University of California
Santa Barbara

New Approaches to the Hierarchy Problem and their Signatures from Microscopic to Cosmic Scales

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

Doctor of Philosophy

in

Physics

by

Seth Koren

Committee in charge:

Professor Nathaniel Craig, Chair
Professor Mark Srednicki
Professor Claudio Campagnari

September 2020

The Dissertation of Seth Koren is approved.

_____

Professor Mark Srednicki

_____

Professor Claudio Campagnari

_____

Professor Nathaniel Craig, Committee Chair

August 2020

New Approaches to the Hierarchy Problem

and their Signatures from Microscopic to Cosmic Scales

"You seem very clever at explaining words, Sir", said Alice.

"Would you kindly tell me the meaning of the poem 'Jabberwocky'?"

"Let's hear it", said Humpty Dumpty.

"I can explain all the poems that ever were invented

— and a good many that haven't been invented just yet."

Lewis Carroll

Through the Looking Glass (1871) [1]

Acknowledgements

Of course I have not worked alone, and very little of this science would have been accomplished were it not for my many collaborators who have helped me learn and problem-solve and provided guidance and been patient when I was overwhelmed with being pulled in too many directions. I especially mention those undergrads I have spent significant time mentoring during my time in graduate school, namely Aidan Herderschee, Samuel Alipour-fard, and Umut Can Öktem. Indeed, they each took a chance on me as well, and working with them has taught me how to be a better teacher and physicist—not to mention all of the great science we worked out together.

My physics knowledge would have also been stunted were it not for the countless hours spent discussing all manner of high energy theory with friends and peers—primarily Matthew Brown, Brianna Grado-White, Alex Kinsella, Robert McGehee, and Gabriel Treviño Verastegui. And my enjoyment of graduate school would have been stunted were it not for the board games, hiking, trivia, art walks, biking, and late-night philosophical discussions with them and with Dillon Cisco, Neelay Fruitwala, Eric Jones, Farzan Vafa, Sicheng Wang, and others. No one has enriched my life here moreso than Nicole Iannaccone—a summary statement of far too few words.

Nor would this have been possible were it not for my 'medical support team' consisting primarily of endocrinologists Dr. Mark Wilson and Dr. Ashley Thorsell, student health physician Dr. Miguel Pedroza, and my therapist Dr. Karen Dias, who have helped me immensely in my time here. Even putting aside physiological issues, graduate school can be and has been incredibly mentally taxing. I truly could not have made it through were it not for the psychological assistance I have received, both psychotherapeutic and pharmacological.

Of course a special 'shout-out' goes to Mother Nature. The idyllic weather, scenic ocean, beautiful mountains, and clear night skies of Santa Barbara may have just been too distracting for me to progress through my degree were it not for the Rey fire, the

Whittier fire, the Cave fire, the Santa Barbara microburst, the recurring January flooding, the Ridgecrest earthquakes, the month of unbreathable air from the Thomas fire, the threat of the Thomas fire itself, the Montecito mudslides, and of course the COVID-19 pandemic, all of which have kept me indoors studying physics.

Finally, I would like to thank all those who kindly read drafts of (parts of) this thesis and provided invaluable feedback, including Samuel Alipour-fard, Changha Choi, Nathaniel Craig, David Grabovsky, Adolfo Holguin, Lucas Johns, Soubhik Kumar, Umut Can Öktem, Robert McGehee, Alex Meiburg, Gabriel Treviño Verastegui, and Farzan Vafa.

Curriculum Vitæ
Seth Koren

Education

2020	Ph.D. in Physics, University of California, Santa Barbara.
2017	M.A. in Physics, University of California, Santa Barbara.
2015	M.S. in Physics, University of Pennsylvania.
2015	B.A. in Mathematics and Physics, Astrophysics Concentration, Honors Distinction in Physics, University of Pennsylvania.

Publications

"Supersoft Stops"
T. Cohen, N. Craig, S. Koren, M. McCullough, J. Tooby-Smith
Submitted to Phys. Rev. Lett., [arXiv:2002.12630 [hep-ph]] [2]

"IR Dynamics from UV Divergences: UV/IR Mixing, NCFT, and the Hierarchy Problem"
N. Craig and S. Koren
JHEP 03 (2020) 037, [arXiv:1909.01365 [hep-ph]] [3]

"Freezing-in Twin Dark Matter"
S. Koren and R. McGehee
Phys. Rev. D101 (2020) 055024, [arXiv:1908.03559 [hep-ph]] [4]

"The Weak Scale from Weak Gravity"
N. Craig, I. Garcia Garcia, S. Koren
JHEP 09 (2019) 081, [arXiv:1904.08426 [hep-ph]] [5]

"Exploring Strong-Field Deviations From General Relativity via Gravitational Waves"
S. Giddings, S. Koren, G. Treviño
Phys. Rev. D100 (2019) 044005, [arXiv:1904.04258 [gr-qc]] [6]

"Neutrino - DM Scattering and Coincident Detections of UHE Neutrinos with EM Sources"
S. Koren
JCAP 09 (2019) 013, [arXiv:1903.05096 [hep-ph]] [7]

"Constructing N=4 Coulomb Branch Superamplitudes"
A. Herderschee, S. Koren, T. Trott

JHEP 08 (2019) 107, [arXiv:1902.07205 [hep-th]] [8]

"Massive On-Shell Supersymmetric Scattering Amplitudes"
A. Herderschee, S. Koren, T. Trott
JHEP 10 (2019) 092, [arXiv:1902.07204 [hep-th]] [9]

"The second Higgs at the lifetime frontier"
S. Alipour-fard, N. Craig, S. Gori, S. Koren, D. Redigolo
JHEP 07 (2020) 029, [arXiv:1812.09315 [hep-ph]] [10]

"Discrete Gauge Symmetries and the Weak Gravity Conjecture"
N. Craig, I. Garcia Garcia, S. Koren
JHEP 05 (2019) 140, [arXiv:1812.08181 [hep-th]] [11]

"Long Live the Higgs Factory: Higgs Decays to Long-Lived Particles at Future Lepton Colliders"
S. Alipour-fard, N. Craig, M. Jiang, S. Koren
Chin. Phys. C43 (2019) 053101, [arXiv:1812.05588 [hep-ph]] [12]

"Cosmological Signals of a Mirror Twin Higgs"
N. Craig, S. Koren, T. Trott
JHEP 05 (2017) 038, [arXiv:1611.07977 [hep-ph]] [13]

"The Low-Mass Astrometric Binary LSR1610-0040"
S. C. Koren, C. H. Blake, C. C. Dahn, H. C. Harris
The Astronomical Journal 151 (2016) 57, [arXiv:1511.02234 [astro-ph.SR]] [14]

"Characterizing Asteroids Multiply-Observed at Infrared Wavelengths"
S. C. Koren, E. L. Wright, A. Mainzer
Icarus 258 (2015) 82-91, [arXiv:1506.04751 [astro-ph.EP]] [15]

Abstract

New Approaches to the Hierarchy Problem

and their Signatures from Microscopic to Cosmic Scales

by

Seth Koren

It's an exciting time in particle theory. Fantastic experimental effort has revealed that the major ideas of the past few decades of particle theory are increasingly in tension with empirical data—underscoring the need for new ideas and strategies. In this thesis we focus on the application of this philosophy toward the hierarchy problem. We begin with a pedagogical introduction aimed at grounding the hierarchy problem in a clear understanding of renormalization and a comprehensive discussion of the problem while debunking common confusions. We survey classic approaches and argue that the LHC data has qualitatively changed the problem. We discuss the Twin Higgs as the avatar of the 'neutral naturalness' strategy, and its exotic signatures which shift the signals of naturalness away from the energy frontier. Finally, we suggest a new trail for the naturalness paradigm: violations of effective field theory—as demanded by quantum gravity—challenge the assumptions of the space of solutions and suggest qualitatively new avenues of pursuit.

Permissions and Attributions

1. The content of Chapter 5 is the result of collaboration with Nathaniel Craig and Timothy Trott, and separately with Robert McGehee. This work previously appeared in the Journal of High Energy Physics (JHEP **05** (2017) 038) and Physical Review D (Phys. Rev. **D101** (2020) 055024), respectively.

2. The content of Chapter 6 is the result of collaboration with Samuel Alipour-fard, Nathaniel Craig, and Minyuan Jiang, and previously appeared in Chinese Physics C (Chin. Phys. **C43** (2019) 053101).

3. The content of Chapter 7 is the result of collaboration with Nathaniel Craig and previously appeared in the Journal of High Energy Physics (JHEP **03** (2020) 037).

Preface

The first four chapters of this thesis are introductory material which has not previously appeared in any public form. My intention has been to write the guide that would have been most useful for me toward the beginning of my graduate school journey as a field theorist interested in the hierarchy problem. My aim has been to make these chapters accessible to beginning graduate students in particle physics and interested parties in related fields—background at the level of a single semester of quantum field theory should be enough for them to be understandable in broad strokes.

Chapter 1 introduces fundamental tools and concepts in quantum field theory which are essential for particle theory, spending especial effort on discussing renormalization from a variety of perspective. Chapter 2 discusses the hierarchy problem and how to think about it—primarily through the pedagogical device of refuting a variety of common misconceptions and pitfalls. Chapter 3 introduces in brief a variety of classic strategies and solutions to the hierarchy problem which also constitute important frameworks in theoretical particle physics beyond the Standard Model. Chapter 4 discusses more-recent ideas about the hierarchy problem in light of the empirical pressure supplied by the lack of observed new physics at the Large Hadron Collider. Throughout I also make note of interesting research programs which, while they lie too far outside the main narrative for me to explain, are too fascinating not to be mentioned.

The first half of this thesis is thus mostly an introduction to and review of material I had no hand in inventing. As always, I am 'standing on the shoulders of giants', and I have benefited enormously from the pedagogical efforts of those who came before me. When my thinking on a topic has been especially informed by a particular exposition, or when I present an example which was discussed in a particular source, I will endeavor to say so and refer to that presentation. As to the rest, it's somewhere between difficult and

impossible to distinguish precisely how and whose ideas I have melded together in my own understanding of the topics—to say nothing of any insight I may have had myself—but I have included copious references to reading material I enjoyed as a guide. Ultimately this is a synthesis of ideas in particle theory aimed toward the particular purpose of understanding the hierarchy problem, and I have attempted to include the most useful and pedagogical explanations of these topics I could find, if not invent.

I then present some work on the subject by myself and my collaborators. Chapter 5 contains work constructing a viable cosmological history for mirror twin Higgs models, an exemplar of the modern Neutral Naturalness approach to the hierarchy problem. Chapter 6 focuses on searching for long-lived particles produced at particle colliders as a discovery channel for a broad class of such models. Chapter 7 is an initial exploration of a new approach to the hierarchy problem which follows a maximalist interpretation of the lack of new observed TeV scale physics, and so relies on questioning and modifying some core assumptions of conventional particle physics. In Chapter 8 we conclude with some parting thoughts.

Contents

Chapter 1

Effective Field Theory

The formulation and understanding of the hierarchy problem is steeped heavily in the principles and application of effective field theory (EFT), so we begin with an introductory overview to set the stage for our main discussion. As is clear from the table of contents, I have prioritized clarity over brevity, especially when it comes to renormalization. The reader with a strong background in particle physics may find much of this to be review, so may wish to skip ahead directly to Chapter 2 and circle back to sections of this chapter if and when the subtleties they discuss become relevant.

We will endeavor to discuss the conceptual points which will be useful later in understanding the hierarchy problem, and more generally to clarify common confusions with ample examples. Of course we will be unable to discuss everything, and will try to provide references to more detailed explanations when we must needs say less than we would like. Some generally useful introductions to effective field theory can be found from Cohen [16] and Georgi [17].

1.1   EFT Basics

Effective field theory is simply the familiar strategy to focus on the important degrees of freedom when understanding a physical situation. For a simple example from an introductory Newtonian mechanics course, consider studying the motion of balls on inclined planes in a freshman lab. It is neither necessary nor useful to model the short-distance physics of the atomic composition of the ball, nor the high-energy physics of special relativity. Inversely, it is also unnecessary to account for the long-distance physics of Hubble expansion or the low-energy physics of air currents in the lab. In quantum field theories this intuitive course of action is formalized in decoupling theorems, showing precisely the sense in which field theories are amenable to this sort of analysis: the effects of short-distance degrees of freedom may be taken into account as slight modifications to the interactions of long-distance degrees of freedom, instead of including explicitly those high-energy modes.

Of course when one returns to the mechanics laboratory armed with an atomic clock and a scanning tunneling microscope, one begins to see deviations from the Newtonian predictions. Indeed, the necessary physics for describing a situation depends not only on the dynamics under consideration but also on the precision one is interested in attaining with the description. So it is crucial that one is able to correct the leading-order description by systematically adding in subdominant effects, as organized in a suitable power series in, for example, (v/c), where v is the ball's velocity and c is the speed of light. Of course when the full description of the physics is known it's in principle possible to just use the full theory to compute observables—but I'd still rather not begin with the QED Lagrangian to predict the deformation of a ball rolling down a ramp.

The construction of an appropriate effective description relies on three ingredients. The first is a list of the important degrees of freedom which specify the system under

consideration—in particle physics this is often some fields ϕ_i. The second is the set of symmetries which these degrees of freedom enjoy. These constrain the allowed interactions between our fields and so control the dynamics of the theory. Finally we need a notion of power counting, which organizes the effects in terms of importance. This will allow us to compute quantities to the desired precision systematically. Frequently in effective field theories of use in particle physics this role is played by E/Λ, where E is an energy and Λ is a heavy mass scale or cutoff above which we expect to require a new description of the physics.

1.1.1   Scale-dependence

We will often be interested in determining the appropriate description of a system at some scale, so it is necessary to understand which degrees of freedom and which interactions will be important as a function of energy. We can gain insight into when certain modes or couplings are important by studying the behavior of our system under scale transformations. Consider for example a theory of a real scalar field ϕ, with action

$$S = \int \mathrm{d}^d x \left(-\frac{1}{2}\partial_\mu \phi \partial^\mu \phi - \frac{1}{2}m^2\phi^2 - \frac{1}{4!}\lambda\phi^4 - \frac{1}{6!}\tau\phi^6 + \dots \right), \qquad (1.1)$$

where d is the dimensionality of spacetime, m, λ, τ are couplings of interactions involving different numbers of ϕ, the $\dots$ denote terms with higher powers of ϕ, and we've imposed a $\mathbb{Z}_2 : \phi \to -\phi$ symmetry for simplicity. We've set $c = \hbar = 1$ which is often called 'working in natural units', leaving us with solely one sort of dimensionality to speak of, mass = energy = length^{-1} = time^{-1}, which is called 'mass dimension'. From the fact that regardless of dimension we have $[S] = 0$ and $[x] = -1$, where $[\cdot]$ denotes the mass dimension, we first calculate from the kinetic term that $[\phi] = \frac{d-2}{2}$, and then we can read off $[m^2] = 2$, $[\lambda] = 4 - d$, $[\tau] = 6 - 2d$. These are known as the classical dimension of

3

the associated operators and are associated to the behavior of the operators at different energies, for the following reason.

If we wish to understand how the physics of this theory varies as a function of scale, we can perform a transformation $x^\mu \to s x'^\mu$ and study the long-distance limit $s \gg 1$ with x'^μ fixed. The measure transforms as $\mathrm{d}^d x \to s^d \mathrm{d}^d x'$ and the derivatives $\partial_\mu \to s^{-1} \partial_{\mu'}$. Then to restore canonical normalization of the kinetic term such that the one-particle states are properly normalized for the LSZ formula to work, we must perform a field redefinition $\phi(x) = s^{\frac{2-d}{2}} \phi'(x')$, and the action becomes

$$S = \int \mathrm{d}^d x' \left(-\frac{1}{2} \partial_{\mu'} \phi' \partial^{\mu'} \phi' - \frac{1}{2} m^2 s^2 \phi'^2 - \frac{1}{4!} \lambda s^{4-d} \phi'^4 - \frac{1}{6!} \tau s^{6-2d} \phi'^6 + \dots \right). \quad (1.2)$$

As a reminder, in the real world (at least at distances $\gtrsim 1~\mu m$) we have $d = 4$. As you look at the theory at longer distances the mass term becomes more important, so is known as a 'relevant' operator. One says that the operator ϕ^2 has classical dimension $\Delta_{\phi^2} = 2$. The quartic interaction is classically constant under such a transformation, so is known as 'marginal' with $\Delta_{\phi^4} = 0$, and interactions with more powers of ϕ shrink at low energies and are termed 'irrelevant', e.g. $\Delta_{\phi^6} = -2$. We have been careful to specify that these are the *classical* dimension of the operators, also called the 'engineering dimension' or 'canonical dimension', which has a simple relation to the mass dimension as $\Delta_\mathcal{O} = d - [\mathcal{O}]$ for some operator $\mathcal{O}$. If the theory is not scale-invariant then quantum corrections modify this classical scaling by an 'anomalous dimension' $\delta_\mathcal{O}(m^2, \lambda, \tau, \dots)$ which is a function of the couplings of the theory, and the full behavior is known as the 'scaling dimension'. The terms 'marginally (ir)relevant' are used for operators whose classical dimension is zero but whose anomalous dimensions push them to one side.

The connection to the typical EFT power counting in (E/Λ) is immediate. In an EFT with UV cutoff Λ, it's natural to normalize all of our couplings with this scale and

rename e.g. $\tau \to \bar{\tau}\Lambda^{6-2d}$ where $\bar{\tau}$ is now dimensionless. It's easy to see that the long-distance limit is equivalently a low-energy limit by considering the derivatives, which pull down a constant p'^{μ} and scale as s^{-1}—or by simply invoking the uncertainty principle. Operators with negative scaling dimension contribute subleading effects at low energies precisely because of these extra powers of a large inverse mass scale.

