrm{d}\phi \,. \tag{C.37}$$

Above, we express the integrand in terms of u and v for convenience, though the integration itself is done in terms of s and t.

Analytic result for the $I(p, q, \tau)$ function. We summarize the results for a radiation-dominated universe (for a more in-depth look, see e.g. [240]). At late times, we have

$$\begin{aligned} I(vk, uk, x/k \to \infty) &= \frac{1}{k^2} I(u, v, x \to \infty) \\ &\simeq \frac{1}{k^2} \frac{1}{x} \tilde{I}_A(u, v) \left(\tilde{I}_B(u, v) \sin x + \tilde{I}_C \cos x \right), \end{aligned} \tag{C.38}$$

where we define

$$\tilde{I}_A(u, v) \equiv \frac{3(u^2 + v^2 - 3)}{4u^3 v^3} \tag{C.39a}$$

$$\tilde{I}_B(u, v) \equiv -4uv + (u^2 + v^2 - 3) \ln \left| \frac{3 - (u + v)^2}{3 - (u - v)^2} \right| \tag{C.39b}$$

$$\tilde{I}_C(u, v) \equiv -\pi (u^2 + v^2 - 3) \Theta(u + v - \sqrt{3}) \,. \tag{C.39c}$$

In the last expression, Θ is the Heaviside theta function. This result redshifts as $1/x \propto 1/a$. Using the above definitions, we compute the quantity given in C.28,

$$\begin{aligned} &\frac{Q_+(\mathbf{k}, \mathbf{q})}{\cos 2\phi} I(|\mathbf{k} - \mathbf{q}|, q, \tau) \\ &= \frac{Q_\times(\mathbf{k}, \mathbf{q})}{\sin 2\phi} I(|\mathbf{k} - \mathbf{q}|, q, \tau) \\ &= \frac{v^2 k^2}{\sqrt{2}} \frac{4v^2 - (1 + v^2 - u^2)^2}{4v^2} I(uk, vk, x/k) \\ &\equiv \frac{\tilde{\mathcal{J}}(u, v)}{\sqrt{2}} k^2 I(uk, vk, x/k), \end{aligned} \tag{C.40}$$

where we have used dimensionless conformal time $x = k\tau$ and defined

$$\tilde{\mathcal{J}}(u, v) = \frac{4v^2 - (1 + v^2 - u^2)^2}{4} \,. \tag{C.41}$$

When computing the GW power spectrum we are generically interested in the time-averaged quantity

$$\begin{aligned}&\overline{k^2 I(v_1 k, u_1 k, x/k \to \infty) k^2 I(v_2 k, u_2 k, x/k \to \infty)} = \\ &\frac{1}{2x^2}\tilde{I}_A(u_1, v_1)\tilde{I}_A(u_2, v_2) \\ &\times \left[\tilde{I}_B(u_1, v_1)\tilde{I}_B(u_2, v_2) + \tilde{I}_C(u_1, v_1)\tilde{I}_C(u_2, v_2)\right].\end{aligned} \tag{C.42}$$

Azimuthal angle integration. In the disconnected contribution (4.72), the only ϕ-dependent factors in the integrands are $\sin 2\phi$ and $\cos 2\phi$, coming from Q_λ factors. For each polarization, we then have

$$\int_0^{2\pi} \mathrm{d}\phi \sin^2(2\phi) = \int_0^{2\pi} \mathrm{d}\phi \cos^2(2\phi) = \pi\,. \tag{C.43}$$

Finally, we are ready to numerically compute the GW energy density (4.75) which is defined in terms of the dimensionless polarization-averaged GW power spectrum

$$\sum_\lambda \Delta^2_\lambda(\tau, k) = \frac{k^3}{2\pi^2}\sum_\lambda P_\lambda(\tau, k). \tag{C.44}$$

Using our recasted variables, the result is

$$\begin{aligned}\Omega_{\rm GW}(k)\Big|_{\rm d} =& \frac{2}{48\alpha^2}\left(\frac{k^3}{2\pi^2}\right)^2 \\ &\int_0^\infty \mathrm{d}t \int_{-1}^1 \mathrm{d}s uv \tilde{\mathcal{J}}(u, v)^2 \tilde{I}_A(u, v)^2 \Big[\tilde{I}_B(u, v)^2 \\ &+ \tilde{I}_C(u, v)^2\Big] P_\zeta(uk) P_\zeta(vk)\end{aligned} \tag{C.45}$$

More compactly,

$$\Omega_{\rm GW}(k)\Big|_{\rm d} = \frac{2}{48\alpha^2}\int_0^\infty \mathrm{d}t \int_{-1}^1 \mathrm{d}s\; \mathcal{K}_{\rm d}(u, v)\Delta^2_\zeta(uk)\Delta^2_\zeta(vk) \tag{C.46}$$

where we define the following the Kernel functions $\mathcal{K}_{\rm d}$ for simplified notation,

$$\mathcal{K}_{\rm d}(u, v) = (uv)^{-2}\tilde{\mathcal{J}}(u, v)^2 \tilde{I}_A(u, v)^2 \left[\tilde{I}_B(u, v)^2 + \tilde{I}_C(u, v)^2\right]. \tag{C.47}$$

C.1.5 Useful formula

The projection operator Q_λ (4.65) is defined as,

$$Q_\lambda(\mathbf{k},\mathbf{q}) \equiv \epsilon_\lambda^{ij}(\mathbf{k}) q_i q_j = -\epsilon_\lambda^{ij}(\mathbf{k})(\mathbf{k}-\mathbf{q})_i q_j, \tag{C.48}$$

where the second equality follows from $\epsilon_\lambda^{ij}(\mathbf{k}) k_i = 0$. If we explicitly set $\hat{k} = \hat{z}$, we have $\mathbf{q} = q(\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$, where θ and ϕ are polar and azimuthal angles. This leads to the expressions,

$$\begin{aligned} Q_+(\mathbf{k},\mathbf{q}) &= \frac{q^2}{\sqrt{2}} \sin^2\theta\cos(2\phi)\,, \\ Q_\times(\mathbf{k},\mathbf{q}) &= \frac{q^2}{\sqrt{2}} \sin^2\theta\sin(2\phi)\,. \end{aligned} \tag{C.49}$$

Since $\epsilon_\lambda(\mathbf{k})$ is orthogonal to $\mathbf{k}$ we have

$$Q_\lambda(\mathbf{k},\mathbf{q}) = Q_\lambda(\mathbf{k},\mathbf{q}+c\mathbf{k})\,, \tag{C.50}$$

for any constant c. $Q_\lambda(\mathbf{k},\mathbf{q})$ is also symmetric under $\mathbf{k}\to-\mathbf{k}$ and $\mathbf{q}\to-\mathbf{q}$:

$$Q_\lambda(\mathbf{k},\mathbf{q}) = Q_\lambda(-\mathbf{k},\mathbf{q}) = Q_\lambda(\mathbf{k},-\mathbf{q}) = Q_\lambda(-\mathbf{k},-\mathbf{q})\,. \tag{C.51}$$

Using (4.71) we see that

$$f(p,q,\tau) = f(q,p,\tau) \tag{C.52}$$

and so

$$I(p,q,\tau) = I(q,p,\tau)\,. \tag{C.53}$$

Bessel functions. The following formulae are helpful for computations involving Bessel functions:

$$
\begin{aligned}
J_{1/2}(x) &= \sqrt{\frac{2}{\pi x}}\sin x\,, \\
Y_{1/2}(x) &= -\sqrt{\frac{2}{\pi x}}\cos x\,, \\
J_{3/2}(x) &= \sqrt{\frac{2}{\pi x}}\left(\frac{\sin x}{x} - \cos x\right)\,, \\
Y_{3/2}(x) &= -\sqrt{\frac{2}{\pi x}}\left(\frac{\cos x}{x} - \sin x\right)\,.
\end{aligned}
\tag{C.54}
$$

Appendix D

Supplementary Material for Chapter 5

D.1 IBP, EOM, and Field Redefinitions in dS: An Explicit Example

We are interested in computing cosmological correlation functions on a late time slice at η_0. Therefore, the boundary terms that arise while employing IBP can potentially contribute to such correlation functions. Here we investigate the nature of these boundary terms by focusing on a concrete example involving a massive field σ having a mass m, defined in terms of the variable $\mu = (m^2/H^2 - 9/4)^{1/2}$. In particular, we focus on the standard quasi-single field operator [280]:

$$\mathcal{O}_1 = \frac{1}{\Lambda}\int \mathrm{d}^4x\sqrt{|g|}(\nabla_\mu\phi)(\nabla^\mu\phi)\sigma. \tag{D.1}$$

By IBP we can write this as,

$$\mathcal{O}_1 = \frac{1}{\Lambda}\int \mathrm{d}^4x\sqrt{|g|}\left[\nabla_\mu(\phi\nabla^\mu\phi\cdot\sigma) - \phi\Box\phi\cdot\sigma - \phi\nabla^\mu\phi\nabla_\mu\sigma\right]. \tag{D.2}$$

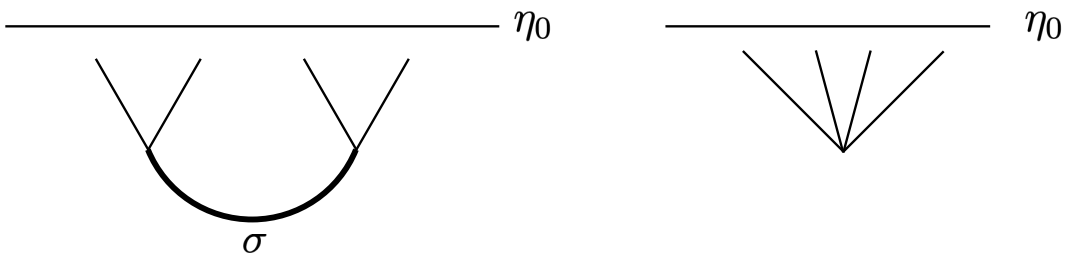

Figure D.1: *Left.* Four point function mediated by $\mathcal{O}_1$ and $\mathcal{O}_2$. *Right.* Contact interaction mediated by $\mathcal{O}_{2,\rm TD}$.

The term involving $\Box\phi$ does not contribute to in-in correlators by virtue of the inflaton EOM $\Box\phi \approx 0$, where we ignore the contribution from the inflaton potential. We rewrite the remaining terms as,

$$\begin{aligned}\mathcal{O}_2 &= \frac{1}{2\Lambda}\int \mathrm{d}^4x\sqrt{|g|}\left[\nabla_\mu(\nabla^\mu(\phi^2)\cdot\sigma) - \nabla^\mu(\phi^2)\nabla_\mu\sigma\right] \\ &= \mathcal{O}_{2,\rm TD} + \mathcal{O}_{2,\rm Bulk}.\end{aligned} \tag{D.3}$$

Here we have separated the total derivative (TD) and the bulk term. As noted in [285], the boundary term can be neglected for equal-time correlation functions in the in-in formalism because such terms are associated with equal-time commutators and can be removed by a redefinition of the local operators. In what follows, we explore this in great detail.

We first check that $\mathcal{O}_1$ and $\mathcal{O}_2$ give exactly the same contribution to in-in correlation functions, as they should. Our goal would then be to understand the contribution from $\mathcal{O}_{2,\rm TD}$. To that end, we focus on a four point function as shown in the left panel of Fig. D.1. There are four terms contributing to this four point function. We call the associated contributions as, $\mathcal{I}_{++}$, $\mathcal{I}_{+-}$, $\mathcal{I}_{-+}$, and $\mathcal{I}_{--}$, of which the third and the fourth are the complex conjugates of the second and the first, respectively (the $+(-)$ sign denotes that the vertex come from a time (anti-time) ordering operator). Therefore, we only consider $\mathcal{I}_{++}$ and $\mathcal{I}_{+-}$ to show the equivalence of $\mathcal{O}_1$ and $\mathcal{O}_2$. For convenience, we

write the TD term as,

$$\mathcal{O}_{2,\rm TD} = \frac{1}{2\Lambda}\int \mathrm{d}^4x\partial_\mu\left(\sqrt{|g|}\partial^\mu(\phi^2)\cdot\sigma\right) = -\frac{1}{2\Lambda}\int \mathrm{d}\eta\partial_\eta\left(\frac{1}{\eta^2}\partial_\eta(\phi^2)\cdot\sigma\right). \tag{D.4}$$

In the last equality, we have dropped the spatial boundary terms, assuming fields decay at spatial infinity. A similar operation can not be naively done for the temporal boundary, since we are interested in computing the correlation functions on the same boundary.

Contribution via $\mathcal{I}_{+-}$. For $\mathcal{I}_{+-}$, the leading contribution involves two separate integrals, one for the time ordering, and the other for anti-time ordering. Schematically they can be written as,

$$\langle\mathcal{O}_1\cdot\phi(\mathbf{k}_1)\phi(\mathbf{k}_2)\phi(\mathbf{k}_3)\phi(\mathbf{k}_4)\cdot\mathcal{O}_1\rangle \text{ and } \langle\mathcal{O}_2\cdot\phi(\mathbf{k}_1)\phi(\mathbf{k}_2)\phi(\mathbf{k}_3)\phi(\mathbf{k}_4)\cdot\mathcal{O}_2\rangle. \tag{D.5}$$

We evaluate these numerically and the comparison between $\mathcal{O}_1$ and $\mathcal{O}_2$ are shown in the left panel of Fig. D.2. The four point function exhibits oscillations as function of the ratio of the momentum of the massive σ particle, $|\mathbf{k}_1+\mathbf{k}_2|$ and the momentum of the inflaton $|\mathbf{k}_1| = k_1$. (For this evaluation we set $k_1 = k_2$.) This confirms a cosmological collider signature from this operator. We can also check explicitly that $\mathcal{O}_{2,\rm TD}$ does not contribute to $\mathcal{I}_{+-}$. To see this, note (D.4) implies $\mathcal{I}_{+-}$ has a contribution from,

$$\propto \frac{1}{\eta_0^2}\left[\partial_\eta(\phi^2)\sigma\right]_{\eta_0}. \tag{D.6}$$

However, this term decays as $\eta_0^{1/2}$ as $\eta_0\to 0$, and does not contribute to $\mathcal{I}_{+-}$.