1.1.2   Bottom-up or Top-down

> Then he made the tank of cast metal, 10 cubits across from brim to brim, completely round; it was 5 cubits high, and it measured 30 cubits in circumference.

> God on the merits of working to finite precision
>
> 1 Kings 7:23, Nevi'im
>
> New Jewish Publication Society Translation (1985) [18]

The procedure of writing down the most general Lagrangian with the given degrees of freedom and respecting the given symmetries up to some degree of power counting is termed 'bottom-up EFT' as we're constructing it entirely generally and will have to fix coefficients by making measurements. A great example is the Standard Model Effective Field Theory (SMEFT), of which the Standard Model itself is described by the SMEFT Lagrangian at zeroth order in the power counting. It is defined by being an $SU(3) \times SU(2) \times U(1)$ gauge theory with three generations of the following representations of left-handed Weyl fermions:

$3x$ Fermions	$SU(3)_C$	$SU(2)_L$	$U(1)_Y$
Q	3	2	$\frac{1}{6}$
$\bar{u}$	$\bar{3}$	-	$-\frac{2}{3}$
$\bar{d}$	$\bar{3}$	-	$\frac{1}{3}$
L	-	2	$-\frac{1}{2}$
$\bar{e}$	-	-	1

In addition the Standard Model contains one scalar, the Higgs boson, which is responsible for implementing the Anderson-Brout-Englert-Guralnik-Hagen-Higgs-Kibble-'t Hooft mechanism [19, 20, 21, 22, 23, 24] to break the electroweak symmetry $SU(2)_L \times U(1)_Y$ down to electromagnetism $U(1)_Q$ at low energies:

H	-	2	$-\frac{1}{2}$

The Standard Model Lagrangian contains all relevant and marginal gauge-invariant operators which can be built out of these fields, and has the following schematic form

$$\mathcal{L}_{\text{kin}} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - i\bar{\psi}\slashed{D}\psi - (D_\mu H)^\dagger (D^\mu H) \tag{1.3}$$

$$\mathcal{L}_{Higgs} = -yH\Psi\psi + \text{ h.c. } + m^2 H^\dagger H - \frac{\lambda}{4}(H^\dagger H)^2, \tag{1.4}$$

with F a gauge field strength, ψ a fermion, D the gauge covariant derivative in the kinetic term Lagrangian on the first line, and the second line containing the Higgs' Yukawa couplings and self-interactions. If a refresher on the Standard Model would be useful, the introduction to its structure toward the end of Srednicki's textbook [25] will suffice for our purposes, while further discussion from a variety of perspectives can be found in Schwartz [26], Langacker [27], and Burgess & Moore [28].

The SMEFT power-counting is in energies divided by an as-yet-unknown UV scale Λ, so the dimension-n SMEFT Lagrangian consists of all gauge-invariant combinations of

these fields with scaling dimension $n - d$. At dimension five there is solely one operator, $\mathcal{L}^{(5)} = (LH)^2/\Lambda +$ h.c., which contains a Majorana mass for neutrinos. In even-more-schematic form, the dimension six Lagrangian contains operators with the field content

$$-\Lambda^2 \mathcal{L}^{(6)} = F^3 + H^6 + D^2 H^4 + \psi^2 H^3 + F^2 H^2 + \psi^2 FH + \psi^2 DH^2 + \psi^4, \qquad (1.5)$$

where for aesthetics we have multiplied through by the scale Λ and haven't bothered writing down couplings. After understanding the structure of the independent symmetry-preserving operators (see e.g. [29, 30, 31]), the job of the bottom-up effective field theorist is to measure or constrain the coefficients of these higher-dimensional operators [32]. Useful data comes from both the energy frontier with searches at colliders for production/decay of high-energy particles through these higher-dimensional operators and from the precision frontier measuring fundamental processes very well to look for deviations from the Standard Model predictions (e.g. [33, 34, 35]). For more detail, see the introduction to SMEFT by Brivio & Trott [36].

Another approach is possible when we already have a theory and just want to focus on some particular degrees of freedom. Then we may construct a 'top-down EFT' by taking our theory and 'integrating out' the degrees of freedom we don't care about—for example by starting with the Standard Model above and getting rid of the electroweak bosons to focus on processes occurring at lower energies [37]. We can't necessarily just ignore those degrees of freedom though; what we need to do is modify the parameters of our EFT such that they reproduce the results of the full theory (to some finite precision) using only the low-energy degrees of freedom. Such a procedure can be illustrated formally by playing with path integrals. Consider the partition function for a theory with some light

fields ϕ and some heavy fields Φ:

$$Z = \int \mathcal{D}\phi\mathcal{D}\Phi \ e^{iS(\phi,\Phi)}. \tag{1.6}$$

This contains all of the physics in our theory, and so in principle we may use it to compute anything we wish. But if we're interested in low-energy processes involving solely the ϕ fields, we could split up our path integral and first do the integral over the Φ fields. The light ϕ fields are the only ones left, so we can then write the partition function as

$$Z = \int \mathcal{D}\phi \ e^{iS_{\text{eff}}(\phi)}, \tag{1.7}$$

where this defines S_{eff}. Thus far this still contains all the same physics, as long as we don't want to know about processes with external heavy fields[1]. But having decided that we are interested in the infrared physics of the ϕ fields, we can say that the effects of the heavy Φ fields will be suppressed by factors of the energies of interest divided by the mass of Φ, and we should expand the Lagrangian $\mathcal{L}_{\text{eff}}$ in an appropriate series:

$$Z = \int \mathcal{D}\phi \ \exp\left[i\int \mathrm{d}^d x \left(\mathcal{L}_0(\phi) + \sum_{n=0}^N M_\Phi^{d-n} \sum_i \lambda_i^{(n)} \mathcal{O}_i^{(n)}(\phi)\right)\right], \tag{1.8}$$

where $\mathcal{L}_0(\phi)$ is the part of the full Lagrangian that had no heavy fields in it, $\mathcal{O}_i^{(n)}(\phi)$ is an operator of classical dimension n, $\lambda_i^{(n)}$ is a dimensionless coupling, and $N \geq d$ defines the precision to which one works in this effective theory. This is the procedure to find a

---

[1]Since we haven't made any approximations and have the same object Z, one may be confused as to why we've lost access to the physics of the Φ fields. In fact I've been a bit sloppy. If we want to compute correlation functions of our fields ϕ, we must couple our fields to classical sources J_ϕ as $\mathcal{L} \supset \phi(x)J_\phi(x)$. Physically, those sources allow us to 'turn on' particular fields so that we can then calculate their expectation values. Mathematically, we really need the partition function as a functional of these sources $Z[J_\phi]$, and we take functional derivatives with respect to these sources as a step to calculating correlation functions or scattering amplitudes. In integrating out our heavy field Φ, we no longer have a source we can put in our Lagrangian to turn on that field, as it no longer appears in the action.

top-down effective field theory in the abstract.

A great example of a top-down EFT is in studying the Standard Model fields in the context of a Grand Unified Theory (GUT). Broadly, Grand Unification is the hope that there is some simpler, more symmetric theory behind the Standard Model which explains its structure. A GUT is a model in which the gauge groups of the SM are (partially [38]) unified in the UV. If there is a full unification to a single gauge factor, then this requires 'gauge coupling unification' in the UV until the symmetry is broken down to the SM gauge group at a high scale [39]. While one's first exposure to this idea today may be in the context of a UV theory like string theory which roughly demands such unification, this was in fact first motivated by the observed infrared SM structure. It is frankly amazing that not only are the values of the SM gauge couplings consistent with this idea, and not only does $SU(3) \times SU(2) \times U(1)$ fit nicely inside $SU(5)$, but the SM fermion representations precisely fit into the $10 \oplus \bar{5}$ representations of $SU(5)$. It's hard to imagine a discovery that would have felt much more like one was obviously learning something deep and important about Nature than when Georgi realized how nicely all of this worked out. I'm reminded of Einstein's words on an analogous situation in the early history of electromagnetism—the original unified theory:

> The precise formulation of the time-space laws of those fields was the work of Maxwell. Imagine his feelings when the differential equations he had formulated proved to him that electromagnetic fields spread in the form of polarised waves, and with the speed of light! To few men in the world has such an experience been vouchsafed. At that thrilling moment he surely never guessed that the riddling nature of light, apparently so completely solved, would continue to baffle succeeding generations.
>
> — Albert Einstein

Considerations concerning the Fundaments of Theoretical Physics, 1940 [40]

And just as with Maxwell, the initial deep insight into Nature was not the end of the story. As of yet, Grand Unification remains an unproven ideal, and indeed further empirical data has brought into question the simplest such schemes. But it's hard to imagine all of this beautiful structure is simply coincidental, and I would wager that most high energy theorists still have a GUT in the back of their minds when they think about the UV structure of the universe, so this is an important story to understand. To learn generally about GUTs, I recommend the classic books by Kounnas, Masiero, Nanopoulos, & Olive [41] and Ross [42] or the recent book by Raby [43] for the more formally-minded. Shorter introductions can be found in Sher's TASI lectures [44] or in the Particle Data Group's Review of Particle Physics [45] from Hebecker & Hisano.

The structure of the simplest $SU(5)$ GUT is that the symmetry group breaks down to the SM at energies $M_{GUT} \sim 10^{16}$ GeV via the Higgs mechanism.[2] More generally, unification may proceed in stages as, for example, $SO(10) \rightarrow SU(4)_c \times SU(2)_L \times SU(2)_R \rightarrow SM$, and the breaking may occur via other mechanisms, as we'll discuss further in Section 3. Back to our simple single-breaking example, as is familiar in the SM this means that the gauge bosons corresponding to broken generators get masses of order this GUT-breaking scale. As this is a far higher scale than we are currently able to directly probe, it is neither necessary nor particularly useful to keep these degrees of freedom fully in our description if we're interesting in understanding the effects of GUT-scale fields. Rather than constructing the complete top-down EFT of the SM from a GUT, let's focus on one particularly interesting effect.

---

[2]The scale M_{GUT} is determined from low-energy data by computing the scale-dependence of the SM gauge couplings, evolving them up to high energies, and looking for an energy scale at which they meet. Since we have three gauge couplings at low energies, it is quite non-trivial that the curves $g_i(\mu), i = 1, 2, 3$ meet at a single scale $\mu = M_{GUT}$. The M_{GUT} so computed is approximate not only due to experimental uncertainties on the low-energy values of parameters in the SM, but also because additional particles with SM charges affect slightly how the couplings evolve toward high energies. Indeed, adding supersymmetry makes the intersection of the three curves even more accurate than it is in the SM itself.

One of the best ways to indirectly probe GUTs is by looking for proton decay. The GUT representations unify quarks and leptons, so the extra $SU(5)$ gauge bosons have nonzero baryon and lepton number and fall under the label of 'leptoquarks'. It's worth considering in detail why proton decay is a feature of GUTs and not of the SM. While $U(1)_B$, the baryon number, is an accidental global symmetry of the SM[3], it's an anomalous symmetry and so is not a symmetry of the quantum world. The 'baryon minus lepton' number, $U(1)_{B-L}$, is non-anomalous, but this is a good symmetry both of the SM and of a GUT. What's really behind the stability of the proton is that, though $U(1)_B$ and $U(1)_L$ are not good quantum symmetries, the fact that they are good classical symmetries means their only violation is nonperturbatively by instantons. Such configurations are proportional to the zero-modes of all the fermions charged under them, leading to baryon number being violated only by three units at once, and the proton with $B = 1$ is stable.

But in GUTs, baryon number and lepton number are no longer accidental symmetries, so no such protection is available and the GUT gauge bosons mediate tree-level proton decay processes as in Figure 1.1. We can find the leading effect by integrating these out—in particular we'll here look just at a four-fermion baryon-number-violating operator. The tree-level amplitude is simply

$$i\mathcal{M}(uude) = (ig)^2 v_{u_1} \gamma^\mu u_{u_2} \frac{-i\left(\eta_{\mu\nu} + \frac{p^\mu p^\nu}{M_{GUT}^2}\right)}{p^2 + M_{GUT}^2} v_e \gamma^\nu u_d \approx ig^2 v_{u_1} \gamma^\mu u_{u_2} \frac{\eta_{\mu\nu}}{M_{GUT}^2} v_e \gamma^\nu u_d + \mathcal{O}(\frac{1}{M_{GUT}^2}),$$
$$(1.9)$$

---

[3]'Accidental' here means that imposing this symmetry on the SM Lagrangian does not forbid any operators which would otherwise be allowed. The SM is defined, as above, by the gauge symmetries $SU(3)_C \times SU(2)_L \times U(1)_Y$ and the field content. Writing down the most general dimension-4 Lagrangian 1.3,1.4 invariant under these symmetries gives a Lagrangian which is *automatically* invariant under $U(1)_B$. This no longer holds at higher order in SMEFT, and indeed the dimension-6 Lagrangian 1.5 does contain baryon-number-violating operators. If one wants to study a baryon-number-conserving version of SMEFT, one needs to explicitly impose that symmetry on the dimension-6 Lagrangian, so $U(1)_B$ is no longer an accidental symmetry of SMEFT.
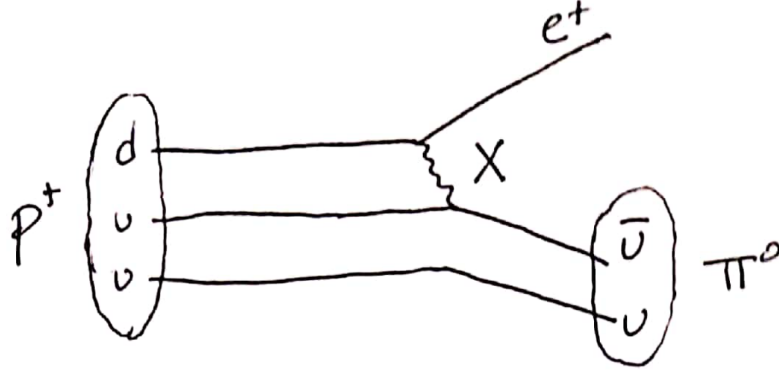
Figure 1.1: A representative diagram contributing to proton decay due to exchange of a heavy GUT gauge boson X.

so integrating out the gauge boson from this diagram gives us one of the contributions to the low-energy operators in Equation 1.8

$$\mathcal{O} = \frac{g^2}{2M_X^2}\epsilon_{\alpha\beta\gamma}\left(u_{R\alpha}\gamma_\mu u_{L\beta}\right)\left(e_R\gamma^\mu d_{L\gamma}\right), \tag{1.10}$$

where, in the notation of Equation 1.8, $\mathcal{O}_i^{(6)} = \epsilon_{\alpha\beta\gamma}\left(u_{R\alpha}\gamma_\mu u_{L\beta}\right)\left(e_R\gamma^\mu d_{L\gamma}\right)$, $\lambda_i^{(6)} = g^2/2$ and $M_\Phi = M_{GUT}$. The calculation of the proton lifetime from this operator is quite complicated, but the dimensional analysis estimate of $\tau_p \sim M_{GUT}^4/g^4 m_p^5$ actually works surprisingly well.

The job of the top-down effective field theorist is to calculate the effects of some particular UV physics on IR observables and by doing so understand how to search for their particular effects. While the effects will, by necessity, be some subset of the operators that the bottom-up effective field theorist has written down, the patterns and correlations present from a particular UV model can suggest or require particular search strategies. In the present context, a GUT may suggest the most promising final states to look for when searching for proton decay. If we wanted to calculate the lifetime and branching ratios more precisely we would have to deal with loop diagrams (among many

12

complications), which of course is a generic feature. So we now turn our attention to the new aspects and challenges that appear once one goes beyond tree-level.

1.2   Renormalization

Renormalization is a notoriously challenging topic for beginning quantum field theorists to grok, and explanations often get bogged down in the details of one particular perspective or scheme or purpose and 'miss the forest for the trees', so to speak.[4] We'll attempt to overcome that issue by discussing a variety of uses for and interpretations of renormalization, as well as how they relate. And, of course, by examining copious examples and pointing out a variety of conceptual pitfalls.

Loops are necessary

At the outset the only fact one needs to have in mind is that renormalization is a procedure which lets quantum field theories make physical predictions given some physical measurements. Such a procedure was not necessary for a classical field theory, which is roughly equivalent to a quantum field theory at tree-level. A natural question for beginners to ask then is why we should bother with loops at all: Why don't we just start off with the physical, measured values in the classical Lagrangian and be done with it? That is, if we measure, say, the mass and self-interaction of some scalar field ϕ, let's just define our theory

$$\mathcal{L}_{\text{exact}} \overset{?}{=} -\frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}m_{\text{phys}}^2\phi^2 - \frac{1}{4!}\lambda_{\text{phys}}\phi^4, \tag{1.11}$$

---

[4]Not that I begrudge QFT textbooks or courses for it, to be clear. There is so much technology to introduce and physics to learn in a QFT class that discussion of all of these various perspectives and issues would be prohibitive.