Contribution via $\mathcal{I}_{++}$. Next we discuss $\mathcal{I}_{++}$ which involves nested time integrals, originating from the time-ordering operation. Schematically, this has the form

$$\langle\mathbb{I}\cdot\phi(\mathbf{k}_1)\phi(\mathbf{k}_2)\phi(\mathbf{k}_3)\phi(\mathbf{k}_4)\cdot\int_{-\infty}^{\eta_0}\mathrm{d}\eta\ \mathcal{O}_2(\eta)\int_{-\infty}^{\eta}\mathrm{d}\eta'\ \mathcal{O}_2(\eta')\rangle, \tag{D.7}$$

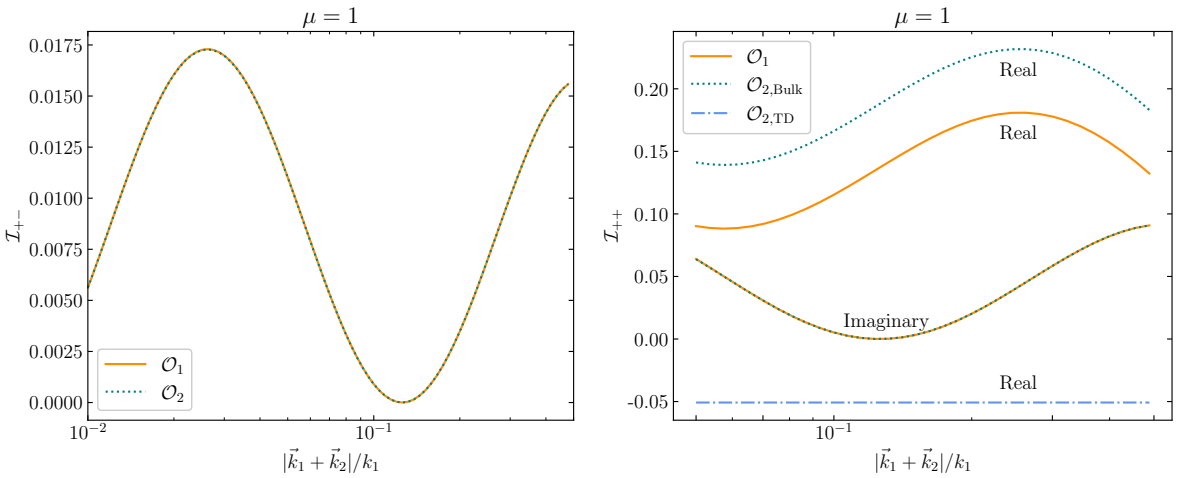

Figure D.2: Comparison between the contributions from $\mathcal{O}_1$ and $\mathcal{O}_2$. *Left.* The $\mathcal{I}_{+-}$ contribution evaluated for the configuration $|\mathbf{k}_1| = |\mathbf{k}_2| = |\mathbf{k}_3| = |\mathbf{k}_4|$ with a varying $|\mathbf{k}_1 + \mathbf{k}_2|$. The contribution from $\mathcal{O}_1$ and $\mathcal{O}_{2,\text{Bulk}}$ match precisely while $\mathcal{O}_{2,\text{TD}}$ does not contribute. *Right.* The real and imaginary parts of the $\mathcal{I}_{++}$ contribution. For the imaginary part, $\mathcal{O}_{2,\text{TD}}$ does not contribute, and $\mathcal{O}_1$ and $\mathcal{O}_{2,\text{Bulk}}$ match, as expected. For the real part, $\mathcal{O}_1$ matches with the sum of $\mathcal{O}_{2,\text{Bulk}}$ and $\mathcal{O}_{2,\text{TD}}$. The contribution from $\mathcal{O}_{2,\text{TD}}$ does not depend upon either the momentum ratio (as seen here) or the mass parameter μ (as can be checked), indicating that it can be thought of as originating from a 'local' contact interaction, such as the one shown schematically in the right panel of Fig. D.1. For both the panels, we do not track the overall factors or momentum dependence.

where we have suppressed the spatial indices. The operator $\mathcal{O}_1$ also contributes via a similar form. From the above form, we note that $\mathcal{O}_{2,\text{TD}}$ could contribute to correlation functions, since the inner integration contributes a non-vanishing integrand for the outer integral. We evaluate this contribution from $\mathcal{O}_{2,\text{TD}}$ numerically and indeed find it to be non-zero, as shown in the right panel of Fig. D.2. There we also find the combined contributions from $\mathcal{O}_{2,\text{TD}}$ and $\mathcal{O}_{2,\text{Bulk}}$ match with $\mathcal{O}_1$.

Importantly, the contribution from $\mathcal{O}_{2,\text{TD}}$ does not vary as a function of $|\mathbf{k}_1+\mathbf{k}_2|/k_1$, unlike the contribution from $\mathcal{O}_{2,\text{Bulk}}$. We also have checked that the contribution from $\mathcal{O}_{2,\text{TD}}$ is independent of the mass of σ for $m > \sqrt{2}H$. These facts indicate that the contribution from $\mathcal{O}_{2,\text{TD}}$ could be thought of as coming from a contact operator that does not involve a σ exchange. This can be understood in two different ways.

First, instead of using the exact mode functions of σ, we can use their sub-horizon, high-energy limits. In detail, we can express σ as,

$$\sigma(\eta,\mathbf{k}) = g_k(\eta)a^\dagger_{\mathbf{k}} + \bar{g}_k(\eta)a_{-\mathbf{k}}, \tag{D.8}$$

with

$$\begin{aligned} g_k(\eta) &= +i\exp(-i\pi/4)\frac{\sqrt{\pi}}{2}\exp(\pi\mu/2)(-\eta)^{3/2}H^{(2)}_{i\mu}(-k\eta), \\ \bar{g}_k(\eta) &= -i\exp(+i\pi/4)\frac{\sqrt{\pi}}{2}\exp(-\pi\mu/2)(-\eta)^{3/2}H^{(1)}_{i\mu}(-k\eta). \end{aligned} \tag{D.9}$$

Using the high-energy limits of the Hankel functions (see, e.g., [320] for explicit expressions), we can evaluate the contribution of $\mathcal{O}_{2,\text{TD}}$. We find,

$$\langle\phi(\mathbf{k}_1)\phi(\mathbf{k}_2)\phi(\mathbf{k}_3)\phi(\mathbf{k}_4)\rangle \propto \left(-\frac{13}{256}+\frac{k_1}{16|\mathbf{k}_1+\mathbf{k}_2|}\right). \tag{D.10}$$

The sub-horizon limit $|\mathbf{k}_1+\mathbf{k}_2|(-\eta)\gg 1$, or equivalently $|\mathbf{k}_1+\mathbf{k}_2|\gg k_1\sim 1/(-\eta)$, then implies the leading contribution is given by the first term in the parenthesis above. This then exactly reproduces the $\mathcal{O}_{2,\text{TD}}$ contribution in the right panel of Fig. D.2.

Alternatively, we can construct a contact term that gives the same contribution as $\mathcal{O}_{2,\mathrm{TD}}$. To that end, we first do a field redefinition:

$$\sigma \to \sigma + \frac{c}{\Lambda}\phi^2. \tag{D.11}$$

Under this redefinition, the σ kinetic term gives rise to,

$$\frac{1}{2}(\nabla_\mu\sigma)(\nabla^\mu\sigma) \to \frac{1}{2}(\nabla_\mu\sigma)(\nabla^\mu\sigma) + \frac{c}{\Lambda}\nabla_\mu\sigma\cdot\nabla^\mu(\phi^2) + \frac{c^2}{2\Lambda^2}\nabla_\mu(\phi^2)\cdot\nabla^\mu(\phi^2). \tag{D.12}$$

For $c = 1/2$, we reproduce the form of the bulk term $\mathcal{O}_{2,\mathrm{Bulk}}$. This indicates that the action of $\mathcal{O}_{2,\mathrm{TD}}$ could be the same as the contact operator,

$$\mathcal{O}_{\mathrm{contact}} = \frac{1}{8\Lambda^2}\nabla_\mu(\phi^2)\cdot\nabla^\mu(\phi^2). \tag{D.13}$$

The temporal component of the above operator gives the same contribution as $\mathcal{O}_{2,\mathrm{TD}}$ in Fig. D.2.

D.2 IBP in Inflationary Spacetime: Examples

In this appendix, we perform some explicit checks of IBP in inflationary spacetimes, taking into account the necessary boundary terms. We first discuss massless fields (both free and interacting) and then massive fields.

D.2.1 Massless Scalars

Free Theory

A massless scalar in dS can be expanded in terms of the mode functions as,

$$\phi(\eta,\mathbf{k}) = f_k(\eta)a^\dagger_{\mathbf{k}} + \bar{f}_k(\eta)a_{-\mathbf{k}}, \tag{D.14}$$

with

$$f_k(\eta) = \frac{(1-ik\eta)e^{ik\eta}}{\sqrt{2k^3}},\ \bar{f}_k(\eta) = \frac{(1+ik\eta)e^{-ik\eta}}{\sqrt{2k^3}}. \tag{D.15}$$

The free theory action, upon using the EOM $\Box\phi = 0$, reduces to:

$$\int \mathrm{d}^4x\sqrt{|g|}(\nabla_\mu\phi)^2 \stackrel{!}{=} \int \mathrm{d}^4x\sqrt{|g|}\nabla_\mu(\phi\nabla^\mu\phi). \tag{D.16}$$

The $\stackrel{!}{=}$ indicates that we are interested in knowing whether both sides of the above equation would give the same correlation function or not. To that end, we evaluate both sides on-shell, using the above mode functions of massless fields in dS. The LHS gives

$$\int \frac{\mathrm{d}\eta\mathrm{d}^3x}{\eta^2}\left[-(\partial_\eta\phi)^2 + (\partial_i\phi)^2\right] = \int_{-\infty}^{\eta_0}\frac{\mathrm{d}\eta}{\eta^2}\left[-(k^2\eta)^2e^{2ik\eta} + k^2(1-ik\eta)^2e^{2ik\eta}\right] \tag{D.17}$$

After performing the integrals we arrive at

$$\int_{-\infty}^{\eta_0}\frac{\mathrm{d}\eta}{\eta^2}\left[-(\partial_\eta\phi)^2 + (\partial_i\phi)^2\right] = -\frac{k^2}{\eta_0} - ik^3. \tag{D.18}$$

The RHS of (D.16) is a boundary term. We can use Stokes' theorem to write it as,

$$\int \mathrm{d}^4x\sqrt{|g|}\nabla_\mu(\phi\nabla^\mu\phi) = \int \mathrm{d}^3x\sqrt{|\gamma|}n_\mu(\phi\nabla^\mu\phi). \tag{D.19}$$

The induced metric on the boundary time slice is denoted by γ, with $\sqrt{|\gamma|} = 1/\eta_0^3$, and n_μ is a unit normal vector $n_\mu = (1/\eta_0, 0, 0, 0)$. This can then be evaluated as $\eta_0 \to 0$,

$$-\frac{1}{\eta_0^2}\phi\partial_\eta\phi\Big|_{\eta_0} = -\frac{k^2}{\eta_0} - ik^3. \tag{D.20}$$

This matches exactly with the LHS contribution. Note the first term is naively divergent as $\eta_0 \to 0$. However it does not contribute to correlation functions, which in this context are the power spectrum. This is because, to evaluate the power spectrum, we need to multiply the above by $(-i)$ (stemming from the $(-i)$ in $\exp(-i\int \mathrm{d}t\mathbb{H})$) and sum with its conjugate. The $1/\eta_0$ piece then would cancel. The remaining term would contribute to the power spectrum as expected.

Interactions

As an example, we consider a massless scalar field ϕ in dS with an interaction term

$$\int \mathrm{d}^4x\sqrt{|g|}\nabla_\mu(\phi^2)\nabla^\mu\phi. \tag{D.21}$$

By IBP we can reduce this as,

$$\int \mathrm{d}^4x\sqrt{|g|}\nabla_\mu(\phi^2)\nabla^\mu\phi \stackrel{!}{=} \int \mathrm{d}^4x\sqrt{|g|}\nabla_\mu(\phi^2\nabla^\mu\phi) = \int \mathrm{d}^3x\sqrt{|\gamma|}n_\mu\phi^2\nabla^\mu\phi. \tag{D.22}$$

In the last relation, we have used Stokes' theorem. We have assumed ϕ vanishes at spatial infinity and dropped contribution from boundary terms at spatial infinity, keeping only the contribution from a late time slice at $\eta = \eta_0 \to 0$, as above. We can rewrite the boundary term as,

$$\int \mathrm{d}^3x\sqrt{|\gamma|}n_\mu\phi^2\nabla^\mu\phi = \int \mathrm{d}^3x\left(-\frac{1}{\eta_0^2}\right)\phi^2\partial_\eta\phi\Big|_{\eta_0}. \tag{D.23}$$

We now compute the contact three-point interaction mediated by this operator and check the relation (D.22).