Figure 1.2: Schematic description of the optical theorem Equation 1.12 as applied to a $2 \rightarrow 2$ process. On the right hand side one must sum over all possible intermediate states, including both various state labels X and their phase space Π_X.

for some definitions of these physical parameters, and compute everything at tree-level. However, this does not constitute a sensible field theory, as the optical theorem tells us this is not consistent. Defining $S = 1 + i\mathcal{M}$ as the S-matrix, then demanding unitarity $S^\dagger S = 1$ tells us the nontrivial part must satisfy

$$i\left(\mathcal{M}^\dagger - \mathcal{M}\right) = \mathcal{M}^\dagger \mathcal{M}. \tag{1.12}$$

Sandwiching this operator equation between initial and final states, we find that the left hand side is the imaginary part of the amplitude $\mathcal{M}(i \rightarrow f)$, which is nonzero solely due to loops. This is depicted schematically in Figure 1.2. We can see why this is by examining a scalar field propagator. Taking the imaginary part one finds

$$\text{Im}\,\frac{1}{p^2 + m^2 + i\epsilon} = \frac{-\epsilon}{(p^2 + m^2)^2 + \epsilon^2}. \tag{1.13}$$

This vanishes manifestly as $\epsilon \rightarrow 0$ except for when $p^2 = -m^2$, and an integral to find the normalization yields

$$\text{Im}\,\frac{1}{p^2 + m^2 + i\epsilon} = -\pi\delta(p^2 + m^2). \tag{1.14}$$

So internal propagators are real except for when the particle is put on-shell. In a tree-level diagram this occurs solely at some measure-zero set of exceptional external momenta, but in a loop-level diagram we integrate over *all* momenta in the loop, so an imaginary

14

part is necessarily present.

The lesson to take away from this is that classical field theories produce correlation functions with some particular momentum dependence, which can be essentially read off from the Lagrangian. But a consistent theory *requires* momentum dependence of a sort that does not appear in such a Lagrangian, which demands that calculations must include loops. In particular it is the analyticity properties of these higher-order contributions that are required by unitarity, and there is an interesting program to understand the set of functions satisfying those properties at each loop order as a way to bootstrap the structure of multi-loop amplitudes (see e.g. [46, 47, 48, 49, 50, 51]).

So far from being 'merely' a way to deal with seemingly unphysical predictions, renormalization is very closely tied to the physics. We begin in the next section with understanding its use for removing divergences, as this is the most basic application and is often the first introduction students receive to renormalization. We will then move on to discuss other, more physical interpretations of renormalization.

1.2.1   To Remove Divergences

Physical input is required

As a first pass, let's look again at a ϕ^4 theory

$$\mathcal{L} = -\frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}m_0^2\phi^2 - \frac{1}{4!}\lambda_0\phi^4, \tag{1.15}$$

and now treat it properly as a quantum field theory. As a simple example, let us consider $2 \to 2$ scattering in this theory, our discussion of which is particularly influenced by Zee [52]. At lowest-order this is extremely simple, and the tree-level amplitude is $i\mathcal{M}(\phi\phi \to \phi\phi) = -i\lambda_0$. But if we're interested in a more precise answer, we go to the next order in
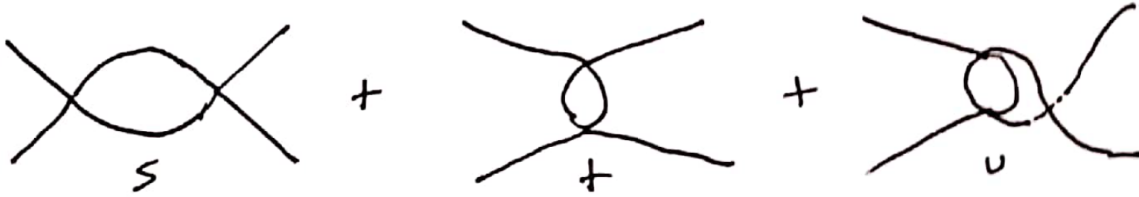
Figure 1.3: One-loop diagrams contributing to $2 \to 2$ scattering in the ϕ^4 theory. Since the external legs are all the same, these three different internal processes must be summed over, corresponding to momentum exchanged in the s, t, u channels, where these are the Mandelstam variables [53].

perturbation theory and we have the one-loop diagrams of Figure 1.3.

Defining $P_s^\mu \equiv p_1^\mu + p_2^\mu$ as the momentum flowing through the loop in the s-channel diagram, that diagram is evaluated as

$$i\mathcal{M}(\phi\phi \to \phi\phi)_{\text{one-loop s-channel}} = (-i\lambda_0)^2(-i)^2 \frac{1}{2} \int \frac{d^4k}{(2\pi)^4} \frac{1}{P_s^2 + m_0^2} \frac{1}{(P_s + k)^2 + m_0^2}. \quad (1.16)$$

Just from power-counting, we can already see that this diagram will be divergent. In the infrared, as $k \to 0$, the diagram is regularized by the mass of the field, but in the ultraviolet $k \to \infty$, the integral behaves as $\sim \int d^4k/k^4 \sim \int dk/k$ which is logarithmically divergent.

Though one might be tempted now to give up, we note that this divergence is appearing from an integral over very high energy modes—far larger than whatever energies we've verified our ϕ^4 model to, so let's try to ignore those modes and see if we can't get a sensible answer. The general term for removing these divergences is 'regularization' and we will here regularize (or 'regulate') this diagram by imposing a hard momentum cutoff Λ in Euclidean momentum space, which is the maximum energy of modes we let propagate in the loop. The loop amplitude may then be calculated with elementary methods detailed in, for example, Srednicki's textbook [25]. First we introduce Feynman parameters to combine the denominators, using $(AB)^{-1} = \int_0^1 dx(xA + (1-x)B)^{-2}$, which

16

here tells us

$$\frac{1}{P_s^2 + m_0^2} \frac{1}{(P_s + k)^2 + m_0^2} = \int_0^1 dx \left[x \left((P_s + k)^2 + m_0^2 \right) + (1 - x)(P_s^2 + m_0^2) \right]^{-2} \quad (1.17)$$

$$= \int_0^1 dx \left[q^2 + D \right]^{-2}, \quad (1.18)$$

where we've skipped the algebra letting us rewrite this with $q = k + xP_s$ and $D = x(1 - x)P_s^2 + m_0^2$. The change of variables $k \to q$ has trivial Jacobian, so the next step is to Wick rotate—Euclideanize the integral by defining $q^0 = i\bar{q}^d$, such that $q^2 \equiv q^\mu \eta_{\mu\nu} q^\nu = \bar{q}^\mu \delta_{\mu\nu} \bar{q}^\nu \equiv \bar{q}^2$. The measure simply picks up a factor of i, $d^d q = i d^d \bar{q}$, and we can then go to polar coordinates via $d^d \bar{q} = \bar{q}^{d-1} d\bar{q} d\Omega_d$. Lorentz invariance then means the angular integral gives us the area of the unit sphere in d dimensions, $\Omega_d = 2\pi^{d/2}/\Gamma(d/2)$, where $\Omega_4 = 2\pi^2$, and the radial integral becomes

$$\int_0^\Lambda d\bar{q} \, \bar{q}^3 \left[q^2 + D \right]^{-2} = \frac{1}{2} \left[\frac{D}{\bar{q}^2 + D} + \log \left(\bar{q}^2 + D \right) \right] \Bigg|_0^\Lambda \quad (1.19)$$

$$= -\frac{1}{2} \left[\frac{\Lambda^2}{\Lambda^2 + D} + \log \frac{D}{\Lambda^2 + D} \right]. \quad (1.20)$$

In fact it is possible to do the x integral analytically here, but we'll take $\Lambda^2 \gg |P^2| \gg m^2$ to find a simple answer

$$-\frac{1}{2} \int_0^1 dx \left[\frac{\Lambda^2}{\Lambda^2 + D} + \log \frac{D}{\Lambda^2 + D} \right] = \frac{1}{2} \left(1 - \log \frac{P_s^2}{\Lambda^2} \right) + \dots \quad (1.21)$$

Now putting all that together and including all the diagrams up to one-loop, we get the form

$$\mathcal{M}(\phi\phi \to \phi\phi) = -\lambda_0 + C\lambda_0^2 \left[\log \left(\frac{\Lambda^2}{s} \right) + \log \left(\frac{\Lambda^2}{t} \right) + \log \left(\frac{\Lambda^2}{u} \right) \right] + \text{ subleading}, \quad (1.22)$$

17

where s, t, u are the Mandelstam variables and C is just a numerical coefficient. Now we see explicitly that the divergence has led to dependence of our amplitude on our regulator Λ. Of course this is problematic because we introduced Λ as a non-physical parameter, and it would not be good if our calculation of a physical low-energy observable depended sensitively on how we dealt with modes in the far UV. But let us try to connect this with an observable anyway. We note that the theory defined by the Lagrangian in Equation 1.15 can not yet be connected to an observable because we have not yet given a numerical value for λ_0. So let's imagine an experimentalist friend of ours prepares some ϕs and measures this scattering amplitude at some particular angles and energies corresponding to values of the Mandelstam variables s_0, t_0, u_0. They find some value λ_{phys}, which is a pure number. If our theory is to describe this measurement accurately, this tells us a relation between our parameters and a physical quantity

$$-\lambda_{\text{phys}} = -\lambda_0 + C\lambda_0^2 \left[\log\left(\frac{\Lambda^2}{s_0}\right) + \log\left(\frac{\Lambda^2}{t_0}\right) + \log\left(\frac{\Lambda^2}{u_0}\right) \right] + \mathcal{O}(\lambda^3). \qquad (1.23)$$

This is known as a 'renormalization condition' which tells us how to relate our quantum field theories to observations at non-trivial loop order. Since the left hand side is a physical quantity, it may worry us that the right hand side contains a non-physical parameter Λ. But we still haven't said what λ_0 is, so perhaps we'll be able to find a sensible answer if we choose $\lambda_0 \equiv \lambda(\Lambda)$ in a correlated way with our regularization scheme. We call this 'promoting λ to a running coupling' by changing from the 'bare coupling' λ_0 to one which depends on the cutoff. So let's solve for λ in terms of λ_{phys}

and Λ. Rearranging we have

$$\lambda_0 = \lambda_{\text{phys}} + C\lambda_0^2 \left[\log\left(\frac{\Lambda^2}{s_0}\right) + \log\left(\frac{\Lambda^2}{t_0}\right) + \log\left(\frac{\Lambda^2}{u_0}\right)\right] + \mathcal{O}(\lambda_0^3) \qquad (1.24)$$

$$\lambda(\Lambda) = \lambda_{\text{phys}} + C\lambda_{\text{phys}}^2 \left[\log\left(\frac{\Lambda^2}{s_0}\right) + \log\left(\frac{\Lambda^2}{t_0}\right) + \log\left(\frac{\Lambda^2}{u_0}\right)\right] + \mathcal{O}(\lambda_{\text{phys}}^3) \qquad (1.25)$$

where in the second line the replacement $\lambda_0^2 \mapsto \lambda_{\text{phys}}^2$ modifies the right side only at higher-order and so that is absorbed into our $\mathcal{O}(\lambda^3)$ uncertainty. To see what this has done for us, let us plug this back into our one-loop amplitude Equation 1.22. This will impose our renormalization condition that our theory successfully reproduces our experimentalist friend's result. We find

$$\mathcal{M}(\phi\phi \to \phi\phi) = -\lambda_{\text{phys}} - C\lambda_{\text{phys}}^2 \left[\log\left(\frac{\Lambda^2}{s_0}\right) + \log\left(\frac{\Lambda^2}{t_0}\right) + \log\left(\frac{\Lambda^2}{u_0}\right)\right] \qquad (1.26)$$
$$+ C\lambda_{\text{phys}}^2 \left[\log\left(\frac{\Lambda^2}{s}\right) + \log\left(\frac{\Lambda^2}{t}\right) + \log\left(\frac{\Lambda^2}{u}\right)\right] + \mathcal{O}(\lambda_{\text{phys}}^3)$$

where again we liberally shunt higher-order corrections into our uncertainty term. Taking advantage of the nice properties of logarithms, we rearrange to get

$$\mathcal{M}(\phi\phi \to \phi\phi) = -\lambda_{\text{phys}} - C\lambda_{\text{phys}}^2 \left[\log\left(\frac{s}{s_0}\right) + \log\left(\frac{t}{t_0}\right) + \log\left(\frac{u}{u_0}\right)\right] + \mathcal{O}(\lambda_{\text{phys}}^3). \quad (1.27)$$

We see that our renormalization procedure of relating our theory to a physical observable has enabled us to write the full amplitude in terms of *physical* quantities, and remove the divergence entirely. This one physical input has enabled us to fully predict any $2 \to 2$ scattering in this theory.

We thus see how the renormalization procedure removes the divergences in a naïve formulation of a field theory and allows us to make finite predictions for physical observables. While we did need to introduce a regulator, once we make the replacement

$\lambda_0 \to \lambda(\Lambda)$ as defined in Equation 1.25 (and similar replacements for the coefficients of the other operators), the one-loop divergences are gone. We are guaranteed that any one-loop correlation function we calculate is finite in the $\Lambda \to \infty$ limit, which removes the regulator. If we wanted to increase our precision and calculate now at two loops, we would first renormalize the theory at two loops analogously to the above, and would find a more precise definition for $\lambda(\Lambda)$ which included terms of order $\mathcal{O}(\lambda^3_{\mathrm{phys}})$. At each loop order, replacing the bare couplings with running couplings suffices to entirely rid the theory of divergences.

Renormalizability

An important question is for which quantum field theories do a finite set of physical inputs will allow the theory to be fully predictive, in analogy to the example above. Such a theory is called 'renormalizable' and means that after some finite number of experimental measurements, we can predict any other physics in terms of those values. Were this not the case, and no finite number of empirical measurements would fix the theory, it would not be of much use. Within the context of perturbation theory, a theory will be renormalizable if its Lagrangian contains solely relevant and marginal operators, and indeed for our ϕ^4 theory three renormalization conditions are needed—one for each such operator.

The simplest way to understand why we must restrict to relevant and marginal operators is that irrelevant operators inevitably lead to the generation of a tower of more-and-more irrelevant operators. To see this, imagine now including a ϕ^6 interaction, as depicted in Figure 1.4a. At one loop this leads to a $4 \to 4$ scattering process with the same sort of divergence we saw in our previous loop diagram. So this loop is probing the UV physics, but we cannot absorb the unphysical divergence into a local interaction in our Lagrangian unless we now include a ϕ^8 term. But then we can draw a similar

20

(a)

(b)
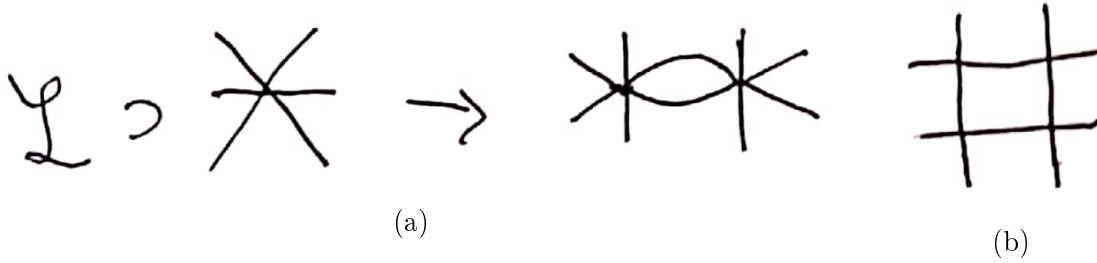
Figure 1.4: In (a), a pictoral representation of the nonrenormalizability of theories described by Lagrangians with irrelevant operators. A tree-level six-point coupling leads to a new divergence in $4 \to 4$ scattering, which must be absorbed into a renormalized eight-point coupling, which would then beget a divergent $6 \to 6$ amplitude .... Note that this does not mean there are not one-loop contributions to $4 \to 4$ scattering. Such a diagram is depicted in (b), but it is finite and does not require additional local operators.

one-loop diagram with the ϕ^8 interaction which will require a ϕ^{10} interaction, and so on. Note that in our ϕ^4 theory we also have $4 \to 4$ scattering at one loop, seen in Figure 1.4b, but there it comes from a box diagram which is finite, and so there is no need to include more local operators.

However, we emphasized above that the most useful description of a system depends on the precision at which one wishes to measure properties of the system. Thus in the study of effective field theories a broader definition of renormalizability should be used. For a theory with cutoff Λ, one decides to work to precision $\mathcal{O}(E/\Lambda)^n$ where E is a typical energy scale of a process and n is some integer. There are then a finite number of operators which contribute to processes to that precision—only those up to scaling dimension n—and so there is a notion of 'effective renormalizability' of the theory. We still require solely a finite number of inputs to set the behavior of the theory to whatever precision we wish, but such a theory nevertheless fails the original criterion, which may be termed 'power counting renormalizability' in comparison.

Wilsonian renormalization of ϕ^4

Above we characterized our cutoff as an unphysical parametrization of physics at high scales that we do not know and we found that its precise value dropped out of our physically observable amplitude. To some extent this is rather surprising, as it's telling us that the high energy modes in our theory have little effect on physics at long distances—we can compensate for their effects by a small shift in a coupling. We can gain insight into the effects of these high energy modes by taking the cutoff seriously and looking at what happens when the cutoff is lowered. This brilliant approach due to Wilson [54] is aimed at providing insight as to the particular effects of these high-energy modes by integrating out 'shells' of high-energy Euclidean momenta and looking at the low-energy effects. This discussion is closely inspired by that in Peskin & Schroeder's chapter 12 [55], as well as Srednicki's chapter 29 [25].

It is easiest to see how to implement this by considering the path integral formulation. We can equally well integrate over position space paths as over momentum modes:

$$Z = \int \mathcal{D}\phi(x)e^{i\int \mathcal{L}(\phi)} = \left(\prod_k \int \mathrm{d}\phi(k)\right)e^{i\int \mathcal{L}(\phi)}, \qquad (1.28)$$

and here it is clear that we may integrate over particular momentum modes separately if we so choose. In the condensed matter application in which Wilson originally worked, a cutoff appears naturally due to the lattice spacing a giving an upper bound on momenta $k_{\max} \sim 1/a$. In a general application we can imagine defining the theory with a fundamental cutoff Λ by including in the path integral only modes with Euclidean momentum $k^2 \leq \Lambda^2$.[5] The theory is defined by the values of the parameters in the theory with that

---

[5]It is necessary to define this cutoff in Euclidean momentum space for the simple fact that in Lorentzian space a mode with arbitrarily high energy k^0 may have tiny magnitude by being close to the light cone $|k^0| \simeq |\vec{k}|$. It is left as an exercise for the reader to determine what deep conclusion should be taken away from the fact that we perform all our QFT calculations in Euclidean space.

cutoff—our familiar relevant and marginal operators $m^2(\Lambda), \lambda(\Lambda)$ and in principle all of the 'Wilson coefficients' of irrelevant operators as well, since the theory is manifestly finite. The idea is to effectively lower the cutoff by explicitly integrating out modes with $b\Lambda \leq k \leq \Lambda$ for some $b < 1$. This will leave us with a path integral over modes with $k^2 \leq b^2\Lambda^2$—which is a theory of the same fields now with a cutoff $b\Lambda$. By integrating out the high-energy modes we'll be able to track precisely their effects in this low-energy theory.

Peskin & Schroeder perform this path integral explicitly by splitting the high energy modes into a different field variable and quantizing it, but since we've already introduced the conceptual picture of integrating fields out we'll take the less-involved approach of Srednicki. To repeat what we discussed above, the diagrammatic idea of integrating out a field is to remove it from the spectrum of the theory and reproduce its effects on low-energy physics by modifying the interactions of the light fields. Performing the path integral over some fields does not change the partition function, so the physics of the other fields must stay the same. We want to do the same thing here, but integrate out solely the high energy modes of a field and reproduce the physics in terms of the light modes.

We'll continue playing around with ϕ^4 theory and define our (finite!) theory with a cutoff of Λ, which in full generality looks like:

$$\mathcal{L}(\Lambda) = -\frac{1}{2}Z(\Lambda)\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}m^2(\Lambda)\phi^2 - \frac{1}{4!}\lambda(\Lambda)\phi^4 - \sum_{n=3}^{\infty}\frac{c_n(\Lambda)}{(2n)!}\phi^{2n}. \tag{1.29}$$

For simplicity we will decree that at our fundamental scale Λ we have a canonically normalized field $Z(\Lambda) = 1$ and no irrelevant interactions $c_n(\Lambda) = 0$, but just some particular $m^2(\Lambda)$ and $\lambda(\Lambda)$.

Let's look first at the one-loop four-point amplitude, which we must ensure is the

same in both the theory with cutoff Λ and that with cutoff $b\Lambda$. In the original theory, the amplitude at zero external momentum is

$$iV_4^\Lambda(0,0,0,0) = -i\lambda(\Lambda) + \frac{3}{2}(-i\lambda(\Lambda))^2 \int_{|k|<\Lambda} \frac{d^4k}{(2\pi)^4} \frac{(-i)^2}{(k^2+m^2(\Lambda))^2} + \mathcal{O}(\lambda^3) \qquad (1.30)$$

When we evaluate this in the theory with a lowered cutoff $b\Lambda$, the modification is simply to everywhere make the replacement $\Lambda \mapsto b\Lambda$. In order for the physics to remain the same without the high-energy modes, the vertex function must not change. We'll take full advantage of the perturbativity of the result—that is, $\lambda(\Lambda)-\lambda(b\Lambda) \sim \mathcal{O}(\lambda^2)$, $m^2(\Lambda) - m^2(b\Lambda) \sim \mathcal{O}(\lambda^2)$—to swap out quantities evaluated at $b\Lambda$ in the second-order term for those evaluated at Λ at the cost solely of higher-order terms which we ignore.

$$0 \equiv V_4^\Lambda(0,0,0,0) - V_4^{b\Lambda}(0,0,0,0) \qquad (1.31)$$

$$= -\lambda(\Lambda) + \frac{3}{2}\lambda(\Lambda)^2 \int_{|\bar{k}|<\Lambda} \frac{d^4\bar{k}}{(2\pi)^4} \frac{1}{(\bar{k}^2+m^2(\Lambda))^2}$$

$$+ \lambda(b\Lambda) - \frac{3}{2}\lambda(b\Lambda)^2 \int_{|\bar{k}|<b\Lambda} \frac{d^4\bar{k}}{(2\pi)^4} \frac{1}{(\bar{k}^2+m^2(b\Lambda))^2}$$

$$= -\lambda(\Lambda) + \lambda(b\Lambda) + \frac{3}{2}\lambda(\Lambda)^2 \int_{|\bar{k}|=b\Lambda}^{\Lambda} \frac{d^4\bar{k}}{(2\pi)^4} \frac{1}{(\bar{k}^2+m^2(\Lambda))^2} + \mathcal{O}(\lambda^3)$$

$$= -\lambda(\Lambda) + \lambda(b\Lambda) + \frac{3}{16\pi^2}\lambda(\Lambda)^2 \log\left(\frac{1}{b}\right) + \mathcal{O}(\lambda^3, \frac{m^2}{\Lambda^2})$$

$$\Rightarrow \lambda(b\Lambda) = \lambda(\Lambda) - \frac{3}{16\pi^2}\lambda(\Lambda)^2 \log\left(\frac{1}{b}\right) + \dots \qquad (1.32)$$

The effect of high energy modes on the four-point vertex function can be simply absorbed into a shift in the coupling constant! This procedure explicitly transfers loop-level physics in the theory defined at Λ into tree-level physics at $b\Lambda$.

We can repeat this for the two point function to find the behavior of $Z(\Lambda), m^2(\Lambda)$.

$$0 \equiv \Sigma^\Lambda(p) - \Sigma^{b\Lambda}(p) \tag{1.33}$$

$$= Z(\Lambda)p^2 + m^2(\Lambda) + \lambda(\Lambda) \int_{|\bar{k}|<\Lambda} \frac{d^4\bar{k}}{(2\pi)^4} \frac{1}{(Z(\Lambda)\bar{k}^2 + m^2(\Lambda))}$$

$$- Z(b\Lambda)p^2 - m^2(b\Lambda) - \lambda(b\Lambda) \int_{|\bar{k}|<b\Lambda} \frac{d^4\bar{k}}{(2\pi)^4} \frac{1}{(Z(b\Lambda)\bar{k}^2 + m^2(b\Lambda))}$$

$$= [Z(\Lambda) - Z(b\Lambda)]\, p^2 + m^2(\Lambda) - m^2(b\Lambda) + \lambda(\Lambda) \int_{|\bar{k}|=b\Lambda}^\Lambda \frac{d^4\bar{k}}{(2\pi)^4} \frac{1}{(Z(\Lambda)\bar{k}^2 + m^2(\Lambda))}$$

$$= [Z(\Lambda) - Z(b\Lambda)]\, p^2 + \left[m^2(\Lambda) - m^2(b\Lambda)\right] + \frac{\lambda(\Lambda)}{16\pi^2} \left[\Lambda^2 - b^2\Lambda^2\right] + \lambda(\Lambda)\frac{m^2(\Lambda)}{8\pi^2} \log\left(\frac{1}{b}\right)$$

$$\Rightarrow Z(b\Lambda) = Z(\Lambda) + \mathcal{O}(\lambda^2) \tag{1.34}$$

$$\Rightarrow m^2(b\Lambda) = m^2(\Lambda) + \frac{\lambda(\Lambda)}{16\pi^2}\Lambda^2 \left[1 - b^2\right] - \lambda(\Lambda)\frac{m^2(\Lambda)}{8\pi^2} \log\left(\frac{1}{b}\right) + \ldots \tag{1.35}$$

We have again liberally ignored subleading terms. We see that the wavefunction normalization Z does not run at one-loop in this theory, since the only one-loop diagram contributing to the two-point function does not have external momentum routed through the loop. This is merely 'accidental' as Z is not symmetry-protected and does run at two-loops. We also see the first hints of a somewhat worrisome situation with scalar masses. The mass $m^2(\Lambda)$ receives large one-loop corrections which tend to raise the mass up to near the cutoff, regardless of whether we originally had $m^2(\Lambda) \ll \Lambda^2$. We will investigate this in great detail later.

Now imagine we want to measure some properties of ϕ particles with external momenta far below our fundamental cutoff $p_i \sim \mu \ll \Lambda$. By construction, our procedure of integrating out high-energy momentum modes keeps the physics of these low-energy particles the same. But if we calculate this scattering amplitude using $\mathcal{L}(\Lambda)$, it is not easy to see from the Lagrangian how these low-energy modes will behave, since important effects are hidden in loop diagrams. If we instead first integrate out momentum shells

down to some $\mathcal{L}(b\Lambda)$ with $\mu < b\Lambda \ll \Lambda$, then the effects of the high energy modes have been absorbed into the parameters of our Lagrangian, and we can read off much more of how ϕ particles will behave at low energies simply by looking at the parameters.

We can see further value in this approach if we consider scattering more low-energy ϕ. Let's look at the 6-point vertex function at zero momentum—in the theory at Λ, we start with $c_6(\Lambda) = 0$ and a one-loop diagram where momenta up to Λ run in the loop:

$$V_6^\Lambda = \frac{\lambda^3(\lambda)}{48} \int_{|k|<\Lambda} \frac{d^4k}{(2\pi)^4} \left(\frac{1}{Z(\Lambda)k^2 + m^2(\Lambda)} \right)^3. \tag{1.36}$$

Now in the theory at $b\Lambda$, the loop only contains momenta up to $b\Lambda$, so we must account for the difference with a contact interaction $c_6(b\Lambda)$:

$$V_6^{b\Lambda} = c_6(b\Lambda) + \frac{\lambda^3(b\lambda)}{48} \int_{|k|<b\Lambda} \frac{d^4k}{(2\pi)^4} \left(\frac{1}{Z(b\Lambda)k^2 + m^2(b\Lambda)} \right)^3. \tag{1.37}$$

Again we should ensure that the physics is the same upon lowering the cutoff:

$$0 \equiv V_6^\Lambda - V_6^{b\Lambda} \tag{1.38}$$

$$c_6(b\Lambda) = \frac{\lambda^3(\lambda)}{48} \int_{b\Lambda}^{\Lambda} \frac{d^4k}{(2\pi)^4} \left(\frac{1}{Z(\Lambda)k^2 + m^2(\Lambda)} \right)^3 \tag{1.39}$$

$$= \frac{\lambda(\Lambda)^3}{3 \times 256\pi^2} \left(\frac{1}{(b\Lambda)^2} - \frac{1}{\Lambda^2} \right) + \ldots \tag{1.40}$$

So this renormalization procedure is especially useful for understanding the behavior of irrelevant interactions. In our original theory nothing about the six-particle interaction was obvious from the Lagrangian, but in our theory with cutoff $b\Lambda$ we can simply read off the strength of this interaction at lowest order.

Note also the inverse behavior to that of the mass corrections—for the irrelevant interaction, the most significant contributions to the infrared behavior come from the

low-energy part of the loop integral, and the UV contributions are suppressed relative to this. Similarly, if we had started with a nonzero $c_6(\Lambda)$ which was small in units of the cutoff $c_6(\Lambda) \ll \Lambda^{-2}$ (so perturbative), such a UV contribution will also be subdominant. Then fully generally here, we have

$$c_6(b\Lambda) \simeq \frac{\lambda(\Lambda)^3}{16\pi^2} \frac{1}{b^2\Lambda^2} + \mathcal{O}(b^0) \tag{1.41}$$

as we evolve to low scales $b \ll 1$.

The Wilsonian approach we have discussed here gives useful intuition for how renormalization works as a coarse-graining procedure wherein one changes the fundamental 'resolution' of the theory, but in practice can make calculations cumbersome. Furthermore, the hard momentum-space cutoff we used is not gauge-invariant, which causes difficulties in applications.

The benefit, however, is that this is a 'physical renormalization scheme' in which the renormalization condition relates the bare parameters to physical observables. For this reason, this renormalization scheme satisfies the Appelquist-Carrazone decoupling theorem [56], which is enormously powerful. This guarantees for us that the effects of massive fields can, at low energies, simply be absorbed into modifications of the parameters in an effective theory containing solely light fields.

In the next section we'll return to a clarifying example of the meaning of the decoupling theorem, and also discuss a renormalization scheme which does not satisfy the requirements for this theorem to operate, but is far simpler to use for calculations. The winning strategy will be to input decoupling by hand, which will allow us to get sensible physical results without the computational difficulty. Before we get to that, though, we'll take a couple detours.

Renormalization and locality

We have seen that the need to remove divergences in our theory led to the introduction of running couplings which change as a function of scale. In our example above we see that renormalization has the operational effect of transferring loop-level physics into the tree-level parameters. This is an interesting perspective which bears further exploration—if there is hidden loop-level physics that really has the same physical effects as the tree-level bare parameters, perhaps this is a sign that there is a better way to organize our perturbation theory. Indeed, at some very general level renormalization can be thought of as a method for improving the quality of perturbation theory. For useful discussions at this level of abstraction of how renormalization operates, see [57] for its natural appearance whenever infinities are encountered in naïve perturbative calculations, and [58] for its usefulness even when infinities are not present. We'll discuss this perspective on renormalization further in the next section.

However it's clear that loops also give rise to physics that is starkly different from the lowest-order result (e.g. non-trivial analytic structure), so how do we know what higher-order physics we *can* stuff into tree-level? In a continuum quantum field theory, a Lagrangian is a *local* object $\mathcal{L}(x)$—that is, it contains operators like $\phi(x)^3$ give an interaction between three ϕ modes at a single spacetime point x. Such effects are known as 'contact interactions', but even at tree-level a local Lagrangian can clearly produce non-local (that is, long-range) physics effects. For example, consider the amplitude for $2 \to 2$ scattering in a ϕ^3 theory at second order in the coupling.

In position space the non-locality here is obvious, as in Figure 1.5: A simple tree-level diagram corresponds to a particle at point x and a particle at point y exchanging a ϕ quantum, but one may forget this important fact when working in momentum space.
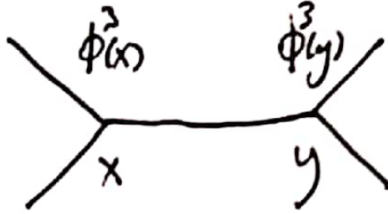
Figure 1.5: A position space Feynman diagram with vertices labeled.

There the result is

$$i\mathcal{M}(\phi\phi \to \phi\phi) = \frac{ig^2}{p^2 + m^2}, \qquad (1.42)$$

and indeed, Fourier transforming the cross-section for this process yields a Yukawa scattering potential for our ϕ particles, showing that they mediate a long-range force over distances $r \sim 1/m$. We obviously cannot redefine the Lagrangian to put this effect into the lowest order of perturbation theory since this is not a local effect.

But if we do have a continuum quantum field theory, then because of locality it describes fluctuations on *all* scales. When we go to loop-level, we must integrate over all possible internal states, which includes integrating over arbitrarily large momenta or equivalently fluctuations on arbitrarily small scales. Heuristically, when the loop integral in sensitive to the ultraviolet of the theory, it is computing effects that operate on all scales—that is, it gives a contribution to *local* physics. This tells us that the pieces of this higher-order contribution which we can reshuffle into our Lagrangian are connected with ultraviolet sensitivity, leading to a close connection of renormalization with divergences.

A couple notes are warranted about the notion of locality we rely on here. Firstly, it's clear that this criterion of local effects appears because we began with a local Lagrangian. If we postulated that our fundamental theory contained nonlocal interactions, say $\mathcal{L}(x) \supset g\phi(x)\left(\int dy\phi(y)\right)\left(\int dz\phi(z)\right)$, then we could clearly absorb further nonlocalities with the same structure into this coupling as well. However, this sort of nonlocality is different

from the nonlocality we saw appearing out of the local theory at tree-level. In particular it would break the standard connection between locality and the analytic structure of amplitudes—see e.g. Schwartz [26] or Weinberg [59] on 'ppology' and locality.

Secondly, our notion of locality should be modified in a low-energy theory with an energy-momentum cutoff Λ, as can be seen in hindsight in our Wilsonian discussion above. As Λ defines a maximum energy scale we can probe, there is equivalently a minimum time and length scale we can probe due to the uncertainty principle, heuristically $\Delta x^\mu \gtrsim 1/\Delta p^\mu \gtrsim 1/\Lambda$. As a result, any fluctuations on shorter length scales are effectively local from the perspective of the low-energy theory. An exchange of a massive field with $M > \Lambda$ or of a light field with high frequency $\omega > \Lambda$ appears instantaneous to low-energy observers. This explains how it's sensible to use renormalization techniques in, for example, condensed matter applications, where systems are fundamentally discrete.

We can see this concretely by imagining the light ϕs in the tree-level example above instead exchanged a heavier scalar Φ with mass $M \gg m$. While the amplitude $\mathcal{M} = g^2/(p^2 + M^2)$ is still nonlocal in the continuum theory, if we're only interested in physics at energies $E \ll M$ we may Taylor expand the result $\mathcal{M} = g^2/M^2 + g^2 p^2/2M^4 + \ldots$. We may then absorb the leading effects of this heavy scalar into an effectively local interaction $g^2 \phi^4/M^2$ among the light fields—as long as we work at energies below M.