LHS. Here we will not write $\int \mathrm{d}^3x$ explicitly to keep the notation simple, and only track time integrals. Then we can write the LHS of (D.22) as,

$$\int_{-\infty}^{\eta_0}\frac{\mathrm{d}\eta}{\eta^2}\left(-2\phi\partial_\eta\phi\partial_\eta\phi + 2\phi\partial_i\phi\partial_i\phi\right) \tag{D.24}$$

Upon using the mode functions this becomes,

$$\int_{-\infty}^{\eta_0}\frac{\mathrm{d}\eta}{\eta^2}\Big(\big[(-2)(1-ik_1\eta)k_2^2\eta k_3^2\eta + \text{perms.}\big] \tag{D.25}$$

$$+\left[(2)(1-ik_1\eta)(1-ik_2\eta)(1-ik_3\eta)(-\mathbf{k}_2\cdot\mathbf{k}_3) + \text{perms.}\right]\Big)\,e^{ik_t\eta} \tag{D.26}$$

Here we have schematically included other momentum permutations and denoted $k_t = k_1 + k_2 + k_3$. Using momentum conservation $\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3 = 0$, the piece involving spatial

derivatives can be simplified as,

$$\int_{-\infty}^{\eta_0} \frac{\mathrm{d}\eta}{\eta^2} \left[(1-ik_1\eta)(1-ik_2\eta)(1-ik_3\eta)(k_1^2+k_2^2+k_3^2)\right] e^{ik_t\eta}. \tag{D.27}$$

This can be evaluated as,

$$(+i)(k_1^2+k_2^2+k_3^2)\left[\frac{i}{\eta_0} - k_t + \frac{k_1k_2+k_1k_3+k_2k_3}{k_t} + \frac{k_1k_2k_3}{k_t^2}\right]. \tag{D.28}$$

The part involving time derivatives can be evaluated as,

$$(+i)(2k_2^2k_3^2)\left[\frac{1}{k_t} + \frac{k_1}{k_t^2}\right] + \text{perms.}. \tag{D.29}$$

Summing the spatial and temporal contributions we have,

$$\int \frac{d\eta}{\eta^2}\left(-2\phi\partial_\eta\phi\partial_\eta\phi + 2\phi\partial_i\phi\partial_i\phi\right) = -\frac{(k_1^2+k_2^2+k_3^2)}{\eta_0} + (-i)(k_1^3+k_2^3+k_3^3). \tag{D.30}$$

RHS. The RHS of (D.22) can be evaluated as (we do not write $\int \mathrm{d}^3x$ for brevity),

$$\left(-\frac{1}{\eta_0^2}\right)(1-ik_1\eta_0)(1-ik_2\eta_0)k_3^2\eta_0 e^{ik_t\eta_0} + \text{perms.}. \tag{D.31}$$

Taking $\eta_0 \to 0$ limit,

$$\left(-\frac{1}{\eta_0}\right)k_3^2(1+ik_3\eta_0) + \text{perms.} = -\frac{(k_1^2+k_2^2+k_3^2)}{\eta_0} + (-i)(k_1^3+k_2^3+k_3^3). \tag{D.32}$$

This matches exactly with (D.30). Note the $1/\eta_0$ piece does not contribute to correlation functions for reasons identical to the free theory case discussed above.

D.2.2 Massive Scalars

Free Theory

We will start with a simple case, the kinetic term

$$\int \mathrm{d}^4x\sqrt{|g|}(\nabla_\mu\sigma)^2 \tag{D.33}$$

for a massive scalar field σ. The mode functions for a massive scalar in dS are given in (D.9). We then check whether

$$\int \mathrm{d}^4x\sqrt{|g|}(\nabla_\mu\sigma)^2 \stackrel{!}{=} \int \mathrm{d}^3x\sqrt{|\gamma|}n_\mu\sigma\nabla^\mu\sigma\Big|_{\eta_0} - \int \mathrm{d}^4x\sqrt{|g|}\sigma\Box\sigma, \tag{D.34}$$

where $\Box\sigma \equiv \nabla_\mu\nabla^\mu\sigma$. To that end, instead of using the explicit forms of the Hankel functions, we will include them schematically and that will be sufficient for our purpose. We will also check the equality (D.34) by analyzing the spatial and the temporal components separately. The spatial part, again omitting $\int \mathrm{d}^3x$ and tracking time integrals, is given by

$$\int \frac{\mathrm{d}\eta}{\eta^2}(\partial_i f_{k_1}\partial_i f_{k_2}) \stackrel{!}{=} -\int \frac{\mathrm{d}\eta}{\eta^2} f_{k_1}\partial_i^2 f_{k_2}, \tag{D.35}$$

Inserting the momentum factors, this becomes

$$\int \frac{\mathrm{d}\eta}{\eta^2}(\mathbf{k_1}\cdot\mathbf{k_2}) f_{k_1}f_{k_2} \stackrel{!}{=} -\int \frac{\mathrm{d}\eta}{\eta^2} k_1^2 f_{k_1}f_{k_2}. \tag{D.36}$$

The two contributions are equal since momentum conservation forces $\mathbf{k}_1 + \mathbf{k}_2 = 0$. The temporal part is given by,

$$-\int \frac{\mathrm{d}\eta}{\eta^2}\left(\partial_\eta f_{k_1}\partial_\eta f_{k_2}\right) \stackrel{!}{=} \left(-\frac{1}{\eta_0^2}\right) f_{k_1}\partial_\eta f_{k_2}\Big|_{\eta_0} + \int \frac{\mathrm{d}\eta}{\eta^2} f_{k_1}\nabla_\eta^2 f_{k_2}, \tag{D.37}$$

where $\nabla_\eta^2 = \partial_\eta^2 - \frac{2}{\eta}\partial_\eta$. Momentum conservation forces $k_1 = k_2$, and by doing an explicit temporal integration-by-parts, we can see the above equality indeed holds.

Interacting Theory

Next, we consider the interaction term

$$\int \mathrm{d}^4x\sqrt{|g|}\nabla_\mu(\phi^2)\nabla^\mu\sigma \tag{D.38}$$

We then follow the same procedure as used in the previous example. To check the validity of IBP for the case of massive scalars in dS, we consider whether

$$\int \mathrm{d}^4x\sqrt{|g|}\nabla_\mu(\phi^2)\nabla^\mu\sigma \stackrel{!}{=} \int \mathrm{d}^4x\sqrt{|g|}\nabla_\mu(\phi^2\nabla^\mu\sigma) - \int \mathrm{d}^4x\sqrt{|g|}\phi^2\Box\sigma$$
$$= \int \mathrm{d}^3x\sqrt{|\gamma|}n_\mu\phi^2\nabla^\mu\sigma\Big|_{\eta_0} - \int \mathrm{d}^4x\sqrt{|g|}\phi^2\Box\sigma. \quad \text{(D.39)}$$

Upon rewriting the boundary term, this condition becomes

$$\int \mathrm{d}^4x\sqrt{|g|}\partial_\mu(\phi^2)\partial^\mu\sigma \stackrel{?}{=} \int \mathrm{d}^3x\left(-\frac{1}{\eta_0^2}\right)\phi^2\partial_\eta\sigma\Big|_{\eta_0} - \int \mathrm{d}^4x\sqrt{|g|}\phi^2\Box\sigma. \quad \text{(D.40)}$$

We first compute the spatial part, again neglecting to write $\int \mathrm{d}^3x$ for brevity. On the LHS, this is

$$\int \frac{\mathrm{d}\eta}{\eta^2}\partial_i(\phi^2)\partial_i\sigma \;=\; -\int \frac{\mathrm{d}\eta}{\eta^2}(\mathbf{k}_1\cdot\mathbf{k}_3 + \mathbf{k}_2\cdot\mathbf{k}_3)f_{k_1}f_{k_2}g_{k_3} \;=\; \int \frac{\mathrm{d}\eta}{\eta^2}k_3^2 f_{k_1}f_{k_2}g_{k_3}, \quad \text{(D.41)}$$

where we have used momentum conservation $\mathbf{k}_1+\mathbf{k}_3+\mathbf{k}_3=0$. We have also considered a particular permutation of momenta, as this will be sufficient for our purpose. On the RHS, we have

$$-\int \frac{\mathrm{d}\eta}{\eta^2}\phi^2\partial_i^2\sigma = \int \frac{\mathrm{d}\eta}{\eta^2}k_3^2 f_{k_1}f_{k_2}g_{k_3}, \quad \text{(D.42)}$$

which matches with the LHS, as expected.

Now, we consider the temporal component. Starting again with the LHS and following a similar procedure as shown with the free theory, we have

$$\begin{aligned} &-\int \frac{\mathrm{d}\eta}{\eta^2}\partial_\eta(f_{k_1}f_{k_2})\partial_\eta g_{k_3} \\ &= \int \mathrm{d}\eta\partial_\eta\left(-\frac{1}{\eta^2}f_{k_1}f_{k_2}\partial_\eta g_{k_3}\right) + \int \mathrm{d}\eta f_{k_1}f_{k_2}\partial_\eta\left(\frac{1}{\eta^2}\partial_\eta g_{k_3}\right) \\ &= -\frac{1}{\eta_0^2}f_{k_1}f_{k_2}\partial_\eta g_{k_3}\Big|_{\eta_0} + \int \mathrm{d}\eta f_{k_1}f_{k_2}\left(-\frac{2}{\eta^3}\partial_\eta g_{k_3} + \frac{1}{\eta^2}\partial_\eta^2 g_{k_3}\right) \end{aligned} \quad \text{(D.43)}$$

We identify the first term as the boundary term and the second as the temporal component of $\Box\sigma$.

In the above two examples — a free theory and an interacting theory — we have seen that IBP indeed holds for massive scalars in dS for contact diagrams.

D.3 Dimension 5 Field Redefinition

The fact that $\mathcal{O}_{5,2}$ does not contribute to cosmological correlation functions can also be seen by performing a field redefinition. In particular, we can redefine:

$$A_\mu \to A_\mu - \frac{\nabla_\mu \phi}{g_A \Lambda}. \tag{D.44}$$

Since we are not interested in late-time correlation functions of A_μ, i.e., A_μ does not appear in the external lines, the above field redefinition does not modify inflaton correlators via contact diagrams. It does have an effect on the other vertices. While the kinetic terms for the inflaton and the gauge boson are unmodified by (D.44),

$$\begin{aligned} |D_\mu \mathcal{H}|^2 &\to |D_\mu \mathcal{H}|^2 + \frac{i}{\Lambda}\nabla^\mu \phi \mathcal{H}^\dagger \nabla_\mu \mathcal{H} - \frac{i}{\Lambda}\nabla^\mu \phi \mathcal{H} \nabla_\mu \mathcal{H}^\dagger - \frac{2g}{\Lambda} A^\mu \nabla_\mu \phi \mathcal{H}^\dagger \mathcal{H} + \frac{1}{\Lambda^2}(\nabla_\mu \phi)^2 \mathcal{H}^\dagger \mathcal{H} \\ &= |D_\mu \mathcal{H}|^2 - \mathcal{O}_{5,2} + \frac{1}{\Lambda^2}(\nabla_\mu \phi)^2 \mathcal{H}^\dagger \mathcal{H}. \end{aligned} \tag{D.45}$$

Therefore, the field redefinition (D.44) eliminates $\mathcal{O}_{5,2}$ and gives a correction to $\mathcal{O}_{6,1}$.

Bibliography

[1] A. Zee, *Group Theory in a Nutshell for Physicists*. Princeton University Press, USA, 3, 2016.

[2] H. Georgi, *Lie algebras in particle physics*, vol. 54. Perseus Books, Reading, MA, 2nd ed. ed., 1999.

[3] D. Tong, *Gauge theory, Lecture notes, DAMTP Cambridge* **10** (2018) 8.

[4] M. Srednicki, *Quantum field theory*. Cambridge University Press, 1, 2007.

[5] M. E. Peskin and D. V. Schroeder, *An Introduction to quantum field theory*. Addison-Wesley, Reading, USA, 1995.

[6] **Planck** Collaboration, Y. Akrami *et. al.*, *Planck 2018 results. IV. Diffuse component separation*, *Astron. Astrophys.* **641** (2020) A4, [arXiv:1807.0620].

[7] **Planck** Collaboration, N. Aghanim *et. al.*, *Planck 2018 results. V. CMB power spectra and likelihoods*, *Astron. Astrophys.* **641** (2020) A5, [arXiv:1907.1287].

[8] A. H. Guth, *The Inflationary Universe: A Possible Solution to the Horizon and Flatness Problems*, *Phys. Rev. D* **23** (1981) 347–356.

[9] M. Tristram *et. al.*, *Improved limits on the tensor-to-scalar ratio using BICEP and Planck data*, *Phys. Rev. D* **105** (2022), no. 8 083524, [arXiv:2112.0796].

[10] N. Craig, *Naturalness: past, present, and future*, *Eur. Phys. J. C* **83** (2023), no. 9 825, [arXiv:2205.0570].

[11] A. Hook, *TASI Lectures on the Strong CP Problem and Axions*, *PoS* **TASI2018** (2019) 004, [arXiv:1812.0266].

[12] S. Weinberg, *Anthropic Bound on the Cosmological Constant*, *Phys. Rev. Lett.* **59** (1987) 2607.

[13] M. Lisanti, *Lectures on Dark Matter Physics*, in *Theoretical Advanced Study Institute in Elementary Particle Physics: New Frontiers in Fields and Strings*, pp. 399–446, 2017. arXiv:1603.0379.

[14] T. Lin, *Dark matter models and direct detection*, *PoS* **333** (2019) 009, [arXiv:1904.0791].

[15] D. Baumann, *Inflation*, in *Theoretical Advanced Study Institute in Elementary Particle Physics: Physics of the Large and the Small*, pp. 523–686, 2011. arXiv:0907.5424.

[16] **Planck** Collaboration, Y. Akrami *et. al.*, *Planck 2018 results. IX. Constraints on primordial non-Gaussianity*, *Astron. Astrophys.* **641** (2020) A9, [arXiv:1905.0569].

[17] N. Arkani-Hamed and J. Maldacena, *Cosmological Collider Physics*, arXiv:1503.0804.

[18] **nEDM** Collaboration, C. Abel *et. al.*, *Measurement of the permanent electric dipole moment of the neutron*, *Phys. Rev. Lett.* **124** (2020), no. 8 081803, [arXiv:2001.1196].

[19] B. Graner, Y. Chen, E. Lindahl, and B. Heckel, *Reduced Limit on the Permanent Electric Dipole Moment of Hg199*, *Phys. Rev. Lett.* **116** (2016), no. 16 161601, [arXiv:1601.0433]. [Erratum: Phys.Rev.Lett. 119, 119901 (2017)].

[20] N. Kaloper and J. Terning, *Landscaping the Strong CP Problem*, *JHEP* **03** (2019) 032, [arXiv:1710.0174].

[21] K. Babu and R. N. Mohapatra, *CP Violation in Seesaw Models of Quark Masses*, *Phys. Rev. Lett.* **62** (1989) 1079.

[22] K. S. Babu and R. N. Mohapatra, *A Solution to the Strong CP Problem Without an Axion*, *Phys. Rev. D* **41** (1990) 1286.

[23] S. M. Barr, D. Chang, and G. Senjanovic, *Strong CP problem and parity*, *Phys. Rev. Lett.* **67** (1991) 2765–2768.

[24] R. N. Mohapatra and G. Senjanovic, *Natural Suppression of Strong p and t Noninvariance*, *Phys. Lett. B* **79** (1978) 283–286.

[25] A. E. Nelson, *Naturally Weak CP Violation*, *Phys. Lett. B* **136** (1984) 387–391.

[26] S. M. Barr, *Solving the Strong CP Problem Without the Peccei-Quinn Symmetry*, *Phys. Rev. Lett.* **53** (1984) 329.

[27] S. Chakdar, K. Ghosh, S. Nandi, and S. K. Rai, *Collider signatures of mirror fermions in the framework of a left-right mirror model*, *Phys. Rev. D* **88** (2013), no. 9 095005, [arXiv:1305.2641].

[28] R. T. D'Agnolo and A. Hook, *Finding the Strong CP problem at the LHC*, *Phys. Lett. B* **762** (2016) 421–425, [arXiv:1507.0033].

[29] L. J. Hall and K. Harigaya, *Implications of Higgs Discovery for the Strong CP Problem and Unification*, *JHEP* **10** (2018) 130, [arXiv:1803.0811].

[30] M. Dine, R. G. Leigh, and D. A. MacIntire, *Of CP and other gauge symmetries in string theory*, *Phys. Rev. Lett.* **69** (1992) 2030–2032, [hep-th/9205011].

[31] K.-w. Choi, D. B. Kaplan, and A. E. Nelson, *Is CP a gauge symmetry?*, *Nucl. Phys. B* **391** (1993) 515–530, [hep-ph/9205202].

[32] **ATLAS** Collaboration, M. Aaboud *et. al.*, *Combination of the searches for pair-produced vector-like partners of the third-generation quarks at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, *Phys. Rev. Lett.* **121** (2018), no. 21 211801, [arXiv:1808.0234].

[33] **CMS** Collaboration, A. M. Sirunyan *et. al.*, *Search for vector-like T and B quark pairs in final states with leptons at* $\sqrt{s} = 13$ *TeV*, *JHEP* **08** (2018) 177, [arXiv:1805.0475].

[34] **ATLAS** Collaboration, G. Aad *et. al.*, *Search for high-mass dilepton resonances using 139* fb^{-1} *of pp collision data collected at* $\sqrt{s}$ *=13 TeV with the ATLAS detector*, *Phys. Lett. B* **796** (2019) 68–87, [arXiv:1903.0624].

[35] **ATLAS** Collaboration, G. Aad *et. al.*, *Search for a heavy charged boson in events with a charged lepton and missing transverse momentum from pp collisions at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, *Phys. Rev. D* **100** (2019), no. 5 052013, [arXiv:1906.0560].

[36] C. Helsens and M. Selvaggi, *Search for high-mass resonances at FCC-hh*, Tech. Rep. CERN-ACC-2019-0028, CERN, Geneva, Oct, 2018.

[37] **FCC** Collaboration, A. Abada *et. al.*, *FCC Physics Opportunities: Future Circular Collider Conceptual Design Report Volume 1*, *Eur. Phys. J. C* **79** (2019), no. 6 474.

[38] Y. Zeldovich, *A New Type of Radioactive Decay: Gravitational Annihilation of Baryons*, *Phys. Lett. A* **59** (1976) 254.

[39] Y. Zeldovich, *A Novel Type of Radioactive Decay: Gravitational Baryon Annihilation*, *Zh. Eksp. Teor. Fiz.* **72** (1977) 18–21.

[40] T. Banks and L. J. Dixon, *Constraints on String Vacua with Space-Time Supersymmetry*, *Nucl. Phys. B* **307** (1988) 93–108.

[41] S. B. Giddings and A. Strominger, *Axion Induced Topology Change in Quantum Gravity and String Theory*, *Nucl. Phys. B* **306** (1988) 890–907.

[42] K.-M. Lee, *Wormholes and Goldstone Bosons*, *Phys. Rev. Lett.* **61** (1988) 263–266.

[43] L. Abbott and M. B. Wise, *Wormholes and Global Symmetries*, *Nucl. Phys. B* **325** (1989) 687–704.

[44] S. R. Coleman and K.-M. Lee, *WORMHOLES MADE WITHOUT MASSLESS MATTER FIELDS*, *Nucl. Phys. B* **329** (1990) 387–409.

[45] R. Kallosh, A. D. Linde, D. A. Linde, and L. Susskind, *Gravity and global symmetries*, *Phys. Rev. D* **52** (1995) 912–935, [hep-th/9502069].

[46] T. Banks and N. Seiberg, *Symmetries and Strings in Field Theory and Gravity*, *Phys. Rev. D* **83** (2011) 084019, [arXiv:1011.5120].

[47] D. Harlow and H. Ooguri, *Symmetries in quantum field theory and quantum gravity*, arXiv:1810.0533.

[48] D. Harlow and H. Ooguri, *Constraints on Symmetries from Holography*, *Phys. Rev. Lett.* **122** (2019), no. 19 191601, [arXiv:1810.0533].

[49] S. Fichet and P. Saraswat, *Approximate Symmetries and Gravity*, *JHEP* **01** (2020) 088, [arXiv:1909.0200].

[50] T. Daus, A. Hebecker, S. Leonhardt, and J. March-Russell, *Towards a Swampland Global Symmetry Conjecture using weak gravity*, *Nucl. Phys. B* **960** (2020) 115167, [arXiv:2002.0245].

[51] R. Peccei and H. R. Quinn, *CP Conservation in the Presence of Instantons*, *Phys. Rev. Lett.* **38** (1977) 1440–1443.

[52] R. Peccei and H. R. Quinn, *Constraints Imposed by CP Conservation in the Presence of Instantons*, *Phys. Rev. D* **16** (1977) 1791–1797.

[53] F. Wilczek, *Problem of Strong P and T Invariance in the Presence of Instantons*, *Phys. Rev. Lett.* **40** (1978) 279–282.

[54] S. Weinberg, *A New Light Boson?*, *Phys. Rev. Lett.* **40** (1978) 223–226.

[55] J. E. Kim, *Weak Interaction Singlet and Strong CP Invariance*, *Phys. Rev. Lett.* **43** (1979) 103.

[56] M. A. Shifman, A. Vainshtein, and V. I. Zakharov, *Can Confinement Ensure Natural CP Invariance of Strong Interactions?*, *Nucl. Phys. B* **166** (1980) 493–506.

[57] M. Dine, W. Fischler, and M. Srednicki, *A Simple Solution to the Strong CP Problem with a Harmless Axion*, *Phys. Lett. B* **104** (1981) 199–202.

[58] S. M. Barr and D. Seckel, *Planck scale corrections to axion models*, *Phys. Rev. D* **46** (1992) 539–549.

[59] M. Kamionkowski and J. March-Russell, *Planck scale physics and the Peccei-Quinn mechanism*, *Phys. Lett. B* **282** (1992) 137–141, [hep-th/9202003].

[60] R. Holman, S. D. Hsu, T. W. Kephart, E. W. Kolb, R. Watkins, and L. M. Widrow, *Solutions to the strong CP problem in a world with gravity*, *Phys. Lett. B* **282** (1992) 132–136, [hep-ph/9203206].

[61] S. Ghigna, M. Lusignoli, and M. Roncadelli, *Instability of the invisible axion*, *Phys. Lett. B* **283** (1992) 278–281.

[62] E. Chun and A. Lukas, *Discrete gauge symmetries in axionic extensions of the SSM*, *Phys. Lett. B* **297** (1992) 298–304, [hep-ph/9209208].

[63] L. Randall, *Composite axion models and Planck scale physics*, *Phys. Lett. B* **284** (1992) 77–80.

[64] H.-C. Cheng and D. E. Kaplan, *Axions and a gauged Peccei-Quinn symmetry*, hep-ph/0103346.

[65] A. Arvanitaki, S. Dimopoulos, S. Dubovsky, N. Kaloper, and J. March-Russell, *String Axiverse*, *Phys. Rev. D* **81** (2010) 123530, [arXiv:0905.4720].

[66] H. Fukuda, M. Ibe, M. Suzuki, and T. T. Yanagida, *A "gauged" $U(1)$ Peccei–Quinn symmetry*, *Phys. Lett. B* **771** (2017) 327–331, [arXiv:1703.0111].

[67] L. Di Luzio, E. Nardi, and L. Ubaldi, *Accidental Peccei-Quinn symmetry protected to arbitrary order*, *Phys. Rev. Lett.* **119** (2017), no. 1 011801, [arXiv:1704.0112].

[68] B. Lillard and T. M. P. Tait, *A Composite Axion from a Supersymmetric Product Group*, *JHEP* **11** (2017) 005, [arXiv:1707.0426].

[69] B. Lillard and T. M. Tait, *A High Quality Composite Axion*, *JHEP* **11** (2018) 199, [arXiv:1811.0308].

[70] G. Senjanovic and R. N. Mohapatra, *Exact Left-Right Symmetry and Spontaneous Violation of Parity*, *Phys. Rev. D* **12** (1975) 1502.

[71] J. R. Ellis and M. K. Gaillard, *Strong and Weak CP Violation*, *Nucl. Phys. B* **150** (1979) 141–162.

[72] Z. Chacko, H.-S. Goh, and R. Harnik, *The Twin Higgs: Natural electroweak breaking from mirror symmetry*, *Phys. Rev. Lett.* **96** (2006) 231802, [hep-ph/0506256].

[73] Z. Chacko, H.-S. Goh, and R. Harnik, *A Twin Higgs model from left-right symmetry*, *JHEP* **01** (2006) 108, [hep-ph/0512088].

[74] A. Albaid, M. Dine, and P. Draper, *Strong CP and SUZ_2*, *JHEP* **12** (2015) 046, [arXiv:1510.0339].

[75] A. Davidson and K. C. Wali, *Universal Seesaw Mechanism?*, *Phys. Rev. Lett.* **59** (1987) 393.

[76] A. Davidson and K. C. Wali, *Family Mass Hierarchy From Universal Seesaw Mechanism*, *Phys. Rev. Lett.* **60** (1988) 1813.

[77] A. Davidson, S. Ranfone, and K. C. Wali, *Quark Masses and Mixing Angles From Universal Seesaw Mechanism*, *Phys. Rev. D* **41** (1990) 208.

[78] S. Ranfone, *The Three generation seesaw model for quarks. 2.*, *Phys. Rev. D* **42** (1990) 3819–3828.

[79] H.-S. Goh and S. Su, *Phenomenology of The Left-Right Twin Higgs Model*, *Phys. Rev. D* **75** (2007) 075010, [hep-ph/0611015].

[80] Y. Kiyo, T. Morozumi, P. Parada, M. Rebelo, and M. Tanimoto, *Quark mass hierarchy, FCNC and CP violation in a seesaw model*, *Prog. Theor. Phys.* **101** (1999) 671–706, [hep-ph/9809333].

[81] A. Maiezza and M. Nemevšek, *Strong P invariance, neutron electric dipole moment, and minimal left-right parity at LHC*, *Phys. Rev. D* **90** (2014), no. 9 095002, [arXiv:1407.3678].

[82] G. Senjanovic and V. Tello, *Strong CP violation: problem or blessing?*, arXiv:2004.0403.