In the next section we'll explore concretely how these insights enable us to transfer loop-level physics to tree-level physics, and so improve our calculations.

1.2.2   To Repair Perturbation Theory

Renormalization group equations

The astute reader may notice a potential issue with our one-loop results in the ϕ^4 theory, Equations 1.32, 1.34, 1.35. We've derived this behavior as the first subleading

terms in a series expansion in the number of loops. In relation to the tree-level results, the one-loop contributions are suppressed by a factor $\sim \lambda/16\pi^2 \log(\Lambda_0/\Lambda)$, where I've switched notation to now have Λ_0 as a high scale and Λ as the lower scale we integrate down to. Higher n-loop contributions will be further suppressed by n loop factors. But what if we wanted to study physics at a scale far lower than Λ_0? Eventually this factor becomes large enough that we will need to compute many loops to obtain high precision, and then large enough that we have reason to question the convergence of the series[6]. Keep in mind that we are in the era of precision measurements of the Standard Model, so these one-loop expressions are very restrictive.

For a concrete example, say we wanted to check the SM prediction for a measurement of a coupling $\lambda(\Lambda)$ with $\lambda(\Lambda_0) = 1$ whose experimental uncertainty was 1%. Let's define a theoretical uncertainty on a perturbative calculation to n^{th} order as

$$\epsilon_n \equiv \frac{n^{\text{th}} \text{ order result} - \text{ estimated size of } (n+1)^{\text{th}} \text{ order result}}{n^{\text{th}} \text{ order result}}, \qquad (1.43)$$

where our Wilsonian calculation in Section 1.2.1 gave at first order, as a reminder (and with modified notation)

$$\lambda(\Lambda) = \lambda(\Lambda_0) - \frac{3}{16\pi^2}\lambda(\Lambda_0)^2 \log \frac{\Lambda_0}{\Lambda} + \mathcal{O}(\lambda^3), \qquad (1.44)$$

and our heuristic estimate for the size of the second order correction is $(\frac{3}{16\pi^2})^2 \lambda(\Lambda_0)^3 \log^2 \frac{\Lambda_0}{\Lambda}$. When the result is simply a series, the uncertainty is very simple to calculate, as the nu-

---

[6]Please excuse my slang. Perturbative series in QFT are quite generally *not* convergent but we can trust the answers anyway to order $n \sim \exp 1/\text{expansion parameter}$ because they are *asymptotic* series. So really when this parameter becomes large enough we worry that our series is not even asymptotic. Thinking deeply about this leads to many interesting topics in field theory, from accounting for nonpertubative instanton effects which are (partially) behind the lack of convergence; to the program of 'resurgence', the idea that there are secret relations between the perturbative and nonperturbative pieces. This is all far outside my purview here, but some introductions aimed at a variety of audiences can be found in [60, 61, 62, 63, 64, 65].

merator is then our estimate of the $(n+1)^{\text{th}}$ order correction, which is roughly the square of the first order correction. Here we have $\epsilon_1 = \frac{3}{16\pi^2}\lambda(\Lambda_0)\log\frac{\Lambda_0}{\Lambda}$.

Then in order for our theoretical uncertainty to be subdominant to the experimental precision, $\epsilon_1 < 0.01$, we must go past the one-loop result if we wish to look at energies below $\Lambda \sim \exp(-16\pi^2 \times 0.01/3)\Lambda_0 \sim \Lambda_0/2$, the two-loop result is only sufficient until $\Lambda \sim \Lambda_0/400$, and if we can manage to calculate the three-loop corrections that only gets us down to something like $\Lambda \sim \Lambda_0/8 \times 10^4$. If we're interested in taking the predictions of a Grand Unified Theory defined at $\Lambda_0 \sim 10^{16}$ GeV and comparing them to predictions at SM energies, how in the world are we to do so?

Fortunately, we can do better by applying our one-loop results more cleverly. It is clear by looking at Equations 1.32,1.34,1.35 that the results have the same form no matter the values of Λ_0, Λ. So if we take Λ to be only slightly smaller than Λ_0 (corresponding to $1-b \ll 1$ in our previous notation) the expansion parameter becomes very small and the one-loop result becomes very trustworthy. What we would like is some sort of iterative procedure to gradually lower the cutoff, which we could then use to find the one-loop result for energies far lower than the range of our perturbative series. This is in fact precisely the sort of problem that a differential equation solves, and we can derive such an equation by differentiating both sides by $\ln\Lambda$ and then taking Λ infinitesimally close to Λ_0. That exercise yields

$$\frac{\mathrm{d}Z(\Lambda)}{\mathrm{d}\log\Lambda} = 0 \tag{1.45}$$

$$\frac{\mathrm{d}m^2(\Lambda)}{\mathrm{d}\log\Lambda} = \frac{\lambda(\Lambda)}{8\pi^2}\left(m^2(\Lambda) - \Lambda^2\right) \tag{1.46}$$

$$\frac{\mathrm{d}\lambda(\Lambda)}{\mathrm{d}\log\Lambda} = \frac{3}{16\pi^2}\lambda(\Lambda)^2 \tag{1.47}$$

These are known as 'renormalization group equations' and they indeed allow us to evolve

the coupling down to low energies—one says we use them to 'resum the logarithm'. Then to study physics at a very low scale we can bring these couplings down to a lower scale and do our loop expansion using those couplings, which is known as 'renormalization group improved perturbation theory', and which we will discuss in more detail soon. Explicitly solving with our boundary condition at Λ_0 yields

$$\lambda(\Lambda) = \frac{\lambda(\Lambda_0)}{1 + \frac{3\lambda(\Lambda_0)}{16\pi^2} \log \frac{\Lambda_0}{\Lambda}} \tag{1.48}$$

Turning back to our effective field theory language, we see that quantum corrections have generated an anomalous dimension for λ, $\delta_{\phi^4} = 3\lambda^2/(16\pi^2)$, correcting the leading order scaling behavior. Since $\delta_{\phi^4} > 0$, we've determined that the quartic interaction is marginally irrelevant, which we will return to later.

We can now look at the theoretical uncertainty in this one-loop resummed calculation by including an estimate of the next order correction to the running of the quartic $\frac{d\lambda(\Lambda)}{d\log\Lambda} = \frac{3}{16\pi^2}\lambda(\Lambda)^2 + \left(\frac{3}{16\pi^2}\right)^2 \lambda(\Lambda)^3$ and resumming that expression. This can no longer be done analytically, but numerical evaluation easily reveals that the theoretical uncertainty here stays below 1% for many, many orders of magnitude below Λ_0. Resumming the logarithmic corrections allows us to use our loop results to far greater effect.

Decoupling

The physical meaning and technical statement of the decoupling theorem commonly confuse even prominent practitioners of effective field theory, so it's worth going clearly through an example to refine our understanding. Indeed, one may be confused just at zeroth order about how decoupling is sensible against the background of the hierarchy problem—which is an issue of sensitivity of a scalar mass to heavy mass scales. How can we claim QFT obeys a decoupling theorem and then go on to worry at length about

quantum corrections $\delta m^2 \propto M^2$?

The correct way to think about the decoupling theorem is *not* whether a top-down calculation could yield a result that depends on heavy mass scales, but whether a bottom-up effective field theorist and low-energy observer could gain information about the heavy mass scales through low-energy measurements. We can clarify this important difference by looking at a one-loop mass correction to a light scalar ϕ of mass m from a heavy scalar Φ of mass M through the interaction $\lambda \phi^2 \Phi^2$. We again take a Wilsonian perspective and begin at a scale $\Lambda_0 > M$. In close analogy to what we had before, we now find

$$0 \equiv \Sigma^{\Lambda_0}(0) - \Sigma^{\Lambda}(0) \tag{1.49}$$

$$= m^2(\Lambda_0) + \lambda(\Lambda_0) \int_{|k| < \Lambda_0} \frac{d^4 k}{(2\pi)^4} \frac{1}{(k^2 + M^2(\Lambda_0))}$$

$$- m^2(\Lambda) - \lambda(\Lambda') \int_{|k| < \Lambda} \frac{d^4 k}{(2\pi)^4} \frac{1}{(k^2 + M^2(\Lambda))}$$

$$= m^2(\Lambda_0) - m^2(\Lambda) + \lambda(\Lambda_0) \int_{|k|=\Lambda}^{\Lambda_0} \frac{d^4 k}{(2\pi)^4} \frac{1}{(k^2 + M^2(\Lambda_0))}$$

$$= \left[m^2(\Lambda_0) - m^2(\Lambda) \right] + \frac{\lambda(\Lambda_0)}{16\pi^2} \left[\Lambda_0^2 - \Lambda^2 \right] - \lambda(\Lambda_0) \frac{M^2(\Lambda_0)}{16\pi^2} \log \left(\frac{\Lambda_0^2 + M^2(\Lambda_0)}{\Lambda^2 + M^2(\Lambda)} \right)$$

$$\Rightarrow m^2(\Lambda) = m^2(\Lambda_0) + \frac{\lambda(\Lambda_0)}{16\pi^2} \left(\left[\Lambda_0^2 - \Lambda^2 \right] + M^2(\Lambda_0) \log \frac{\Lambda_0^2 + M^2(\Lambda_0)}{\Lambda^2 + M^2(\Lambda)} \right) + \mathcal{O}(\lambda^2). \tag{1.50}$$

And we may already exhibit the confusion. If we use this to calculate the mass at a low scale $\Lambda \ll M$, we see that $m^2(\Lambda)$ does depend on the heavy mass scale, and gets a contribution which goes like $m^2(\Lambda) \sim M^2(\Lambda_0) \log \left(1 + \Lambda_0^2 / M^2(\Lambda_0) \right)$.

However, the effect on the light scalar is an additive shift of the mass. If we go out and measure the mass at a single scale $m^2(\Lambda)$ we can't tell empirically which 'parts' of that came from $m^2(\Lambda_0)$ and which came from $M^2(\Lambda_0)$ or whatever else is in there, so we have no idea of how this low-energy measurement depends on heavy scales. To get information about the various contributions to the light scalar mass, we can measure it at

different scales and look at how it changes. Of course this information is contained in the renormalization group equation for $m^2(\Lambda)$. At $\mathcal{O}(\lambda)$, we can find this by differentiating the above, and we find

$$\frac{\mathrm{d}m^2(\Lambda)}{\mathrm{d}\log\Lambda} = \frac{\lambda(\Lambda)}{8\pi^2}\Lambda^2\left[\frac{M^2(\Lambda)}{\Lambda^2 + M^2(\Lambda)} - 1\right]. \qquad (1.51)$$

Now we can see the difference. If we perform low-energy observations where we can take the cutoff below the mass of the heavy scalar $\Lambda \ll M$, then the physics of the heavy scalar decouples from the *running* of the light scalar mass. It is only by studying this running at low energies that we can gain information about the ultraviolet, and we see that this information is contained solely in small corrections scaling as Λ^2/M^2. At low energies, to learn about short-distance physics we must make very precise measurements of the low-energy physics. *This* is the sense in which heavy mass scales decouple from the theory in the infrared.

Renormalized perturbation theory

Now let us study another, slightly more complex theory and apply renormalization techniques to simplify our calculations. We avoid the complication of gauge symmetries and focus instead on a Yukawa theory of a Dirac fermion interacting with a parity-odd scalar.

Our first improvement to perturbation theory will be to switch from 'bare' to 'renormalized perturbation theory'. Let's first recap our procedure in Section 1.2.1. We began with a Lagrangian with bare parameters $m_0, \lambda_0, \dots$, introduced a regulator, computed the physical parameters $m_{\mathrm{phys}}, \lambda_{\mathrm{phys}}$ in terms of the bare ones, inverted those relationships, and then plugged in for the bare parameters in terms of the renormalized ones, after which we were left with an amplitude which remains finite as we remove the regu-

lator. This procedure works to remove the divergences in any renormalizable theory, but is obviously rather cumbersome.

Furthermore one may question the validity of performing a perturbative expansion in a bare parameter which we later discover is formally infinite in the continuum theory $\lambda_0 \sim \log(\Lambda_0^2) \to \infty$. It is both conceptually and computationally easier to instead start off by performing perturbation theory in terms of the renormalized parameters which we know to be finite by definition. Fortunately we can improve our accounting simply by reshuffling the Lagrangian as follows.

In terms of the bare parameters and fields, the Lagrangian reads

$$\mathcal{L}_0 = \bar{\psi}_0 \left(i \slashed{\partial} - M_0 \right) \psi_0 + \frac{1}{2} \phi_0 \left(\Box - m_0^2 \right) \phi_0 \tag{1.52}$$

$$\mathcal{L}_1 = i g_0 \phi_0 \bar{\psi}_0 \gamma_5 \psi_0 - \frac{1}{24} \lambda_0 \phi_0^4 \tag{1.53}$$

where we've split up the free and interaction parts. Just as in our earlier example, when we compute at one-loop these parameters will get corrections such that the bare parameters are no longer the physical parameters we measure. Anticipating that fact, let us rewrite the Lagrangian to explicitly account for those corrections from the outset.

Although it was not a feature of our simple example above, in general there will be 'wavefunction renormalization' which changes the normalization of our field operators, so we define $\phi_0 = Z_\phi^{1/2} \phi, \psi_0 = Z_\psi^{1/2} \psi$ where ψ, ϕ are now renormalized fields, We do the same to define renormalized masses related to the bare masses as $M_0 = Z_M Z_\psi^{-1} M, m_0^2 = Z_\phi^{-1} Z_m m^2$, and for the couplings $g_0 = Z_\phi^{-1/2} Z_\psi^{-1} Z_g g, \lambda_0 = Z_\phi^{-2} Z_\lambda \lambda$. Next we use the brilliant strategy of adding zero to split these Z-factors into a piece with the same form we started with and a 'counterterm' proportional to $(Z - 1)$. Since at tree-level there's no renormalization needed, we know $Z = 1 + \mathcal{O}(\text{couplings})$. At nontrivial loop level, we must choose the Z-factors to implement our chosen renormalization scheme.

The Lagrangian now takes the form

$$\mathcal{L}_0 = \bar{\psi}\left(i\slashed{\partial} - M\right)\psi + \frac{1}{2}\phi\left(\Box - m^2\right)\phi \tag{1.54}$$

$$\mathcal{L}_1 = iZ_g g\phi\bar{\psi}\gamma_5\psi - \frac{1}{24}Z_\lambda\lambda\phi^4 + \mathcal{L}_{\text{ct}} \tag{1.55}$$

$$\mathcal{L}_{\text{ct}} = i(Z_\psi - 1)\bar{\psi}\slashed{\partial}\psi - (Z_M - 1)M\bar{\psi}\psi - \frac{1}{2}(Z_\phi - 1)\partial^\mu\phi\partial_\mu\phi - \frac{1}{2}(Z_m - 1)m^2\phi^2 \tag{1.56}$$

where we've split off the counterterms into $\mathcal{L}_{\text{ct}}$. We can now treat the terms in $\mathcal{L}_{\text{ct}}$ simply as additional lines and vertices contributing to our Feynman diagrams. We'll see how useful this is once we begin renormalizing the theory. This is done in full in Srednicki's chapters 51-52 [25], so we will not go through every detail.

Continuum renormalization

We'll regulate this theory using 'dimensional regularization' (dim reg) which analytically continues the theory to general dimension $d = 4 - \epsilon$. That this will regulate our theory is not obvious, but I recommend Georgi [17] to convince yourself of this and Collins [66] for a full construction of dim reg; we'll content ourselves with seeing it in action. Our renormalization scheme will be 'modified minimal subtraction' and denoted $\overline{\text{MS}}$, where 'minimal subtraction' means we'll choose our counterterms solely to cancel off the divergent pieces (rather than to enforce some relation to physical observables, as we did previously) and 'modified' means that actually it's a bit nicer if we cancel off a couple annoying constants as well. Since we're using $\overline{\text{MS}}$, the mass parameters m, M will not quite be the physical masses, which are always the locations of the poles in the propagators, and the fields will not be normalized to have unity residue on those poles. So we'll have to relate these parameters to the physical ones later.

We'll briefly go through renormalizing the scalar two-point function at one loop to evince dim reg and $\overline{\text{MS}}$. In our one-loop diagrams we use propagators given by $\mathcal{L}_0$,
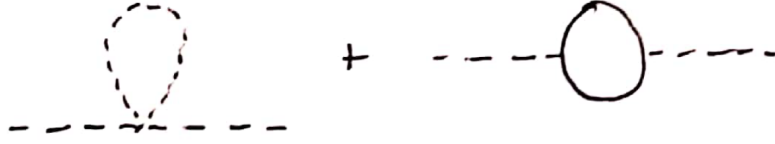
Figure 1.6: Diagrams giving the one-loop correction to the scalar propagator in Yukawa theory.

since we know the counterterms begin at higher order. The full details of the one-loop renormalization of this theory can be found in Srednicki's Chapter 51.