[83] A. Khodjamirian, T. Mannel, A. Pivovarov, and Y.-M. Wang, *Charm-loop effect in $B \to K^{(*)}\ell^+\ell^-$ and $B \to K^*\gamma$*, *JHEP* **09** (2010) 089, [arXiv:1006.4945].

[84] **Particle Data Group** Collaboration, P. Zyla *et. al.*, *Review of Particle Physics*, *PTEP* **2020** (2020), no. 8 083C01.

[85] N. Hutzler *et. al.*, *Searches for new sources of CP violation using molecules as quantum sensors*, arXiv:2010.0870.

[86] **ACME** Collaboration, V. Andreev *et. al.*, *Improved limit on the electric dipole moment of the electron*, *Nature* **562** (2018), no. 7727 355–360.

[87] Z. G. Berezhiani, R. N. Mohapatra, and G. Senjanovic, *Planck scale physics and solutions to the strong CP problem without axion*, *Phys. Rev. D* **47** (1993) 5565–5570, [hep-ph/9212318].

[88] T. Kibble, *Topology of Cosmic Domains and Strings*, *J. Phys. A* **9** (1976) 1387–1398.

[89] G. R. Dvali and G. Senjanovic, *Is there a domain wall problem?*, *Phys. Rev. Lett.* **74** (1995) 5178–5181, [hep-ph/9501387].

[90] G. R. Dvali, A. Melfo, and G. Senjanovic, *Nonrestoration of spontaneously broken P and CP at high temperature*, *Phys. Rev. D* **54** (1996) 7857–7866, [hep-ph/9601376].

[91] Y. Zeldovich, I. Kobzarev, and L. Okun, *Cosmological Consequences of the Spontaneous Breakdown of Discrete Symmetry*, *Zh. Eksp. Teor. Fiz.* **67** (1974) 3–11.

[92] B. Rai and G. Senjanovic, *Gravity and domain wall problem*, *Phys. Rev. D* **49** (1994) 2729–2733, [hep-ph/9301240].

[93] W. H. Press, B. S. Ryden, and D. N. Spergel, *Dynamical Evolution of Domain Walls in an Expanding Universe*, *Astrophys. J.* **347** (1989) 590–604.

[94] T. Garagounis and M. Hindmarsh, *Scaling in numerical simulations of domain walls*, *Phys. Rev. D* **68** (2003) 103506, [hep-ph/0212359].

[95] J. Oliveira, C. Martins, and P. Avelino, *The Cosmological evolution of domain wall networks*, *Phys. Rev. D* **71** (2005) 083509, [hep-ph/0410356].

[96] P. Avelino, C. Martins, and J. Oliveira, *One-scale model for domain wall network evolution*, *Phys. Rev. D* **72** (2005) 083506, [hep-ph/0507272].

[97] A. Leite and C. Martins, *Scaling Properties of Domain Wall Networks*, *Phys. Rev. D* **84** (2011) 103523, [arXiv:1110.3486].

[98] M. Hindmarsh, *Analytic scaling solutions for cosmic domain walls*, *Phys. Rev. Lett.* **77** (1996) 4495–4498, [hep-ph/9605332].

[99] M. Hindmarsh, *Level set method for the evolution of defect and brane networks*, *Phys. Rev. D* **68** (2003) 043510, [hep-ph/0207267].

[100] A. Vilenkin, *Gravitational Field of Vacuum Domain Walls and Strings*, *Phys. Rev. D* **23** (1981) 852–857.

[101] T. Vachaspati, A. E. Everett, and A. Vilenkin, *Radiation From Vacuum Strings and Domain Walls*, *Phys. Rev. D* **30** (1984) 2046.

[102] M. Gleiser and R. Roberts, *Gravitational waves from collapsing vacuum domains*, *Phys. Rev. Lett.* **81** (1998) 5497–5500, [astro-ph/9807260].

[103] J. F. Dufaux, A. Bergman, G. N. Felder, L. Kofman, and J.-P. Uzan, *Theory and Numerics of Gravitational Waves from Preheating after Inflation*, *Phys. Rev. D* **76** (2007) 123517, [arXiv:0707.0875].

[104] T. Hiramatsu, M. Kawasaki, and K. Saikawa, *Gravitational Waves from Collapsing Domain Walls*, *JCAP* **05** (2010) 032, [arXiv:1002.1555].

[105] M. Kawasaki and K. Saikawa, *Study of gravitational radiation from cosmic domain walls*, *JCAP* **09** (2011) 008, [arXiv:1102.5628].

[106] K. Saikawa, *A review of gravitational waves from cosmic domain walls*, *Universe* **3** (2017), no. 2 40, [arXiv:1703.0257].

[107] G. Janssen *et. al.*, *Gravitational wave astronomy with the SKA*, *PoS* **AASKA14** (2015) 037, [arXiv:1501.0012].

[108] **NANOGRAV** Collaboration, Z. Arzoumanian *et. al.*, *The NANOGrav 11-year Data Set: Pulsar-timing Constraints On The Stochastic Gravitational-wave Background*, *Astrophys. J.* **859** (2018), no. 1 47, [arXiv:1801.0261].

[109] L. Lentati *et. al.*, *European Pulsar Timing Array Limits On An Isotropic Stochastic Gravitational-Wave Background*, *Mon. Not. Roy. Astron. Soc.* **453** (2015), no. 3 2576–2598, [arXiv:1504.0369].

[110] **LISA** Collaboration, P. Amaro-Seoane *et. al.*, *Laser Interferometer Space Antenna*, arXiv:1702.0078.

[111] N. Seto, S. Kawamura, and T. Nakamura, *Possibility of direct measurement of the acceleration of the universe using 0.1-Hz band laser interferometer gravitational wave antenna in space*, *Phys. Rev. Lett.* **87** (2001) 221103, [astro-ph/0108011].

[112] J. Preskill and A. Vilenkin, *Decay of metastable topological defects*, *Phys. Rev. D* **47** (1993) 2324–2342, [hep-ph/9209210].

[113] **Planck** Collaboration, P. A. R. Ade *et. al.*, *Planck 2015 results. XIII. Cosmological parameters*, *Astron. Astrophys.* **594** (2016) A13, [arXiv:1502.0158].

[114] J. M. Pendlebury *et. al.*, *Revised experimental upper limit on the electric dipole moment of the neutron*, *Phys. Rev. D* **92** (2015), no. 9 092003, [arXiv:1509.0441].

[115] G. 't Hooft, *Symmetry Breaking Through Bell-Jackiw Anomalies*, *Phys. Rev. Lett.* **37** (1976) 8–11.

[116] V. Baluni, *CP Violating Effects in QCD*, *Phys. Rev. D* **19** (1979) 2227–2230.

[117] R. J. Crewther, P. Di Vecchia, G. Veneziano, and E. Witten, *Chiral Estimate of the Electric Dipole Moment of the Neutron in Quantum Chromodynamics*, *Phys. Lett. B* **88** (1979) 123. [Erratum: Phys.Lett.B 91, 487 (1980)].

[118] M. Pospelov and A. Ritz, *Theta vacua, QCD sum rules, and the neutron electric dipole moment*, *Nucl. Phys. B* **573** (2000) 177–200, [hep-ph/9908508].

[119] D. B. Kaplan and A. V. Manohar, *Current Mass Ratios of the Light Quarks*, *Phys. Rev. Lett.* **56** (1986) 2004.

[120] T. Banks, Y. Nir, and N. Seiberg, *Missing (up) mass, accidental anomalous symmetries, and the strong CP problem*, in *2nd IFT Workshop on Yukawa Couplings and the Origins of Mass*, pp. 26–41, 2, 1994. hep-ph/9403203.

[121] A. Hook, *Anomalous solutions to the strong cp problem*, *Phys. Rev. Lett.* **114** (Apr, 2015) 141801.

[122] P. Agrawal and K. Howe, *A Flavorful Factoring of the Strong CP Problem*, *JHEP* **12** (2018) 035, [arXiv:1712.0580].

[123] **Flavour Lattice Averaging Group (FLAG)** Collaboration, Y. Aoki *et. al.*, *FLAG Review 2021*, *Eur. Phys. J. C* **82** (2022), no. 10 869, [arXiv:2111.0984].

[124] A. Valenti and L. Vecchi, *The CKM phase and $\overline{\theta}$ in Nelson-Barr models*, *JHEP* **07** (2021), no. 203 203, [arXiv:2105.0912].

[125] J. de Vries, P. Draper, and H. H. Patel, *Do Minimal Parity Solutions to the Strong CP Problem Work?*, arXiv:2109.0163.

[126] A. Valenti and L. Vecchi, *Perturbative running of the topological angles*, *JHEP* **01** (2023) 131, [arXiv:2210.0932].

[127] J. Hisano, T. Kitahara, N. Osamura, and A. Yamada, *Novel loop-diagrammatic approach to QCD θ parameter and application to the left-right model*, arXiv:2301.1340.

[128] J. Preskill, M. B. Wise, and F. Wilczek, *Cosmology of the Invisible Axion*, *Phys. Lett.* **120B** (1983) 127–132.

[129] L. F. Abbott and P. Sikivie, *A Cosmological Bound on the Invisible Axion*, *Phys. Lett.* **120B** (1983) 133–136.

[130] M. Dine and W. Fischler, *The Not So Harmless Axion*, *Phys. Lett.* **B120** (1983) 137–141. [,URL(1982)].

[131] L. Di Luzio, M. Giannotti, E. Nardi, and L. Visinelli, *The landscape of QCD axion models*, *Phys. Rept.* **870** (2020) 1–117, [arXiv:2003.0110].

[132] M. A. B. Beg and H. S. Tsao, *Strong P, T Noninvariances in a Superweak Theory*, *Phys. Rev. Lett.* **41** (1978) 278.

[133] L. Lavoura, *A New type of spontaneous CP breaking*, *Phys. Lett. B* **400** (1997) 152–156, [hep-ph/9701221].

[134] P.-H. Gu, *Mirror left-right symmetry*, *Phys. Lett. B* **713** (2012) 485–489, [arXiv:1201.3551].

[135] J. Kawamura, S. Okawa, Y. Omura, and Y. Tang, *WIMP dark matter in the parity solution to the strong CP problem*, *JHEP* **04** (2019) 162, [arXiv:1812.0700].

[136] D. Dunsky, L. J. Hall, and K. Harigaya, *Higgs Parity, Strong CP, and Dark Matter*, *JHEP* **07** (2019) 016, [arXiv:1902.0772].

[137] L. J. Hall and K. Harigaya, *Higgs Parity Grand Unification*, *JHEP* **11** (2019) 033, [arXiv:1905.1272].

[138] N. Craig, I. Garcia Garcia, G. Koszegi, and A. McCune, *P not PQ*, *JHEP* **09** (2021) 130, [arXiv:2012.1341].

[139] M. Redi and A. Tesi, *Neutrinos, Dark Matter and Higgs Vacua in Parity Solutions of the strong CP problem*, arXiv:2307.0316.

[140] J. Carrasco-Martinez, D. I. Dunsky, L. J. Hall, and K. Harigaya, *Leptogenesis in Parity Solutions to the Strong CP Problem and Standard Model Parameters*, arXiv:2307.1573.

[141] J. A. Dror, D. Dunsky, L. J. Hall, and K. Harigaya, *Sterile Neutrino Dark Matter in Left-Right Theories*, *JHEP* **07** (2020) 168, [arXiv:2004.0951].

[142] D. Dunsky, L. J. Hall, and K. Harigaya, *Sterile Neutrino Dark Matter and Leptogenesis in Left-Right Higgs Parity*, *JHEP* **01** (2021) 125, [arXiv:2007.1271].

[143] Q. Bonnefoy, L. Hall, C. A. Manzari, and C. Scherb, *A Colorful Mirror Solution to the Strong CP Problem*, arXiv:2303.0615.

[144] S. Koren and R. McGehee, *Freezing-in twin dark matter*, *Phys. Rev. D* **101** (2020), no. 5 055024, [arXiv:1908.0355].

[145] D. Dunsky, L. J. Hall, and K. Harigaya, *CHAMP Cosmic Rays*, *JCAP* **07** (2019) 015, [arXiv:1812.1111].

[146] T. P. Cheng, E. Eichten, and L.-F. Li, *Higgs Phenomena in Asymptotically Free Gauge Theories*, *Phys. Rev. D* **9** (1974) 2259.

[147] G.-y. Huang and S. Zhou, *Precise Values of Running Quark and Lepton Masses in the Standard Model*, *Phys. Rev. D* **103** (2021), no. 1 016010, [arXiv:2009.0485].

[148] **Particle Data Group** Collaboration, R. L. Workman *et. al.*, *Review of Particle Physics*, *PTEP* **2022** (2022) 083C01.

[149] T. van Ritbergen, J. A. M. Vermaseren, and S. A. Larin, *The Four loop beta function in quantum chromodynamics*, *Phys. Lett. B* **400** (1997) 379–384, [hep-ph/9701390].

[150] B. Grzadkowski, M. Iskrzynski, M. Misiak, and J. Rosiek, *Dimension-Six Terms in the Standard Model Lagrangian*, *JHEP* **10** (2010) 085, [arXiv:1008.4884].

[151] B. Audren, J. Lesgourgues, G. Mangano, P. D. Serpico, and T. Tram, *Strongest model-independent bound on the lifetime of Dark Matter*, *JCAP* **12** (2014) 028, [arXiv:1407.2418].

[152] M. G. Baring, T. Ghosh, F. S. Queiroz, and K. Sinha, *New Limits on the Dark Matter Lifetime from Dwarf Spheroidal Galaxies using Fermi-LAT*, *Phys. Rev. D* **93** (2016), no. 10 103009, [arXiv:1510.0038].

[153] Y. Mambrini, S. Profumo, and F. S. Queiroz, *Dark Matter and Global Symmetries*, *Phys. Lett. B* **760** (2016) 807–815, [arXiv:1508.0663].

[154] T. R. Slatyer and C.-L. Wu, *General Constraints on Dark Matter Decay from the Cosmic Microwave Background*, *Phys. Rev. D* **95** (2017), no. 2 023010, [arXiv:1610.0693].

[155] A. J. Benson, A. Farahi, S. Cole, L. A. Moustakas, A. Jenkins, M. Lovell, R. Kennedy, J. Helly, and C. Frenk, *Dark matter halo merger histories beyond cold dark matter – i. methods and application to warm dark matter*, *Monthly Notices of the Royal Astronomical Society* **428** (nov, 2012) 1774–1789.

[156] M. R. Lovell, C. S. Frenk, V. R. Eke, A. Jenkins, L. Gao, and T. Theuns, *The properties of warm dark matter haloes*, *Mon. Not. Roy. Astron. Soc.* **439** (2014) 300–317, [arXiv:1308.1399].

[157] R. Kennedy, C. Frenk, S. Cole, and A. Benson, *Constraining the warm dark matter particle mass with Milky Way satellites*, *Mon. Not. Roy. Astron. Soc.* **442** (2014), no. 3 2487–2495, [arXiv:1310.7739].

[158] S. D. McDermott, H.-B. Yu, and K. M. Zurek, *Turning off the Lights: How Dark is Dark Matter?*, *Phys. Rev. D* **83** (2011) 063509, [arXiv:1011.2907].

[159] F. J. Sanchez-Salcedo, E. Martinez-Gomez, and J. Magana, *On the fraction of dark matter in charged massive particles (CHAMPs)*, *JCAP* **02** (2010) 031, [arXiv:1002.3145].

[160] P. Agrawal, F.-Y. Cyr-Racine, L. Randall, and J. Scholtz, *Make Dark Matter Charged Again*, *JCAP* **05** (2017) 022, [arXiv:1610.0461].

[161] Y. Bai and B. A. Dobrescu, *Collider Tests of the Renormalizable Coloron Model*, *JHEP* **04** (2018) 114, [arXiv:1802.0300].

[162] **CDMS-II** Collaboration, Z. Ahmed *et. al.*, *Dark Matter Search Results from the CDMS II Experiment*, *Science* **327** (2010) 1619–1621, [arXiv:0912.3592].

[163] **XENON** Collaboration, E. Aprile *et. al.*, *Dark Matter Search Results from a One Ton-Year Exposure of XENON1T*, *Phys. Rev. Lett.* **121** (2018), no. 11 111302, [arXiv:1805.1256].

[164] R. Essig, T. Volansky, and T.-T. Yu, *New Constraints and Prospects for sub-GeV Dark Matter Scattering off Electrons in Xenon*, *Phys. Rev. D* **96** (2017), no. 4 043017, [arXiv:1703.0091].

[165] **Majorana** Collaboration, S. I. Alvis *et. al.*, *First Limit on the Direct Detection of Lightly Ionizing Particles for Electric Charge as Low as e/1000 with the Majorana Demonstrator*, *Phys. Rev. Lett.* **120** (2018), no. 21 211804, [arXiv:1801.1014].

[166] **MACRO** Collaboration, M. Ambrosio *et. al.*, *Final search for lightly ionizing particles with the MACRO detector*, hep-ex/0402006.

[167] F. Kajino, S. Matsuno, T. Kitamura, T. Aoki, Y. K. Yuan, K. Mitsui, Y. Ohashi, and A. Okada, *New limit on the flux of slowly moving magnetic monopoles*, *Journal of Physics G: Nuclear Physics* **10** (apr, 1984) 447.

[168] E. N. Alekseev, M. M. Boliev, A. E. Chudakov, S. P. Mikheyev, and O. Y. Shkvorets, *SEARCH FOR SLOWLY MOVING PENETRATING PARTICLES AT BAKSAN UNDERGROUND TELESCOPE*, in *18th International Cosmic Ray Conference*, vol. HE2.2-1, pp. 52–55, 1983.

[169] W. G. Jones, P. F. Smith, G. J. Homer, J. D. Lewin, and H. E. Walford, *Searches for Fractional Electric Charge in Meteorite Samples*, *Z. Phys. C* **43** (1989) 349–355.

[170] M. C. Digman, C. V. Cappiello, J. F. Beacom, C. M. Hirata, and A. H. G. Peter, *Not as big as a barn: Upper bounds on dark matter-nucleus cross sections*, *Phys. Rev. D* **100** (2019), no. 6 063013, [arXiv:1907.1061]. [Erratum: Phys.Rev.D 106, 089902 (2022)].

[171] M. W. Goodman and E. Witten, *Detectability of Certain Dark Matter Candidates*, *Phys. Rev. D* **31** (1985) 3059.

[172] **XENON** Collaboration, E. Aprile *et. al.*, *First Dark Matter Search with Nuclear Recoils from the XENONnT Experiment*, *Phys. Rev. Lett.* **131** (2023), no. 4 041003, [arXiv:2303.1472].

[173] **LZ** Collaboration, J. Aalbers *et. al.*, *First Dark Matter Search Results from the LUX-ZEPLIN (LZ) Experiment*, *Phys. Rev. Lett.* **131** (2023), no. 4 041002, [arXiv:2207.0376].

[174] D. Dunsky, L. J. Hall, and K. Harigaya, *Dark Matter, Dark Radiation and Gravitational Waves from Mirror Higgs Parity*, *JHEP* **02** (2020) 078, [arXiv:1908.0275].

[175] Z. G. Berezhiani, A. D. Dolgov, and R. N. Mohapatra, *Asymmetric inflationary reheating and the nature of mirror universe*, *Phys. Lett. B* **375** (1996) 26–36, [hep-ph/9511221].

[176] P. Gondolo and G. Gelmini, *Cosmic abundances of stable particles: Improved analysis*, *Nuclear Physics B* **360** (1991), no. 1 145–179.

[177] G. Pancheri, J. P. Revol, and C. Rubbia, *Precise measurement of the toponium mass from the observation of its two photon decay at the LHC*, *Phys. Lett. B* **277** (1992) 518–523.

[178] V. D. Barger, E. W. N. Glover, K. Hikasa, W.-Y. Keung, M. G. Olsson, C. J. Suchyta, III, and X. R. Tata, *Superheavy Quarkonium Production and Decays: A New Higgs Signal*, *Phys. Rev. D* **35** (1987) 3366. [Erratum: Phys.Rev.D 38, 1632 (1988)].

[179] D. Kahawala and Y. Kats, *Distinguishing spins at the LHC using bound state signals*, *JHEP* **09** (2011) 099, [arXiv:1103.3503].

[180] Y. Kats and M. J. Strassler, *Probing Colored Particles with Photons, Leptons, and Jets*, *JHEP* **11** (2012) 097, [arXiv:1204.1119]. [Erratum: JHEP 07, 009 (2016)].

[181] T. Hambye, M. H. G. Tytgat, J. Vandecasteele, and L. Vanderheyden, *Dark matter from dark photons: a taxonomy of dark matter production*, *Phys. Rev. D* **100** (2019), no. 9 095018, [arXiv:1908.0986].

[182] G. Bélanger, C. Delaunay, A. Pukhov, and B. Zaldivar, *Dark matter abundance from the sequential freeze-in mechanism*, *Phys. Rev. D* **102** (2020), no. 3 035017, [arXiv:2005.0629].

[183] R. T. Co, E. Gonzalez, and K. Harigaya, *Increasing Temperature toward the Completion of Reheating*, *JCAP* **11** (2020) 038, [arXiv:2007.0432].

[184] M. Fukugita and T. Yanagida, *Baryogenesis Without Grand Unification*, *Phys. Lett. B* **174** (1986) 45–47.

[185] M. A. Luty, *Baryogenesis via leptogenesis*, *Phys. Rev. D* **45** (1992) 455–465.

[186] **Planck** Collaboration, Y. Akrami *et. al.*, *Planck 2018 results. X. Constraints on inflation*, *Astron. Astrophys.* **641** (2020) A10, [arXiv:1807.0621].

[187] K. K. Boddy *et. al.*, *Snowmass2021 theory frontier white paper: Astrophysical and cosmological probes of dark matter*, *JHEAp* **35** (2022) 112–138, [arXiv:2203.0638].

[188] G. Domènech, *Scalar Induced Gravitational Waves Review*, *Universe* **7** (2021), no. 11 398, [arXiv:2109.0139].

[189] A. M. Green and B. J. Kavanagh, *Primordial Black Holes as a dark matter candidate*, *J. Phys. G* **48** (2021), no. 4 043001, [arXiv:2007.1072].

[190] B. Carr and F. Kuhnel, *Primordial Black Holes as Dark Matter: Recent Developments*, *Ann. Rev. Nucl. Part. Sci.* **70** (2020) 355–394, [arXiv:2006.0283].

[191] P. Ivanov, P. Naselsky, and I. Novikov, *Inflation and primordial black holes as dark matter*, *Phys. Rev. D* **50** (1994) 7173–7178.

[192] J. Garcia-Bellido and E. Ruiz Morales, *Primordial black holes from single field models of inflation*, *Phys. Dark Univ.* **18** (2017) 47–54, [arXiv:1702.0390].

[193] G. Ballesteros and M. Taoso, *Primordial black hole dark matter from single field inflation*, *Phys. Rev. D* **97** (2018), no. 2 023501, [arXiv:1709.0556].

[194] N. C. Tsamis and R. P. Woodard, *Improved estimates of cosmological perturbations*, *Phys. Rev. D* **69** (2004) 084005, [astro-ph/0307463].

[195] W. H. Kinney, *Horizon crossing and inflation with large eta*, *Phys. Rev. D* **72** (2005) 023515, [gr-qc/0503017].

[196] S. Hooshangi, A. Talebian, M. H. Namjoo, and H. Firouzjahi, *Multiple field ultraslow-roll inflation: Primordial black holes from straight bulk and distorted boundary*, *Phys. Rev. D* **105** (2022), no. 8 083525, [arXiv:2201.0725].

[197] S. Kasuya and M. Kawasaki, *Axion isocurvature fluctuations with extremely blue spectrum*, *Phys. Rev. D* **80** (2009) 023516, [arXiv:0904.3800].

[198] M. Kawasaki, N. Kitajima, and T. T. Yanagida, *Primordial black hole formation from an axionlike curvaton model*, *Phys. Rev. D* **87** (2013), no. 6 063519, [arXiv:1207.2550].

[199] D. J. H. Chung and H. Yoo, *Elementary Theorems Regarding Blue Isocurvature Perturbations*, *Phys. Rev. D* **91** (2015) 083530, [arXiv:1501.0561].

[200] D. J. H. Chung and A. Upadhye, *Search for strongly blue axion isocurvature*, *Phys. Rev. D* **98** (2018), no. 2 023525, [arXiv:1711.0673].

[201] D. J. H. Chung and S. C. Tadepalli, *Analytic treatment of underdamped axionic blue isocurvature perturbations*, *Phys. Rev. D* **105** (2022), no. 12 123511, [arXiv:2110.0227].

[202] A. Talebian, A. Nassiri-Rad, and H. Firouzjahi, *Stochastic effects in axion inflation and primordial black hole formation*, *Phys. Rev. D* **105** (2022), no. 10 103516, [arXiv:2202.0206].

[203] P. W. Graham, J. Mardon, and S. Rajendran, *Vector Dark Matter from Inflationary Fluctuations*, *Phys. Rev. D* **93** (2016), no. 10 103520, [arXiv:1504.0210].

[204] A. L. Erickcek and K. Sigurdson, *Reheating Effects in the Matter Power Spectrum and Implications for Substructure*, *Phys. Rev. D* **84** (2011) 083503, [arXiv:1106.0536].

[205] J. Barir, M. Geller, C. Sun, and T. Volansky, *Gravitational Waves from Incomplete Inflationary Phase Transitions*, arXiv:2203.0069.

[206] D. J. H. Chung, E. W. Kolb, A. Riotto, and L. Senatore, *Isocurvature constraints on gravitationally produced superheavy dark matter*, *Phys. Rev. D* **72** (2005) 023511, [astro-ph/0411468].

[207] A. A. Starobinsky, *STOCHASTIC DE SITTER (INFLATIONARY) STAGE IN THE EARLY UNIVERSE*, *Lect. Notes Phys.* **246** (1986) 107–126.

[208] A. A. Starobinsky and J. Yokoyama, *Equilibrium state of a selfinteracting scalar field in the De Sitter background*, *Phys. Rev. D* **50** (1994) 6357–6368, [astro-ph/9407016].

[209] A. D. Linde and V. F. Mukhanov, *Nongaussian isocurvature perturbations from inflation*, *Phys. Rev. D* **56** (1997) R535–R539, [astro-ph/9610219].

[210] K. Enqvist and M. S. Sloth, *Adiabatic CMB perturbations in pre - big bang string cosmology*, *Nucl. Phys. B* **626** (2002) 395–409, [hep-ph/0109214].

[211] T. Moroi and T. Takahashi, *Effects of cosmological moduli fields on cosmic microwave background*, *Phys. Lett. B* **522** (2001) 215–221, [hep-ph/0110096]. [Erratum: Phys.Lett.B 539, 303–303 (2002)].