At one-loop, the scalar two-point function gets corrections due to both interactions, as seen in Figure 1.6. There is the diagram we had in the ϕ^4 theory above, but we must recompute this in dim reg

$$i\Sigma_{\phi \text{ loop}}(p^2) = -i\frac{1}{2}\lambda \int \frac{d^4 k}{(2\pi)^4} \frac{-i}{k^2 + m^2} \tag{1.57}$$

$$= -i\frac{1}{2}\lambda\tilde{\mu}^\epsilon \int \frac{d^d \bar{k}}{(2\pi)^d} \frac{1}{\bar{k}^2 + m^2} \tag{1.58}$$

$$= -i\frac{\lambda}{2}\Gamma\left(-1 + \frac{\epsilon}{2}\right) \frac{m^2}{(4\pi)^2} \left(\frac{4\pi\tilde{\mu}^2}{m^2}\right)^{\epsilon/2} \tag{1.59}$$

$$\simeq i\frac{\lambda}{2(4\pi)^2}m^2 \left(\frac{2}{\epsilon} + 1 - \gamma_E\right)\left(1 + \frac{\epsilon}{2}\log\frac{4\pi\tilde{\mu}^2}{m^2}\right) \tag{1.60}$$

$$\simeq i\frac{\lambda}{2(4\pi)^2}m^2 \left(\frac{2}{\epsilon} + 1 - \gamma_E - \log\frac{4\pi\tilde{\mu}^2}{m^2}\right) \tag{1.61}$$

$$= i\frac{\lambda}{(4\pi)^2}m^2 \left(\frac{1}{2} + \frac{1}{\epsilon} + \frac{1}{2}\log\frac{\mu^2}{m^2}\right) + \mathcal{O}(\epsilon), \tag{1.62}$$

where we have analytically continued to $d = 4 - \epsilon$ dimensions including replacing $\lambda \mapsto \lambda\tilde{\mu}^\epsilon$, with $\tilde{\mu}$ a mass scale, to keep λ dimensionless; performed the integral in general dimension; expanded for $\epsilon \simeq 0$; and defined $\mu^2 \equiv 4\pi\tilde{\mu}^2 e^{-\gamma_E}$, where γ_E is the Euler-Mascheroni constant, to simplify the expression. Details on these calculational steps are laid out in

Srednicki's Chapter 14. There's another diagram with a ψ loop

$$i\Sigma_{\psi \text{ loop}}(p^2) = (-1)(ig)^2 \int \frac{d^4k}{(2\pi)^4}(-i)^2 \frac{\text{Tr}\left[(-\slashed{k}+M)\gamma_5(-\slashed{k}-\slashed{p}+M)\gamma_5\right]}{(k^2+M^2)((k+p)^2+M^2)} \tag{1.63}$$

$$= -\frac{g^2}{4\pi^2}\left[\frac{1}{\epsilon}(k^2+2M^2) + \frac{1}{6}k^2 + M^2 - \int_0^1 dx (3x(1-x)k^2+M^2)\ln\frac{D}{\mu^2}\right], \tag{1.64}$$

with $D = x(1-x)k^2 + m^2$, whose evaluation follows similar steps but we skip for brevity. Adding these together, $\overline{\text{MS}}$ tells us the ϕ counterterms must take the values

$$Z_\phi = 1 - \frac{g^2}{4\pi^2}\frac{1}{\epsilon} \tag{1.65}$$

$$Z_m = 1 + \left(\frac{\lambda}{16\pi^2} - \frac{g^2}{2\pi^2}\frac{M^2}{m^2}\right)\frac{1}{\epsilon}. \tag{1.66}$$

For the fermion, evaluating the one-loop diagrams gives us the counterterms

$$Z_\psi = 1 - \frac{g^2}{16\pi^2}\frac{1}{\epsilon} \tag{1.67}$$

$$Z_M = 1 - \frac{g^2}{8\pi^2}\frac{1}{\epsilon}. \tag{1.68}$$

Since we didn't choose the counterterm to keep the location of the pole in the propagator fixed, m is no longer the physical scalar mass. But we can find the physical, 'pole' mass precisely from that condition:

$$0 \equiv \Delta^{-1}(k^2 = -m_{\text{phys}}^2) \tag{1.69}$$

$$= k^2 + m^2 + \Sigma(k^2)\big|_{k^2=-m_{\text{phys}}^2} \tag{1.70}$$

$$= -m_{\text{phys}}^2 + m^2 - \Sigma(-m_{\text{phys}}^2) \tag{1.71}$$

$$\Rightarrow m_{\text{phys}}^2 = m^2 - \Sigma(-m^2) + \mathcal{O}(\lambda^2, g^4, \lambda g^2), \tag{1.72}$$
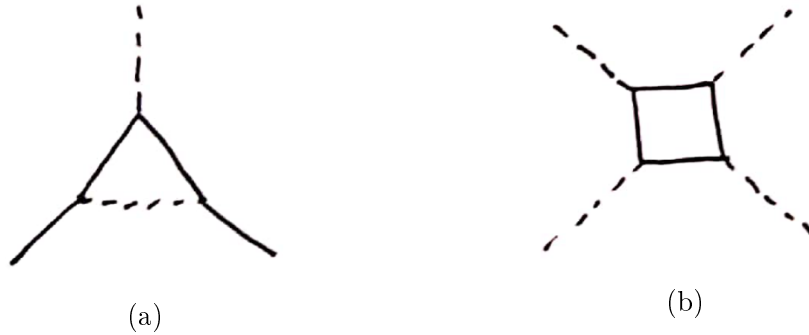
Figure 1.7: Some one-loop diagrams in Yukawa theory correcting the interactions. In (a), a triangle diagram correcting the Yukawa interaction. In (b), a box diagram correcting the scalar quartic interaction.

where we have used our favorite trick to replace m^2_{phys} with m^2 in the one-loop correction, since it is already higher order in couplings.

As for the interactions, we have a triangle diagram for the Yukawa coupling and a new contribution to the quartic with a fermion running in the loop, as depicted in Figures 1.7a and 1.7b respectively. These lead to the counterterms

$$Z_g = 1 + \frac{g^2}{8\pi^2}\frac{1}{\epsilon} \tag{1.73}$$

$$Z_\lambda = 1 + \left(\frac{3\lambda}{16\pi^2} - \frac{3g^4}{\pi^2\lambda}\right)\frac{1}{\epsilon}, \tag{1.74}$$

from which we'll be able to understand how the strength of the interactions changes as a function of the energy at which the theory is probed.

Renormalization group improvement

Now the second improvement to perturbation theory is the RG-improved perturbation theory we mentioned above. This takes on an even more useful role in our continuum renormalization scheme here. In the Wilsonian picture, Λ was a high cutoff and we ensured the physics was invariant under evolution of Λ, but this scale still needed to stay

far above the scales of interest in the problem $\Lambda \gg m, M, -k^2, \ldots$. Now the scale μ is entirely unphysical and we are free to bring it all the way down to the scales of kinematic interest—in fact doing so will vastly simplify calculations. As a result we are able to make even more use of the RG-improvement than we could above.

We find the running couplings by again using the fact that the bare parameters are independent of the unphysical renormalization scale μ. Having utilized a mass-independent regulator, a Wilsonian interpretation of couplings running with the value of the regulator is nonsensical here and so renormalization group improvement is the way to extract predictions from this theory. We already have the relations between the bare and renormalized quantities, e.g.

$$g_0 = Z_\phi^{-1/2} Z_\psi^{-1} Z_g g \tilde{\mu}^{\epsilon/2}. \tag{1.75}$$

And since we know that the bare parameters are independent of μ by definition, we have

$$\ln g_0 = \ln\left(1 + \frac{g^2}{8\pi^2}\frac{1}{\epsilon}\right) + \ln\left(1 + \frac{g^2}{16\pi^2}\frac{1}{\epsilon}\right) + \ln\left(1 + \frac{g^2}{8\pi^2}\frac{1}{\epsilon}\right) + \ln g + \frac{1}{2}\epsilon\ln\tilde{\mu} \tag{1.76}$$

$$= \frac{5g^2}{16\pi^2}\frac{1}{\epsilon} + \ln g + \frac{1}{2}\epsilon\ln\tilde{\mu} + \mathcal{O}(g^4) \tag{1.77}$$

$$\frac{\mathrm{d}\ln g_0}{\mathrm{d}\ln\mu} = 0 \tag{1.78}$$

$$= \frac{10g}{16\pi^2}\frac{1}{\epsilon}\frac{\mathrm{d}g}{\mathrm{d}\ln\mu} + \frac{\mathrm{d}\ln g}{\mathrm{d}\ln\mu} + \frac{1}{2}\epsilon \tag{1.79}$$

$$= \frac{\mathrm{d}g}{\mathrm{d}\ln\mu}\left(1 + \frac{5g^2}{8\pi^2}\frac{1}{\epsilon}\right) + \epsilon g. \tag{1.80}$$

If we expand $\frac{\mathrm{d}g}{\mathrm{d}\ln\mu} = a_1\epsilon + a_0 + \ldots$ order by order in ϵ, then matching the $\mathcal{O}(\epsilon)$ terms gives $a_1 = -g/2$ and matching the $\mathcal{O}(\epsilon^0)$ terms tells us that, in the $\epsilon \to 0$ limit, we have

$$\frac{\mathrm{d}g}{\mathrm{d}\ln\mu} = \frac{5}{16\pi^2}g^3 + \mathcal{O}(g^4). \tag{1.81}$$

This ϵ-independent piece is known as the 'beta function' for the coupling, $\beta(g) = \frac{5}{16\pi^2}g^3$. Of course there are higher-order terms in $\frac{\mathrm{d}g}{\mathrm{d}\ln\mu}$ which are needed to match the $\mathcal{O}(\epsilon^{-n})$ terms, and which one can solve for. But these vanish in the $\epsilon \to 0$ limit, so will not contribute to the running of $g(\mu)$.

We can now resum this logarithm to find the evolution of this coupling with renormalization scale

$$g(\mu) = \frac{\bar{g}}{\sqrt{1 - \frac{5\bar{g}^2}{8\pi^2}\log\frac{\mu}{\bar{\mu}}}}, \tag{1.82}$$

where we've used the boundary condition $g(\bar{\mu}) = \bar{g}$. As before, the resummed version will allow us to maintain precision to far lower scale than we could with simply its leading order approximation.

It's useful to keep in mind the Wilsonian picture as a clearer example because our regulator had a physical interpretation. The point is that the logarithms are really what's encoding how couplings change as a function of scale; in the Wilsonian calculation it was obvious that the logarithmic contribution $\log\frac{1}{b}$ is present no matter the initial cutoff. One says that couplings which receive logarithmic quantum corrections 'get contributions from all scales'. Then it's clear why this RG-improvement is sensible—though we may start at some particular Λ or μ, a one-loop calculation offers information on the lowest-order logarithmic running over *all* momenta, and we may sum up those modifications to improve our perturbation theory.

1.2.3   To Relate Theories

Mass-independent schemes and matching

We've seen already the necessity of renormalization when a theory produces naïvely divergent results, and its enormous use in improving the precision of perturbative calcu-

lations in a given theory. The last facet we'll discuss is its use in connecting theories. This is necessary to use the computationally-simple scheme of dim reg with $\overline{\text{MS}}$ in theories with different mass scales, and is very closely related to the effective field theory philosophy we discussed in Section 1.1. Cohen's monograph [16] goes into far more depth than I will be able to, and is a fantastic introduction to these ideas and their application. This perspective on renormalization has also been of enormous use in condensed matter to understand behaviors that appear in many distinct systems in the long-distance limit, and has applications in formal field theory to understand better the properties of QFT itself.

In the previous section we derived the beta function for Yukawa theory in the $\overline{\text{MS}}$ scheme. As promised by our terming of this as a 'mass-independent' scheme, the beta functions indeed have no reliance on the masses. But this should seem remarkably peculiar, as it suggests that there is no decoupling at all. Were that the case, by measuring the low-energy beta functions of QED we could tell how many charged particles existed up to arbitrarily large mass scales! What has gone wrong is that $\overline{\text{MS}}$ does not meet the criterion of a physical scheme which is necessary for the Appelquist-Carrazone theorem to operate. In $\overline{\text{MS}}$ the renormalization condition has nothing to do with physical values of the parameters so, while it makes calculations far simpler, $\overline{\text{MS}}$ has broken decoupling.

To restore decoupling and allow us to properly use a mass-independent scheme, we must implement the mass-dependence ourselves by 'matching' the Yukawa theory at energies above the fermion mass to a theory of solely light scalars at energies below the spinor mass. To 'match', we consider some process which exists in both theories—for example, ϕ^4 scattering—and ensure that at the matching scale M both theories agree on the physics.

In the high-energy Yukawa theory we can run the RG scale all the way down from a high scale $\bar{\mu}$ to $\mu = M_{\text{phys}}$, the physical, pole mass of the fermion. To get simple closed-

form expressions, we'll take the couplings small enough that working to lowest order gives a good approximation. We'll denote all of the UV values with bars, e.g. $M(\mu = \bar{\mu}) = \bar{M}$. Firstly, we use the counterterms to find the anomalous dimension of the fermion mass

$$\frac{d\log M}{d\log \mu} = -\frac{d}{d\log \mu}\left(Z_\psi^{-1}Z_M\right) \tag{1.83}$$

$$= -\frac{g^2}{16\pi^2} \tag{1.84}$$

$$\Rightarrow M(\mu) \simeq \bar{M}\left(\frac{\bar{\mu}}{\mu}\right)^{\frac{\bar{g}^2}{16\pi^2}} \tag{1.85}$$

$$\simeq \bar{M}\left(1 + \frac{\bar{g}^2}{16\pi^2}\log\frac{\bar{\mu}}{\mu}\right) + \dots \tag{1.86}$$

We then find the fermion pole mass as

$$0 = -M_{\mathrm{phys}} + M(\mu = M_{\mathrm{phys}}) - \Sigma(-M_{\mathrm{phys}}) \tag{1.87}$$

$$M_{\mathrm{phys}} = \bar{M}\left(1 + \frac{\bar{g}^2}{16\pi^2}\log\frac{\bar{\mu}}{\bar{M}}\right) - \Sigma(-\bar{M}) + \text{ subleading} \tag{1.88}$$

$$= \bar{M}\left(1 + \frac{\bar{g}^2}{16\pi^2}\log\frac{\bar{\mu}}{\bar{M}}\right) - \frac{\bar{g}^2}{16\pi^2}\bar{M}\int_0^1 dx\, x \log\left(\frac{x^2\bar{M}^2 + (1-x)\bar{m}^2}{\bar{M}}\right) + \dots$$
$$\tag{1.89}$$

$$= \bar{M}\left[1 + \frac{\bar{g}^2}{16\pi^2}\left(\frac{1}{2} + \log\frac{\bar{\mu}}{\bar{M}}\right)\right] + \dots \tag{1.90}$$

Now we need the value of the other parameters at that mass threshold

$$\frac{1}{m}\frac{dm}{d\log\mu} = \frac{1}{4\pi^2}\left(\frac{1}{8}\lambda - g^2\frac{m^2}{M^2} + \frac{1}{2}g^2\right) \tag{1.91}$$

$$\Rightarrow m(M_{\text{phys}}) = \bar{m}\left(1 - \frac{1}{4\pi^2}\left[\frac{1}{8}\bar{\lambda} - \bar{g}^2\frac{m^2}{M^2} + \frac{1}{2}\bar{g}^2\right]\log\frac{\bar{\mu}}{\bar{M}}\right) + \ldots \tag{1.92}$$

$$\frac{dg}{d\log\mu} = \frac{5}{16\pi^2}g^3 \tag{1.93}$$

$$\Rightarrow g(M_{\text{phys}}) = \bar{g}\left(1 - \frac{5\bar{g}^2}{16\pi^2}\log\frac{\bar{\mu}}{\bar{M}}\right) + \ldots \tag{1.94}$$

$$\frac{d\lambda}{d\log\mu} = \frac{1}{16\pi^2}\left(3\lambda^2 + 8\lambda g^2 - 48g^4\right) \tag{1.95}$$

$$\Rightarrow \lambda(M_{\text{phys}}) = \bar{\lambda}\left(1 - \frac{1}{16\pi^2}\left[3\bar{\lambda}^2 + 8\bar{\lambda}\bar{g}^2 - 48\bar{g}^4\right]\log\frac{\bar{\mu}}{\bar{M}}\right) + \ldots \tag{1.96}$$

Now we are ready to proceed to even lower energies. We enforce decoupling by matching to the low-energy theory of just a self-interacting scalar. We have

$$\mathcal{L}_0 = \frac{1}{2}\phi\left(\square - m^2\right)\phi \tag{1.97}$$

$$\mathcal{L}_1 = -\frac{1}{24}Z_\lambda\lambda\phi^4 + \mathcal{L}_{\text{ct}} \tag{1.98}$$

$$\mathcal{L}_{\text{ct}} = -\frac{1}{2}(Z_\phi - 1)\partial^\mu\phi\partial_\mu\phi - \frac{1}{2}(Z_m - 1)m^2\phi^2 \tag{1.99}$$

The counterterms and beta functions in this theory can be conveniently found by truncating those found above. Of course, by construction, we find that the heavy fermion ψ no longer contributes to the running of parameters at low-energies. To make sure we're getting the physics correct, we must impose the boundary condition that the predictions match at $\mu = M_{\text{phys}}$, which here is quite simple—we just use the values at M_{phys} in the UV theory as literal boundary conditions for our running in the IR theory. In the IR

theory, for $\mu \leq M_{\text{phys}}$, we have

$$\frac{d\lambda}{d\log\mu} = \frac{3}{16\pi^2}\lambda^2 \tag{1.100}$$

$$\Rightarrow \lambda(\mu) = \frac{\lambda(M_{\text{phys}})}{1 + \frac{3}{16\pi^2}\lambda(M_{\text{phys}})\log\frac{M_{\text{phys}}}{\mu}} \tag{1.101}$$

$$\simeq \bar{\lambda}\left(1 - \frac{1}{16\pi^2}\left[3\bar{\lambda}^2 + 8\bar{\lambda}\bar{g}^2 - 48\bar{g}^4\right]\log\frac{\bar{\mu}}{\bar{M}}\right)\left(1 - \frac{3}{16\pi^2}\bar{\lambda}\log\frac{\bar{M}}{\mu}\right) + \ldots \tag{1.102}$$

$$\simeq \bar{\lambda}\left(1 - \frac{1}{16\pi^2}\left[3\bar{\lambda}^2 + 8\bar{\lambda}\bar{g}^2 - 48\bar{g}^4\right]\log\frac{\bar{\mu}}{\bar{M}} - \frac{3}{16\pi^2}\bar{\lambda}\log\frac{\bar{M}}{\mu}\right) + \ldots \tag{1.103}$$

The benefit is now clear. While the RGE in the UV theory were very complicated, the running of λ in the low energy theory is simple. Our mass-independent scheme allows us to explicitly factorize these and contain all the UV physics in the boundary condition, which lets us study the low-energy theory in a simple manner.