[212] D. H. Lyth and D. Wands, *Generating the curvature perturbation without an inflaton*, *Phys. Lett. B* **524** (2002) 5–14, [hep-ph/0110002].

[213] E. W. Kolb and M. S. Turner, *The Early Universe*, vol. 69. 1990.

[214] K. A. Malik and D. Wands, *Cosmological perturbations*, *Phys. Rept.* **475** (2009) 1–51, [arXiv:0809.4944].

[215] 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. D* **62** (2000) 043527, [astro-ph/0003278].

[216] M. Sasaki, Y. Nambu, and K.-i. Nakao, *Classical Behavior of a Scalar Field in the Inflationary Universe*, *Nucl. Phys. B* **308** (1988) 868–884.

[217] Y. Nambu and M. Sasaki, *Stochastic Stage of an Inflationary Universe Model*, *Phys. Lett. B* **205** (1988) 441–446.

[218] P. W. Graham and A. Scherlis, *Stochastic axion scenario*, *Phys. Rev. D* **98** (2018), no. 3 035017, [arXiv:1805.0736].

[219] T. Markkanen, A. Rajantie, S. Stopyra, and T. Tenkanen, *Scalar correlation functions in de Sitter space from the stochastic spectral expansion*, *JCAP* **08** (2019) 001, [arXiv:1904.1191].

[220] A. R. Liddle and S. M. Leach, *How long before the end of inflation were observable perturbations produced?*, *Phys. Rev. D* **68** (2003) 103503, [astro-ph/0305263].

[221] S. Dodelson and L. Hui, *A Horizon ratio bound for inflationary fluctuations*, *Phys. Rev. Lett.* **91** (2003) 131301, [astro-ph/0305113].

[222] L. F. Abbott, E. Farhi, and M. B. Wise, *Particle Production in the New Inflationary Cosmology*, *Phys. Lett. B* **117** (1982) 29.

[223] A. D. Dolgov and A. D. Linde, *Baryon Asymmetry in Inflationary Universe*, *Phys. Lett. B* **116** (1982) 329.

[224] A. Albrecht, P. J. Steinhardt, M. S. Turner, and F. Wilczek, *Reheating an Inflationary Universe*, *Phys. Rev. Lett.* **48** (1982) 1437.

[225] D. I. Podolsky, G. N. Felder, L. Kofman, and M. Peloso, *Equation of state and beginning of thermalization after preheating*, *Phys. Rev. D* **73** (2006) 023501, [hep-ph/0507096].

[226] J. B. Munoz and M. Kamionkowski, *Equation-of-State Parameter for Reheating*, *Phys. Rev. D* **91** (2015), no. 4 043521, [arXiv:1412.0656].

[227] L. Dai, M. Kamionkowski, and J. Wang, *Reheating constraints to inflationary models*, *Phys. Rev. Lett.* **113** (2014) 041302, [arXiv:1404.6704].

[228] K. D. Lozanov and M. A. Amin, *Equation of State and Duration to Radiation Domination after Inflation*, *Phys. Rev. Lett.* **119** (2017), no. 6 061301, [arXiv:1608.0121].

[229] D. Maity and P. Saha, *(P)reheating after minimal Plateau Inflation and constraints from CMB*, *JCAP* **07** (2019) 018, [arXiv:1811.1117].

[230] S. Antusch, D. G. Figueroa, K. Marschall, and F. Torrenti, *Energy distribution and equation of state of the early Universe: matching the end of inflation and the onset of radiation domination*, *Phys. Lett. B* **811** (2020) 135888, [arXiv:2005.0756].

[231] R. Allahverdi, R. Brandenberger, F.-Y. Cyr-Racine, and A. Mazumdar, *Reheating in Inflationary Cosmology: Theory and Applications*, *Ann. Rev. Nucl. Part. Sci.* **60** (2010) 27–51, [arXiv:1001.2600].

[232] J. Chluba, R. Khatri, and R. A. Sunyaev, *CMB at 2x2 order: The dissipation of primordial acoustic waves and the observable part of the associated energy release*, *Mon. Not. Roy. Astron. Soc.* **425** (2012) 1129–1169, [arXiv:1202.0057].

[233] J. Chluba *et. al.*, *Spectral Distortions of the CMB as a Probe of Inflation, Recombination, Structure Formation and Particle Physics: Astro2020 Science White Paper*, *Bull. Am. Astron. Soc.* **51** (2019), no. 3 184, [arXiv:1903.0421].

[234] V. S. H. Lee, A. Mitridate, T. Trickle, and K. M. Zurek, *Probing Small-Scale Power Spectra with Pulsar Timing Arrays*, *JHEP* **06** (2021) 028, [arXiv:2012.0985].

[235] K. Van Tilburg, A.-M. Taki, and N. Weiner, *Halometry from Astrometry*, *JCAP* **07** (2018) 041, [arXiv:1804.0199].

[236] D. J. Fixsen and J. C. Mather, *The Spectral Results of the Far-Infrared Absolute Spectrophotometer Instrument on COBE*, **581** (Dec., 2002) 817–822.

[237] A. Riotto, *The Primordial Black Hole Formation from Single-Field Inflation is Not Ruled Out*, arXiv:2301.0059.

[238] K. N. Ananda, C. Clarkson, and D. Wands, *The Cosmological gravitational wave background from primordial density perturbations*, *Phys. Rev. D* **75** (2007) 123518, [gr-qc/0612013].

[239] D. Baumann, P. J. Steinhardt, K. Takahashi, and K. Ichiki, *Gravitational Wave Spectrum Induced by Primordial Scalar Perturbations*, *Phys. Rev. D* **76** (2007) 084019, [hep-th/0703290].

[240] K. Kohri and T. Terada, *Semianalytic calculation of gravitational wave spectrum nonlinearly induced from primordial curvature perturbations*, *Phys. Rev. D* **97** (2018), no. 12 123532, [arXiv:1804.0857].

[241] J. R. Espinosa, D. Racco, and A. Riotto, *A Cosmological Signature of the SM Higgs Instability: Gravitational Waves*, *JCAP* **09** (2018) 012, [arXiv:1804.0773].

[242] P. Adshead, K. D. Lozanov, and Z. J. Weiner, *Non-Gaussianity and the induced gravitational wave background*, *JCAP* **10** (2021) 080, [arXiv:2105.0165].

[243] J.-P. Li, S. Wang, Z.-C. Zhao, and K. Kohri, *Primordial Non-Gaussianity and Anisotropies in Gravitational Waves induced by Scalar Perturbations*, arXiv:2305.1995.

[244] Z.-C. Chen, C. Yuan, and Q.-G. Huang, *Pulsar Timing Array Constraints on Primordial Black Holes with NANOGrav 11-Year Dataset*, *Phys. Rev. Lett.* **124** (2020), no. 25 251101, [arXiv:1910.1223].

[245] S. Garcia-Saenz, L. Pinol, S. Renaux-Petel, and D. Werth, *No-go theorem for scalar-trispectrum-induced gravitational waves*, *JCAP* **03** (2023) 057, [arXiv:2207.1426].

[246] C. Unal, *Imprints of Primordial Non-Gaussianity on Gravitational Wave Spectrum*, *Phys. Rev. D* **99** (2019), no. 4 041301, [arXiv:1811.0915].

[247] V. Atal and G. Domènech, *Probing non-Gaussianities with the high frequency tail of induced gravitational waves*, *JCAP* **06** (2021) 001, [arXiv:2103.0105].

[248] M. Maggiore, *Gravitational wave experiments and early universe cosmology*, *Phys. Rept.* **331** (2000) 283–367, [gr-qc/9909001].

[249] K. Schmitz, *New Sensitivity Curves for Gravitational-Wave Signals from Cosmological Phase Transitions*, *JHEP* **01** (2021) 097, [arXiv:2002.0461].

[250] A. Sesana *et. al.*, *Unveiling the gravitational universe at μ-Hz frequencies*, *Exper. Astron.* **51** (2021), no. 3 1333–1383, [arXiv:1908.1139].

[251] M. Braglia and S. Kuroyanagi, *Probing prerecombination physics by the cross-correlation of stochastic gravitational waves and CMB anisotropies*, *Phys. Rev. D* **104** (2021), no. 12 123547, [arXiv:2106.0378].

[252] **NANOGrav** Collaboration, G. Agazie *et. al.*, *The NANOGrav 15 yr Data Set: Evidence for a Gravitational-wave Background*, *Astrophys. J. Lett.* **951** (2023), no. 1 L8, [arXiv:2306.1621].

[253] **EPTA** Collaboration, J. Antoniadis *et. al.*, *The second data release from the European Pulsar Timing Array III. Search for gravitational wave signals*, arXiv:2306.1621.

[254] **EPTA** Collaboration, J. Antoniadis *et. al.*, *The second data release from the European Pulsar Timing Array - I. The dataset and timing analysis*, *Astron. Astrophys.* **678** (2023) A48, [arXiv:2306.1622].

[255] D. J. Reardon *et. al.*, *Search for an Isotropic Gravitational-wave Background with the Parkes Pulsar Timing Array*, *Astrophys. J. Lett.* **951** (2023), no. 1 L6, [arXiv:2306.1621].

[256] H. Xu *et. al.*, *Searching for the Nano-Hertz Stochastic Gravitational Wave Background with the Chinese Pulsar Timing Array Data Release I*, *Res. Astron. Astrophys.* **23** (2023), no. 7 075024, [arXiv:2306.1621].

[257] **NANOGrav** Collaboration, A. Afzal *et. al.*, *The NANOGrav 15 yr Data Set: Search for Signals from New Physics*, *Astrophys. J. Lett.* **951** (2023), no. 1 L11, [arXiv:2306.1621].

[258] **EPTA** Collaboration, J. Antoniadis *et. al.*, *The second data release from the European Pulsar Timing Array: V. Implications for massive black holes, dark matter and the early Universe*, arXiv:2306.1622.

[259] M. Baumgart *et. al.*, *Snowmass Theory Frontier: Effective Field Theory*, in *Snowmass 2021*, 10, 2022. arXiv:2210.0319.

[260] H. D. Politzer, *Power Corrections at Short Distances*, *Nucl. Phys. B* **172** (1980) 349–382.

[261] H. Georgi, *On-shell effective field theory*, *Nucl. Phys. B* **361** (1991) 339–350.

[262] C. Arzt, *Reduced effective Lagrangians*, *Phys. Lett. B* **342** (1995) 189–195, [hep-ph/9304230].

[263] B. Henning, X. Lu, T. Melia, and H. Murayama, *Operator bases, S-matrices, and their partition functions*, *JHEP* **10** (2017) 199, [arXiv:1706.0852].

[264] C. Cheung, P. Creminelli, A. L. Fitzpatrick, J. Kaplan, and L. Senatore, *The Effective Field Theory of Inflation*, *JHEP* **03** (2008) 014, [arXiv:0709.0293].

[265] S. Weinberg, *Effective Field Theory for Inflation*, *Phys. Rev. D* **77** (2008) 123541, [arXiv:0804.4291].

[266] J. M. Maldacena, *Non-Gaussian features of primordial fluctuations in single field inflationary models*, *JHEP* **05** (2003) 013, [astro-ph/0210603].

[267] D. Seery and J. E. Lidsey, *Non-Gaussian Inflationary Perturbations from the dS/CFT Correspondence*, *JCAP* **06** (2006) 001, [astro-ph/0604209].

[268] F. Arroja and T. Tanaka, *A note on the role of the boundary terms for the non-Gaussianity in general k-inflation*, *JCAP* **05** (2011) 005, [arXiv:1103.1102].

[269] G. Rigopoulos, *Gauge invariance and non-Gaussianity in Inflation*, *Phys. Rev. D* **84** (2011) 021301, [arXiv:1104.0292].

[270] S. Renaux-Petel, *On the redundancy of operators and the bispectrum in the most general second-order scalar-tensor theory*, *JCAP* **02** (2012) 020, [arXiv:1107.5020].

[271] R. H. Ribeiro and D. Seery, *Decoding the bispectrum of single-field inflation*, *JCAP* **10** (2011) 027, [arXiv:1108.3839].

[272] C. de Rham and R. H. Ribeiro, *Riding on irrelevant operators*, *JCAP* **11** (2014) 016, [arXiv:1405.5213].

[273] L. Bordin, G. Cabass, P. Creminelli, and F. Vernizzi, *Simplifying the EFT of Inflation: generalized disformal transformations and redundant couplings*, *JCAP* **09** (2017) 043, [arXiv:1706.0375].

[274] S. Garcia-Saenz, L. Pinol, and S. Renaux-Petel, *Revisiting non-Gaussianity in multifield inflation with curved field space*, *JHEP* **01** (2020) 073, [arXiv:1907.1040].

[275] L. Bordin and G. Cabass, *Graviton non-Gaussianities and Parity Violation in the EFT of Inflation*, *JCAP* **07** (2020) 014, [arXiv:2004.0061].

[276] D. Green and E. Pajer, *On the Symmetries of Cosmological Perturbations*, *JCAP* **09** (2020) 032, [arXiv:2004.0958].

[277] E. Pajer, *Building a Boostless Bootstrap for the Bispectrum*, *JCAP* **01** (2021) 023, [arXiv:2010.1281].

[278] D. Ghosh, K. Panchal, and F. Ullah, *Mixed graviton and scalar bispectra in the EFT of inflation: Soft limits and Boostless Bootstrap*, *JHEP* **07** (2023) 233, [arXiv:2303.1692].

[279] H.-L. Li, Z. Ren, M.-L. Xiao, J.-H. Yu, and Y.-H. Zheng, *On-shell operator construction in the effective field theory of gravity*, *JHEP* **10** (2023) 019, [arXiv:2305.1048].

[280] X. Chen and Y. Wang, *Quasi-Single Field Inflation and Non-Gaussianities*, *JCAP* **04** (2010) 027, [arXiv:0911.3380].

[281] D. Baumann and D. Green, *Equilateral Non-Gaussianity and New Physics on the Horizon*, *JCAP* **09** (2011) 014, [arXiv:1102.5343].

[282] D. Baumann and D. Green, *Signatures of Supersymmetry from the Early Universe*, *Phys. Rev. D* **85** (2012) 103520, [arXiv:1109.0292].

[283] X. Chen and Y. Wang, *Quasi-Single Field Inflation with Large Mass*, *JCAP* **09** (2012) 021, [arXiv:1205.0160].

[284] T. Noumi, M. Yamaguchi, and D. Yokoyama, *Effective field theory approach to quasi-single field inflation and effects of heavy fields*, *JHEP* **06** (2013) 051, [arXiv:1211.1624].

[285] V. Assassi, D. Baumann, D. Green, and L. McAllister, *Planck-Suppressed Operators*, *JCAP* **01** (2014) 033, [arXiv:1304.5226].

[286] N. Craig and D. Green, *Testing Split Supersymmetry with Inflation*, *JHEP* **07** (2014) 102, [arXiv:1403.7193].

[287] E. Dimastrogiovanni, M. Fasiello, and M. Kamionkowski, *Imprints of Massive Primordial Fields on Large-Scale Structure*, *JCAP* **02** (2016) 017, [arXiv:1504.0599].

[288] P. D. Meerburg, M. Münchmeyer, J. B. Muñoz, and X. Chen, *Prospects for Cosmological Collider Physics*, *JCAP* **03** (2017) 050, [arXiv:1610.0655].

[289] H. Lee, D. Baumann, and G. L. Pimentel, *Non-Gaussianity as a Particle Detector*, *JHEP* **12** (2016) 040, [arXiv:1607.0373].

[290] X. Chen, Y. Wang, and Z.-Z. Xianyu, *Standard Model Background of the Cosmological Collider*, *Phys. Rev. Lett.* **118** (2017), no. 26 261302, [arXiv:1610.0659].

[291] X. Chen, Y. Wang, and Z.-Z. Xianyu, *Loop Corrections to Standard Model Fields in Inflation*, *JHEP* **08** (2016) 051, [arXiv:1604.0784].

[292] X. Chen, Y. Wang, and Z.-Z. Xianyu, *Standard Model Mass Spectrum in Inflationary Universe*, *JHEP* **04** (2017) 058, [arXiv:1612.0812].

[293] H. An, M. McAneny, A. K. Ridgway, and M. B. Wise, *Quasi Single Field Inflation in the non-perturbative regime*, *JHEP* **06** (2018) 105, [arXiv:1706.0997].

[294] X. Chen, Y. Wang, and Z.-Z. Xianyu, *Schwinger-Keldysh Diagrammatics for Primordial Perturbations*, *JCAP* **12** (2017) 006, [arXiv:1703.1016].

[295] S. Kumar and R. Sundrum, *Heavy-Lifting of Gauge Theories By Cosmic Inflation*, *JHEP* **05** (2018) 011, [arXiv:1711.0398].

[296] D. Baumann, G. Goon, H. Lee, and G. L. Pimentel, *Partially Massless Fields During Inflation*, *JHEP* **04** (2018) 140, [arXiv:1712.0662].

[297] A. Moradinezhad Dizgah, H. Lee, J. B. Muñoz, and C. Dvorkin, *Galaxy Bispectrum from Massive Spinning Particles*, *JCAP* **05** (2018) 013, [arXiv:1801.0726].

[298] X. Chen, Y. Wang, and Z.-Z. Xianyu, *Neutrino Signatures in Primordial Non-Gaussianities*, *JHEP* **09** (2018) 022, [arXiv:1805.0265].

[299] N. Arkani-Hamed, D. Baumann, H. Lee, and G. L. Pimentel, *The Cosmological Bootstrap: Inflationary Correlators from Symmetries and Singularities*, *JHEP* **04** (2020) 105, [arXiv:1811.0002].

[300] S. Kumar and R. Sundrum, *Seeing Higher-Dimensional Grand Unification In Primordial Non-Gaussianities*, *JHEP* **04** (2019) 120, [arXiv:1811.1120].

[301] Y.-P. Wu, *Higgs as heavy-lifted physics during inflation*, *JHEP* **04** (2019) 125, [arXiv:1812.1065].

[302] C. Sleight and M. Taronna, *Bootstrapping Inflationary Correlators in Mellin Space*, *JHEP* **02** (2020) 098, [arXiv:1907.0114].

[303] S. Lu, Y. Wang, and Z.-Z. Xianyu, *A Cosmological Higgs Collider*, *JHEP* **02** (2020) 011, [arXiv:1907.0739].

[304] A. Hook, J. Huang, and D. Racco, *Searches for other vacua. Part II. A new Higgstory at the cosmological collider*, *JHEP* **01** (2020) 105, [arXiv:1907.1062].

[305] A. Hook, J. Huang, and D. Racco, *Minimal signatures of the Standard Model in non-Gaussianities*, *Phys. Rev. D* **101** (2020), no. 2 023519, [arXiv:1908.0001].

[306] S. Kumar and R. Sundrum, *Cosmological Collider Physics and the Curvaton*, *JHEP* **04** (2020) 077, [arXiv:1908.1137].

[307] L.-T. Wang and Z.-Z. Xianyu, *In Search of Large Signals at the Cosmological Collider*, *JHEP* **02** (2020) 044, [arXiv:1910.1287].

[308] D. Baumann, C. Duaso Pueyo, A. Joyce, H. Lee, and G. L. Pimentel, *The cosmological bootstrap: weight-shifting operators and scalar seeds*, *JHEP* **12** (2020) 204, [arXiv:1910.1405].

[309] L. Li, T. Nakama, C. M. Sou, Y. Wang, and S. Zhou, *Gravitational Production of Superheavy Dark Matter and Associated Cosmological Signatures*, *JHEP* **07** (2019) 067, [arXiv:1903.0884].

[310] S. Alexander, S. J. Gates, L. Jenks, K. Koutrolikos, and E. McDonough, *Higher Spin Supersymmetry at the Cosmological Collider: Sculpting SUSY Rilles in the CMB*, *JHEP* **10** (2019) 156, [arXiv:1907.0582].

[311] K. Kogai, K. Akitsu, F. Schmidt, and Y. Urakawa, *Galaxy imaging surveys as spin-sensitive detector for cosmological colliders*, *JCAP* **03** (2021) 060, [arXiv:2009.0551].

[312] A. Bodas, S. Kumar, and R. Sundrum, *The Scalar Chemical Potential in Cosmological Collider Physics*, *JHEP* **02** (2021) 079, [arXiv:2010.0472].

[313] S. Aoki and M. Yamaguchi, *Disentangling mass spectra of multiple fields in cosmological collider*, *JHEP* **04** (2021) 127, [arXiv:2012.1366].

[314] S. Lu, *Axion isocurvature collider*, *JHEP* **04** (2022) 157, [arXiv:2103.0595].

[315] Q. Lu, M. Reece, and Z.-Z. Xianyu, *Missing scalars at the cosmological collider*, *JHEP* **12** (2021) 098, [arXiv:2108.1138].

[316] L.-T. Wang, Z.-Z. Xianyu, and Y.-M. Zhong, *Precision calculation of inflation correlators at one loop*, *JHEP* **02** (2022) 085, [arXiv:2109.1463].

[317] X. Tong, Y. Wang, and Y. Zhu, *Cutting rule for cosmological collider signals: a bulk evolution perspective*, *JHEP* **03** (2022) 181, [arXiv:2112.0344].

[318] Y. Cui and Z.-Z. Xianyu, *Probing Leptogenesis with the Cosmological Collider*, *Phys. Rev. Lett.* **129** (2022), no. 11 111301, [arXiv:2112.1079].

[319] G. L. Pimentel and D.-G. Wang, *Boostless cosmological collider bootstrap*, *JHEP* **10** (2022) 177, [arXiv:2205.0001].

[320] X. Chen, R. Ebadi, and S. Kumar, *Classical Cosmological Collider Physics and Primordial Features*, arXiv:2205.0110.

[321] N. Anil Kumar, G. Sato-Polito, M. Kamionkowski, and S. C. Hotinli, *Primordial trispectrum from kinetic Sunyaev-Zel'dovich tomography*, *Phys. Rev. D* **106** (2022), no. 6 063533, [arXiv:2205.0342].

[322] S. Jazayeri and S. Renaux-Petel, *Cosmological bootstrap in slow motion*, *JHEP* **12** (2022) 137, [arXiv:2205.1034].

[323] Z. Qin and Z.-Z. Xianyu, *Helical inflation correlators: partial Mellin-Barnes and bootstrap equations*, *JHEP* **04** (2023) 059, [arXiv:2208.1379].

[324] X. Niu, M. H. Rahat, K. Srinivasan, and W. Xue, *Gravitational wave probes of massive gauge bosons at the cosmological collider*, *JCAP* **02** (2023) 013, [arXiv:2211.1433].

[325] X. Chen, J. Fan, and L. Li, *New inflationary probes of axion dark matter*, arXiv:2303.0340.

[326] Z. Qin and Z.-Z. Xianyu, *Inflation correlators at the one-loop order: nonanalyticity, factorization, cutting rule, and OPE*, *JHEP* **09** (2023) 116, [arXiv:2304.1329].

[327] S. Aoki, T. Noumi, F. Sano, and M. Yamaguchi, *Analytic Formulae for Inflationary Correlators with Dynamical Mass*, arXiv:2312.0964.

[328] M. M. Anber and L. Sorbo, *Naturally inflating on steep potentials through electromagnetic dissipation*, *Phys. Rev. D* **81** (2010) 043534, [arXiv:0908.4089].

[329] A. Maleknejad and M. M. Sheikh-Jabbari, *Non-Abelian Gauge Field Inflation*, *Phys. Rev. D* **84** (2011) 043515, [arXiv:1102.1932].

[330] A. Maleknejad, M. M. Sheikh-Jabbari, and J. Soda, *Gauge Fields and Inflation*, *Phys. Rept.* **528** (2013) 161–261, [arXiv:1212.2921].

[331] P. Adshead and M. Wyman, *Chromo-Natural Inflation: Natural inflation on a steep potential with classical non-Abelian gauge fields*, *Phys. Rev. Lett.* **108** (2012) 261302, [arXiv:1202.2366].

[332] P. Adshead, L. Pearce, M. Peloso, M. A. Roberts, and L. Sorbo, *Phenomenology of fermion production during axion inflation*, *JCAP* **06** (2018) 020, [arXiv:1803.0450].

[333] V. Domcke and K. Mukaida, *Gauge Field and Fermion Production during Axion Inflation*, *JCAP* **11** (2018) 020, [arXiv:1806.0876].

[334] D. Baumann, D. Green, A. Joyce, E. Pajer, G. L. Pimentel, C. Sleight, and M. Taronna, *Snowmass White Paper: The Cosmological Bootstrap*, in *Snowmass 2021*, 3, 2022. arXiv:2203.0812.

[335] J. C. Criado and M. Pérez-Victoria, *Field redefinitions in effective theories at higher orders*, *JHEP* **03** (2019) 038, [arXiv:1811.0941].

[336] B. Gripaios and D. Sutherland, *DEFT: A program for operators in EFT*, *JHEP* **01** (2019) 128, [arXiv:1807.0754].

[337] R. M. Fonseca, *The Sym2Int program: going from symmetries to interactions*, *J. Phys. Conf. Ser.* **873** (2017), no. 1 012045, [arXiv:1703.0522].

[338] R. M. Fonseca, *Enumerating the operators of an effective field theory*, *Phys. Rev. D* **101** (2020), no. 3 035040, [arXiv:1907.1258].

[339] L.-T. Wang and Z.-Z. Xianyu, *Gauge Boson Signals at the Cosmological Collider*, *JHEP* **11** (2020) 082, [arXiv:2004.0288].

[340] M. B. Einhorn and J. Wudka, *The Bases of Effective Field Theories*, *Nucl. Phys. B* **876** (2013) 556–574, [arXiv:1307.0478].

[341] N. Craig, M. Jiang, Y.-Y. Li, and D. Sutherland, *Loops and Trees in Generic EFTs*, *JHEP* **08** (2020) 086, [arXiv:2001.0001].

[342] P. Creminelli, *On non-Gaussianities in single-field inflation*, *JCAP* **10** (2003) 003, [astro-ph/0306122].

[343] Z.-Z. Xianyu and J. Zang, *Inflation Correlators with Multiple Massive Exchanges*, arXiv:2309.1084.

[344] T. Inami and C. Lim, *Effects of Superheavy Quarks and Leptons in Low-Energy Weak Processes $K_L \to \bar{\mu}\mu$, $K^+ \to \pi^+\nu\bar{\nu}$ and $K^0 \leftrightarrow \bar{K}^0$*, *Prog. Theor. Phys.* **65** (1981) 297. [Erratum: Prog.Theor.Phys. 65, 1772 (1981)].

[345] G. Ecker and W. Grimus, *CP Violation and Left-Right Symmetry*, *Nucl. Phys. B* **258** (1985) 328–360.