The general procedure of renormalization group evolution in mass-independent schemes is called 'running and matching'. The parameters in the Lagrangian run as you evolve down in energies, but at a mass threshold M we must match the UV theory at $\mu = M$ from above to a theory without this field at $\mu = M$ from below. When we match we ensure that the physics of the low-energy fields stays constant as we cross that threshold and remove that particle from the spectrum of our theory. This becomes less trivial when we have multiple mass scales, so consider now upgrading our Yukawa theory with additional fermions.

$$\mathcal{L} \supset \sum_{i=1}^{N} M_i \bar{\psi}_i \psi_i + g^{(N)}\phi \sum_{i=1}^{N} \bar{\psi}_i \gamma_5 \psi_i \tag{1.104}$$

Now if we are given the theory at very high energies $\mu \gg M_i$ and we want to understand what it looks like at very low energies, there is a cascade of EFTs we evolve through. As depicted in Figure 1.8, we run the parameters down to the largest fermion mass, match to a theory with one less fermion, run down again until the next mass scale, and so on.

As an example which is closer to the real world, consider QED with our three generations of leptons (and ignoring the strong sector for simplicity). At low energies we measure the asymptotic value $\alpha(\mu \simeq 0) \simeq 1/137$, and in colliders we measure the value of the gauge coupling at the Z mass $m_Z \simeq 91$ GeV. To compare these, we must run the high-energy value all the way down into the infrared. Above $m_\tau \simeq$ 1.7 GeV we have a theory where all of e, μ, τ run in loops and giving an $\overline{\text{MS}}$ beta function $\beta_\alpha = 2\alpha^2/\pi$. But at m_τ we should remove the τ from our theory, such that from m_τ down to $m_\mu \simeq 105$ MeV we have $\beta_\alpha = 4\alpha^2/3\pi$. Below the μ we solely have the electron and recover the textbook $\beta_\alpha = 2\alpha^2/3\pi$, and finally as we cross the electron mass threshold $m_e \simeq 511$ keV we remove the electron from the spectrum and find that the gauge coupling stops running $\beta_\alpha \equiv 0$. Physically this corresponds to the fact that pure QED is

Figure 1.8: A schematic depiction of the cascade of effective field theories as one transitions from a Yukawa theory of N fermions at high energies through integrating these out sequentially until one finds an effective ϕ^4 theory at low energies.

scale-invariant, meaning that the coupling will not evolve at all in a theory with no charged particles. This is the regime in which classical electrodynamics holds precisely (up to the presence of additional interactions suppressed by powers of m_e, that is).

A possible confusion is to conflate the mass-independence of the regularization scheme with that of the renormalization scheme, and conclude that dimensional regularization cannot be used if one wants decoupling without having to integrate out and match. So lest one confuse the roles let's quickly look at an example of using dimensional regularization with a renormalization scheme which *does* satisfy decoupling, known as 'off-shell momentum subtraction'. For simplicity, we'll look at the anomalous dimension of our Yukawa scalar ϕ, and we'll perform wavefunction renormalization by subtracting the

47

value of the graphs at the off-shell momentum scale $k^2 = \mu_E^2$. In symbols this amount to the prescription

$$Z_\phi = 1 + \Sigma_{\text{loop}}(k^2 = \mu_E^2), \quad \text{where } \Pi_{\text{loop}}(k^2) = k^2 \Sigma_{\text{loop}}(k^2) + \text{ mass renorm pieces.}$$

$$(1.105)$$

Since we're still using the same regularization scheme, we have the same result for $\Pi_{\text{loop}}(k^2)$ as above. We can then simply calculate the anomalous scaling dimension as defined by $\gamma_\phi \equiv \frac{1}{2}\frac{\text{d}\log Z_\phi}{\text{d}\log \mu_E}$,

$$\gamma_\phi = \frac{1}{2}\frac{\text{d}\Sigma_{\text{loop}}(\mu_E^2)}{\text{d}\log \mu_E} \tag{1.106}$$

$$= \frac{3g^2}{4\pi^2}\int\limits_0^1 \text{d}x \frac{x^2(1-x)^2\mu_E^2}{M^2 + (1-x)x\mu_E^2}. \tag{1.107}$$

This integral can be performed analytically, but the full expression is unilluminating. However, it is useful to look at the limits

$$\gamma_\phi = \begin{cases} \frac{g^2}{8\pi^2}, & \mu_E^2 \gg M^2 \\[2mm] \frac{g^2}{40\pi^2}\frac{\mu_E^2}{M^2}, & \mu_E^2 \ll M^2 \end{cases}$$

to check if they agree with our expectations. At energies far above the fermion mass its contribution to the scalar anomalous dimension cannot know about that scale, and at energies far below its mass we expect inverse dependence on the mass for decoupling to occur. This is precisely what we find, so the lesson is that even if we didn't want to go through the trouble of integrating out the fermion and matching, we could still make use of the magical regularization scheme that is dim reg.

Flowing in theory space

Our interpretation of the renormalization group thus far has been as a way of under-standing what a particular theory looks like at different energies. But there is another way of looking at it that is also useful, for which we shall follow an example of Peskin & Schroeder. Let's return to the idea of the Wilsonian path integral and successively integrating out Euclidean momentum shells. In the previous section we began with a scalar field theory

$$Z = \left(\int \prod_{k=0}^{\Lambda} \mathrm{d}\phi(k) \right) \exp\left[-\int \mathrm{d}^d x \left(\frac{1}{2}(\partial_\mu \phi)^2 + \frac{1}{2}m^2\phi^2 + \frac{1}{4!}\lambda\phi^4 \right) \right]. \qquad (1.108)$$

We then integrated over momentum shells from Λ down to $b\Lambda$ with $0 < b < \Lambda$, and found we could express our result as (schematically; see 1.32,1.34,1.35,1.41)

$$Z = \left(\int \prod_{k=0}^{b\Lambda} \mathrm{d}\phi(k) \right) \exp\left[-\int \mathrm{d}^d x \mathcal{L}_{\text{eff}} \right] \qquad (1.109)$$

$$\mathcal{L}_{\text{eff}} = \left(\frac{1}{2}(1+\Delta Z)(\partial_\mu \phi)^2 + \frac{1}{2}(m^2 + \Delta m^2)\phi^2 + \frac{1}{4!}(\lambda + \Delta\lambda)\phi^4 + \frac{1}{6!}\Delta c_6 \phi^6 + \dots \right).$$

$$(1.110)$$

Above we interpreted this in terms of looking at the same theory at lower energies, having coarse-grained over the largest momentum modes, which is a useful way of comparing the two path integrals. Another useful way to compare is to get them to a form where they look similar, so let's now define a change of variables $k' = k/b, x' = xb$, in terms of

which the path integral now looks like

$$Z = \left(\int \prod_{k'=0}^{\Lambda} \mathrm{d}\phi(k') \right) \exp \left[-\int \mathrm{d}^d x' \mathcal{L}_{\text{eff}} \right] \tag{1.111}$$

$$\mathcal{L}_{\text{eff}} = b^{-d} \left(\frac{1}{2}(1 + \Delta Z)b^2(\partial_{\mu'}\phi)^2 + \frac{1}{2}(m^2 + \Delta m^2)\phi^2 + \frac{1}{4!}(\lambda + \Delta\Lambda)\phi^4 + \frac{1}{6!}\Delta c_6\phi^6 + \dots \right).$$
$$\tag{1.112}$$

We can transform the kinetic terms back to the canonical form with the field redefinition $\phi'(x') = b^{(2-d)/2}(1 + \Delta Z)^{1/2}\phi(x')$, after which we can write the effective Lagrangian as

$$\mathcal{L}_{\text{eff}} = \frac{1}{2}(\partial_{\mu'}\phi')^2 + \frac{1}{2}m'^2\phi'^2 + \frac{1}{4!}\lambda'\phi'^4 + \frac{1}{6!}c_6'\phi'^6 + \dots. \tag{1.113}$$

with $m'^2 = (m^2 + \Delta m^2)(1 + \Delta Z)^{-1}b^2$, $\lambda' = (\lambda + \Delta\lambda)(1 + \Delta Z)^{-2}b^{d-4}$, $c_6' = \Delta c_6(1 + \Delta Z)^{-3}b^{2d-6}, \dots$, and it's clear that we could write such an effective action regardless of what sort of coefficients we began with before integrating out this momentum shell.

Now our series of transformations has effected the change (up to normalization)

$$Z = \left(\int \prod_{k=0}^{\Lambda} \mathrm{d}\phi(k) \right) \exp \left[-\int \mathrm{d}^d x \mathcal{L} \right] \rightarrow Z = \left(\int \prod_{k'=0}^{\Lambda} \mathrm{d}\phi'(k') \right) \exp \left[-\int \mathrm{d}^d x' \mathcal{L}' \right].$$
$$\tag{1.114}$$

Since all of our dynamical variables are integrated over in calculating the partition function, we can view this as a transition in the space of Lagrangians, $\mathcal{L} \rightarrow \mathcal{L}'$. So this gives us an interpretation of the renormalization group as a flow in 'theory space'.

This interpretation invites us to conceptualize renormalization group flow as a path through theory space between two conformal field theories (CFTs), as depicted in Figure 1.9. CFTs are quantum field theories with an enlarged spacetime symmetry group[7], consisting essentially of a scaling symmetry [8]. CFTs are fixed points of RG flows—since

---

[7]This does not violate the Coleman-Mandula theorem [67] because CFTs do not have S-matrices.

[8]It is not known whether scale invariance in fact implies conformality in four-dimensional QFTs, the
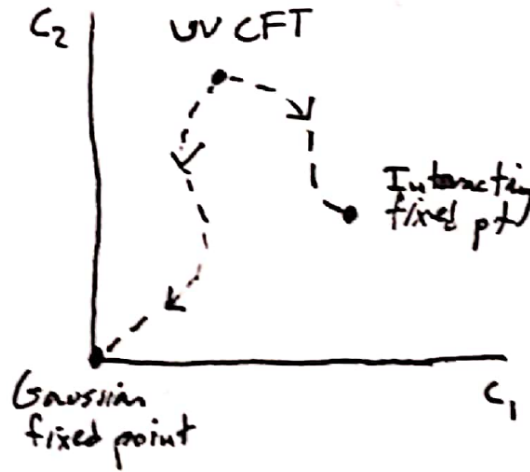
Figure 1.9: Schematic two-dimensional projection of some RG flow trajectories from an interacting UV fixed point to IR fixed points that are either interacting or free.

they possess scaling symmetry they look the same at all energy scales, so if an RG flow is to have an endpoint it clearly must be a CFT.

There's a terminological confusion here, which is that the 'renormalization group' isn't actually a group at all, since the operation of integrating out a momentum shell is irreversible. This came up already above when we saw that integrating out heavy fields means we can no longer compute processes which have them as external fields (see Footnote 1). Flowing to lower energies, or toward the IR, is really a coarse-graining operation which does lose information about small scales, in precise analogy to decreasing the resolution of an image. This means that RG evolution is a directed flow, so there is a difference between fixed points in the UV and in the IR.

Quantum field theories can have different sorts of fixed points. If the theory has a 'mass gap'—no zero-energy excitations in the infrared, which may be because one began solely with massive fields or through dynamical mechanisms like Higgsing and confinement—then one finds a 'trivial' fixed point. In the far infrared, everything has

latter of which includes also invariance roughly under inversion of spacetime through a point. Polchinski's early paper on the topic is a classic [68], and a recent review can be found from Nakayama [69].

been integrated out and there is not enough energy to excite any modes. We know phenomenologically that this happens in QCD. Alternatively, one can have a 'Gaussian' or 'free' fixed point if the theory contains massless fields which don't interact, such as in QED. At energies far below the mass of the lightest charged particle this is a theory of free electromagnetism, though one can excite photons of arbitrarily-long wavelength.

Such a Gaussian fixed point occurs in the UV for QCD—the celebrated result of 'asymptotic freedom' [70, 71]—because the strong coupling flows toward zero, giving a free theory. This famously cannot occur for $U(1)$ gauge theories whose couplings necessarily grow with increasing energies, leading them to herald their own breakdown with the prediction of a 'Landau pole' [72], a finite UV energy where the perturbative theory predicts the coupling becomes infinite. From one perspective this is an inverse to the prediction of a confinement scale in QCD, where the perturbative prediction is a blowup of the coupling at low energies, as we'll discuss further in Section 1.3.3. In either case the theory cannot make predictions for energies above the Landau pole or below the confinement scale, respectively, and so can be called inconsistent.

It's clear that the divergence of a coupling either in the IR or the UV is problematic for a complete, consistent interpretation of a QFT. This is precisely why the formal perspective on a well-defined QFT is that it describes an RG flow between two CFTs, such that in neither direction does a coupling grow uncontrollably. In fact this provides an incredibly important perspective on renormalizability, which we'll get to momentarily.

First, let us introduce somewhat of a generalization of the 'relevant/irrelevant' terminology which we introduced in Section 1.1. We implicitly had in mind that we were studying a theory in the vicinity of the Gaussian fixed point which we perturbed with various field operators—indeed, this is precisely how we normally carry out perturbative calculations—and our terminology depended on that. Generally one wishes to imagine perturbing a CFT by a particular operator, flowing down in energy, and seeing whether

the operator grows in importance—a relevant operator—or shrinks—an irrelevant operator. Perturbing a CFT by an irrelevant operator does not induce an RG flow (to a different CFT), so interesting dynamical RG flows come from CFTs perturbed by relevant operators. This clearly agrees with our power-counting notion of relevance when we're near the Gaussian fixed point, but works also if one is near a strongly-coupled, interacting fixed point where one may not know how to do such power-counting and anomalous dimensions of operators may grow to overpower their classical dimensions.

The importance of this language was realized in particular by Polchinski in his pioneering article [73]. He showed that in fact the intuitive notion of 'power-counting renormalizability' that the field had been building—that for a theory near the Gaussian fixed point we could see whether it was renormalizable merely by checking whether it has any coefficients of negative mass dimensions—in fact maps on to a very general statement. This is enormously powerful, as prior arguments for renormalizability were made on a case by case basis and were complicated and messy and graph-theoretic. His derivation of this fact is brilliant but requires much work, so we'll merely try to get a sense for why it should be true by building on our intuition from our Wilsonian renormalization of ϕ^4 theory in Section 1.2.1.

So long as your theory has a finite number of fields, all of which have mass dimension $[\phi_i] > 0$ when canonically normalized, then there are a finite number of relevant or marginal operators. As we flow down in energy through theory space, we saw above the sense in which those coefficients are UV-sensitive—the IR coefficients of those operators are determined primarily by their UV values and IR corrections are subdominant. Contrariwise, the coefficients of the infinite number of possible irrelevant operators are UV-insensitive, being determined primarily by the IR physics of the relevant and marginal operators. That is, the RG flows are attracted to a finite-dimensional submanifold of the space of Lagrangians.

Consider a Wilsonian RG flow where we start off, as above, by specifying a cutoff Λ_0 and the values of coefficients λ_0^i of all the n marginal and relevant operators at that scale, as well as the coefficients c_0^i of however many irrelevant operators we wish to turn on. We can then flow downwards in energy as normal, and at an energy scale $\Lambda_R \ll \Lambda_0$ let's say we measure the relevant and marginal coefficients λ_R^i at that scale. The coefficients of the irrelevant operators are then dominated by the infrared λ_R^i and Λ_R up to precision $(\Lambda_R/\Lambda_0)^{\Delta_i}$ from subleading corrections, with $\Delta_i > 0$ the scaling dimension of the irrelevant operator. So indeed, the RG flow is attracted to the n-dimensional surface described by $c_i = c_R^i(\lambda_R^i; \Lambda_R)$ and separate trajectories through theory space as a function of scale Λ which reach the same λ_R^i will differ only by positive powers of (Λ_R/Λ_0). For less abstract discussion, Polchinski goes through a simple example which may provide further insight, and Schwartz discusses the same example in Chapter 23 of his textbook [26].

To see why this implies renormalizability, recall that the program of Wilsonian renormalization is to define renormalized, 'running' couplings as a function of scale to keep infrared physics at Λ_R fixed while 'removing the cutoff'. A bit more formally, we want a family of Lagrangians $\mathcal{L}(\Lambda_0; \lambda_0^i, c_0^i)$ with coefficients chosen as a function of Λ_0 such that each Lagrangian yields the same low energy physics λ_R^i, in terms of which all the IR observables can be calculated up to subleading corrections in Λ_R/Λ_0. When the cutoff is removed by taking $\Lambda_0 \to \infty$, one thus recovers precisely the correct physics, specified by those chosen values of the λ_R^i, which are the renormalization conditions. If we can find such a family of Lagrangians, then we say this theory is power counting renormalizable. Polchinski's argument shows that this can be done so long as one wishes solely to fix the IR values of relevant and marginal couplings.

Trivialities

The attentive reader may at this point notice an inconsistency due to imprecise language. We've seen now that the criterion for renormalizability, which Polchinski provided a robust basis for, is power-counting of the operators near the IR fixed point. This would suggest that the $\lambda\phi^4$ theory we've studied by means of an example is renormalizable. However, recall the result of resumming its renormalization group equation:

$$\lambda(\Lambda) = \frac{\lambda(\Lambda_0)}{1 + \frac{3\lambda(\Lambda_0)}{16\pi^2}\log\frac{\Lambda_0}{\Lambda}}, \tag{1.115}$$

which has a Landau pole for $\Lambda = \Lambda_0 \exp(\frac{16\pi^2}{3\lambda(\Lambda_0)})$, preventing us from taking the limit we required above. A mathematical physicist would say $\lambda\phi^4$ theory is 'trivial' or 'quantum trivial', as if we demand the existence of a continuum limit, that sets $\lambda(\Lambda) = \lambda(\Lambda_0) = 0$.

The issue is that the tree-level, classical scaling dimension captures only the scaling of the operators infinitesimally close to the infrared Gaussian fixed point. If we move a finite distance upward in energy scale, we've seen above that the ϕ^4 operator gets an anomalous dimension $\delta_{\phi^4} > 0$ and so is marginally irrelevant. So it's clear that Polchinski's picture of renormalization is only getting at a perturbative sense of renormalizability, and cannot tell us whether there truly exists an RG flow from a UV fixed point down to the IR theory we want to study.

So what are we to make of $\lambda\phi^4$ theory—or for that matter of QED, which has the same problem? Of course we know empirically that QED works fantastically well and we can absolutely make finite, accurate predictions after a finite number of inputs. To understand this, we must appeal the language of effective field theory, which we've already discussed. In fact the feature we're really relying on is effective renormalizability, which tells us we require solely a finite number of inputs to set the behavior of the theory to a given precision. It's clear that in this sense QED itself is an effective field theory whose

validity breaks down somewhere below its Landau pole.

Finally, let me mention another reason not to be too worried that our most beloved quantum field theories don't exist in the continuum limit: Our universe is not described by a QFT at its smallest scales! It's indeed true that only RG flows between a UV CFT and an IR CFT can hope to define fully consistent and mathematically well-defined QFTs. But the existence of gravity—and the very strong evidence that a quantum field theory of gravity is inconsistent—means that at some energy scale effects not present in quantum field theory must become relevant. And since gravity couples universally to everything [74, 75], we have no strict empirical need for a UV complete, interacting quantum field theory that does not include gravity. It is entirely consistent, and overwhelmingly likely, that a quantum-field-theoretic description of the world works only approximately and some inherently quantum gravitational theory provides a sensible UV complete theory.

1.2.4   To Reiterate

Before moving on, let us reiterate what we've discussed about renormalization. As we've seen, renormalization is *so* important and *so* useful and fulfills *so many* purposes that an entirely general statement risks becoming vague. But if a single sentence summary is demanded: Renormalization reveals for us the scale-dependence of a quantum mechanical field theory.

The effects of this seemingly innocuous statement, however, are powerful and manifold, including:

- Correctly accounting for this scale-dependence is necessary to have well-defined quantum field theories, which otherwise appear nonpredictive.

- Bringing the non-scale-invariance of the quantum mechanical theory into clear scale-dependence of the couplings makes it simple to read off the qualitative behavior at

different scales from the renormalized Lagrangian.

- Including this scale-dependence in the couplings allows us to reorganize our perturbative series such that we can efficiently calculate the behavior of the theory over a far wider range of energies than a naïve treatment allows.

- Properly accounting for the scale dependence allows us to harness the full power of effective field theory, as we can study a theory of low-energy fields which correctly accounts for corrections from the high-energy physics.

- Understanding the perspective of single quantum field theories as flows through theory space as a function of the scale allowed us to develop a nonperturbative definition of how a fully UV-complete quantum field theory behaves and what it is.

All of these various perspectives will be of use in the following chapters as we apply this technology to understanding the hierarchy problem and how we can solve it.

1.3   Naturalness

Naturalness is the notion that we have the right to ask about the origins of the dimensionless numbers in our theories—past solely fitting them to the data. It was Dirac who first introduced such a notion to particle physics in 1938 [76]. In modern language, what is referred to as 'Dirac naturalness' consists of the idea that dimensionless parameters in a fundamental physical theory should take values of order 1. In the language of EFT, in a theory with a cutoff Λ and an operator O with scaling dimension Δ, we expect its coefficient to taken a value $c_O \sim \mathcal{O}(1)\Lambda^{d-\Delta}$. As stated this is essentially dimensional analysis, as Λ is the only scale we've introduced, but we will discuss below that quantum mechanics gives additional credence to this expectation—indeed we've already seen this feature in our one-loop examples above.

't Hooft pointed out a refinement of this principle [77], which has come to be called 'technical naturalness'. If the operator O breaks a symmetry which is respected by the action in the limit $c_O \to 0$, then one says it is 'technically natural' for c_O to take on a small value. The reasoning here is simple—as we saw above, in a quantum field theory defined at a high scale, one finds corrections δc_O to such coupling constants as they run to low energies. If there is an enhanced symmetry of the theory in the limit that $c_O \to 0$, then such quantum mechanical corrections cannot generate that operator and break the symmetry, so we know that $\delta c_O \propto c_O$. The low-energy effective field theorist says of such couplings that one can 'set it and forget it': if one has $c_O \ll 1$ at the cutoff Λ, that coupling will remain small as one flows to lower energies.

Conversely, we can emphasize the connection to Dirac naturalness by looking at this picture in reverse. We know of mechanisms to generate small technically natural couplings at low energies from Dirac natural ones, as we will discuss in detail below. Imagine one measures a small coupling c_O at long distances in the low-energy theory with cutoff Λ that does not have a Dirac natural explanation. If that parameter is technically natural, it remains small up to the cutoff, and so the next generation of physicists can explain its small size at Λ in the UV theory, as emphasized nicely by Zhou [78]. If c_O is not technically natural, then its RG evolution up the cutoff yields a value to which the low-energy physics is very sensitive, and we must explain why it has a very specific value such that the correct physics emerges at long distances.

1.3.1 Technical Naturalness and Fine-Tuning

> A model is fine-tuned if a plot of the allowed parameter
> space makes you wanna puke.
>
> David E. Kaplan (2007)

It's useful to make this less abstract by looking at a simple example. Consider a $d = 6$ dimensional effective field theory of a real scalar field ϕ of mass m which is odd under a $\mathbb{Z}_2$ symmetry, which we expect is a good description of our system up to a cutoff Λ with $m \ll \Lambda$. If we add a small explicit breaking $\sigma\phi^3$ with $\sigma \ll 1$ at low energies, σ is technically natural and stays small up to the cutoff, so we can easily write down a UV completion which generates this small value Dirac naturally.

However, if we add another $\mathbb{Z}_2$-odd real scalar Φ and give it a large $\mathbb{Z}_2$-breaking interaction with ϕ, then σ is no longer technically natural. Its low-energy value becomes extremely sensitive to the values of the parameters at the cutoff. It then becomes difficult to understand an ultraviolet reason for why those values take the precise values they need to realize small σ in the far IR. Consider the bare action

$$S = \int \mathrm{d}^6 x \left[-\frac{1}{2}(\partial\phi_0)^2 - \frac{1}{2}m_0^2\phi_0^2 - \frac{1}{2}(\partial\Phi_0)^2 - \frac{1}{2}m_0^2\Phi_0^2 - \frac{\sigma_0}{3!}\phi_0^3 - \frac{y_0}{2}\phi_0\Phi_0^2 \right], \quad (1.116)$$

where we've given the two fields the same mass for simplicity. This is not stable under radiative corrections, but that's a higher-order effect which will not come into our one-loop calculation of the RG evolution of the cubic couplings.

We will renormalize this theory at one loop using dim reg with $\overline{\mathrm{MS}}$. As discussed above, we will compute the one-loop 1PI diagrams and add counterterms to cancel solely

the $\frac{1}{\epsilon}$ pieces of the results. With counterterms, the action is

$$S = \int \mathrm{d}^d x \left[-\frac{1}{2} Z_\phi (\partial \phi)^2 - \frac{1}{2} Z_{m\phi} m^2 \phi^2 - \frac{1}{2} Z_\Phi (\partial \Phi)^2 - \frac{1}{2} Z_{m\Phi} m^2 \Phi^2 + \frac{\sigma}{3!} Z_\sigma \phi^3 + \frac{y}{2} Z_y \phi \Phi^2 \right]$$
(1.117)

where these parameters and fields are the renormalized parameters, and for compactness we have not written down the split of these terms as we did above in Equation 1.97. At tree level the relation to the bare quantities is trivial and so $Z = 1 + \ldots$. To get an accurate picture of how the strength of the interactions vary as they're probed at different energy scales, we must fully renormalize the theory. Since our focus is on the interactions, we simply state the results for the quadratic part of the action, where we have

$$Z_\phi = 1 - \frac{1}{6(4\pi)^3} \left(\sigma^2 + y^2 \right) \frac{1}{\epsilon} + \ldots$$
(1.118)

$$Z_\Phi = 1 - \frac{1}{3(4\pi)^3} \left(\sigma^2 + y^2 \right) \frac{1}{\epsilon} + \ldots$$
(1.119)

$$Z_{m\phi} = 1 - \frac{1}{(4\pi)^3} \left(\sigma^2 + y^2 \right) \frac{1}{\epsilon} + \ldots$$
(1.120)

$$Z_{m\Phi} = 1 - \frac{2}{(4\pi)^3} y^2 \frac{1}{\epsilon} + \ldots$$
(1.121)

These tell us that the physical mass of the fields and the normalization of their one-particle states has changed. The relation to these can be found using the quantum-corrected propagator $\Delta(k^2)^{-1}$ as $\Delta(-m_{\mathrm{phys}}^2)^{-1} \equiv 0$ to define the mass and $R^{-1} = \frac{d}{dk^2} \left[\Delta(k^2)^{-1} \right]\big|_{k^2 = -m_{\mathrm{phys}}^2}$ to define the normalization R, but solving for these relations explicitly will not be necessary for our purposes.

The one-loop three-point functions each have two diagrams, whose evaluation only differs in the coupling constants. For the correction to σ, we have triangle diagrams with
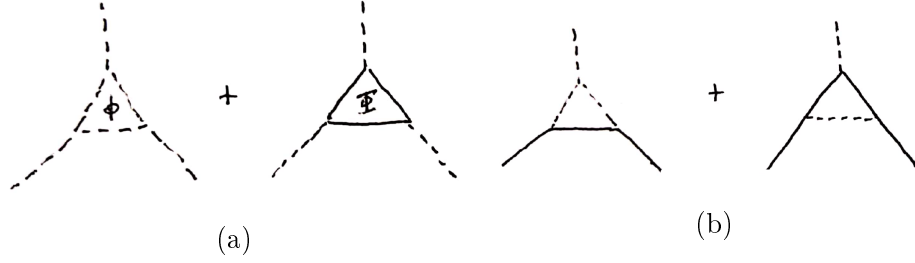
Figure 1.10: Diagrams contributing to the one-loop corrections to the three-point functions in our $6d$ scalars with cubic interactions. Dashed lines denote ϕ and full lines denote Φ. In (a), the diagrams renormalizing σ, and in (b), the diagrams renormalizing y.

either ϕ or Φ running in the loop. We can evaluate them in $d = 6 - \epsilon$ dimensions as

$$i\frac{1}{\sigma}V_\sigma = \frac{1}{\sigma}\left[(i\sigma)^3 + (iy)^3\right]\int\frac{d^6q}{(2\pi)^6}\frac{(-i)^3}{(q^2+m^2)^3} \tag{1.122}$$

$$= \frac{i}{\sigma}\left[\sigma^3 + y^3\right]\tilde{\mu}^\epsilon\int\frac{d^d\bar{q}}{(2\pi)^d}\frac{1}{(\bar{q}^2+m^2)^3} \tag{1.123}$$

$$= \frac{i}{\sigma}\left[\sigma^3 + y^3\right]\tilde{\mu}^\epsilon\frac{\Gamma\left(\frac{\epsilon}{2}\right)}{2(4\pi)^3}\left(\frac{m^2}{4\pi}\right)^{-\frac{\epsilon}{2}} \tag{1.124}$$

$$= \frac{i}{2(4\pi)^3\sigma}\left[\sigma^3 + y^3\right]\left(\frac{2}{\epsilon} - \gamma_E\right)\left(1 + \frac{\epsilon}{2}\log\frac{4\pi\tilde{\mu}^2}{m^2}\right) \tag{1.125}$$

$$= \frac{i}{2(4\pi)^3\sigma}\left[\sigma^3 + y^3\right]\left(\frac{2}{\epsilon} + \log\frac{\mu^2}{m^2}\right). \tag{1.126}$$

The counterterm vertex contributes to this as $-i(Z_\sigma - 1)$, meaning that $\overline{\text{MS}}$ prescribes we set

$$Z_\sigma = 1 + \frac{1}{(4\pi)^3}\left(\sigma^2 + \frac{y^3}{\sigma}\right)\frac{1}{\epsilon} \tag{1.127}$$

For the other vertex correction there are diagrams with either one or two of each internal line, which give

$$i\frac{1}{y}V_y = \frac{1}{y}\left[(iy)^3 + (i\sigma)(iy)^2\right]\int\frac{d^6q}{(2\pi)^6}\frac{(-i)^3}{(q^2+m^2)^3} \tag{1.128}$$

$$= \frac{i}{2(4\pi)^3\sigma}\left[y^2 + \sigma y\right]\left(\frac{2}{\epsilon} + \log\frac{\mu^2}{m^2}\right), \tag{1.129}$$

leading to the counterterm

$$Z_y = 1 + \frac{1}{(4\pi)^3} \left(y^2 + \sigma y\right) \frac{1}{\epsilon}. \tag{1.130}$$

This gives us the beta functions

$$\beta_\sigma = \frac{1}{4(4\pi)^3} \left(-3\sigma^3 - 4y^3 + y^2\sigma\right) + \dots \tag{1.131}$$

$$\beta_y = \frac{1}{12(4\pi)^3} \left(\sigma^2 y - 12\sigma y^2 - 7y^3\right) + \dots \tag{1.132}$$

Now we are finally in a place to mathematically evince our physics point about technical naturalness. Without Φ, the coupling σ is the only one which violates the $\mathbb{Z}_2$ and so the beta function is necessarily proportional to σ. [9] Let's say we recruit an experimentalist friend of ours to measure the $2 \to 2$ scattering cross-section of ϕs and we find that at $\mu = m$, the theory fits the data for $\sigma(m) = \sigma_0$ with $\sigma_0 \ll 1$. Solving the beta function, we find that to lowest order

$$\sigma(\mu) \simeq \sigma_0 \left(1 - \frac{3}{4(4\pi)^3}\sigma_0^2 \log\left[\frac{\mu}{m}\right]\right). \tag{1.133}$$

So σ is indeed radiatively stable. If $\sigma(m) = \sigma_0$ is small, then it takes until the enormous scale $\mu \simeq m \exp\left((4\pi)^3/\sigma_0^2\right)$ for σ to change by an order one fraction. So running $\sigma(\mu)$ up to wherever the cutoff Λ of our theory is, $\sigma(\Lambda)$ will still be small. If by Λ we haven't discovered any explanation for the size of $\sigma(m)$, we can ask the theory above Λ to produce this small value of $\sigma(\Lambda)$ from Dirac natural parameters at yet higher energies. Perhaps the high-energy Dirac natural value of σ has been relaxed toward zero by an axion-type

---

[9]This is trivial in our case as σ would be the only interaction, but you'll find the feature persists if you add other symmetry-respecting interactions, for example a $\mathbb{Z}_2$-even scalar ψ with an interaction $\phi^2\psi$. The general argument for this fact is given in the next section.

mechanism, or perhaps σ is the vev of another $\mathbb{Z}_2$-odd field which spontaneously broke the $\mathbb{Z}_2$ via confinement. We don't need to know a particular mechanism; the fact that σ is technically natural means that if we don't find an explanation for its size there is hope yet that our academic descendants will. One says that here σ is UV-insensitive as its low-energy value does not depend strongly on the physics at high energies.

On the other hand, if σ is *not* technically natural, we have a much more difficult issue. If we now have the theory with both ϕ and Φ, and our experimentalist measures $\sigma(m) = \sigma_0 \ll 1$ and $y(m) = y_0 = \mathcal{O}(1)$, then to lowest order the RG evolution of σ will be $\beta_\sigma \simeq -y^3/(4\pi)^3$ leading to

$$\sigma(\mu) = \sigma_0 - \frac{1}{(4\pi)^3} y_0^3 \log\left[\frac{\mu}{m}\right]. \tag{1.134}$$

And we can see the issue, as we are no longer guaranteed that a small $\sigma(m)$ is related to a small $\sigma(\Lambda)$. For concreteness, if $\sigma_0 = 10^{-3}$, $y_0 = -5$ and $\Lambda = 10^5 m$, then (using the full one-loop RG), we have $\sigma(\Lambda) \simeq 0.55$ and $y(\Lambda) \simeq 4.16$. How are we to ensure *these* values at Λ? We know how to produce small numbers, but not incredibly specific ones.

To see that we do need to produce these values very precisely, let's switch directions and consider the RG evolution down in energy from Λ to m. In the theory with solely ϕ, the coupling σ runs incredibly slowly, so an $\mathcal{O}(1)$ change in $\sigma(\Lambda)$, evolved down to the scale m, results in an $\mathcal{O}(1)$ change to $\sigma(m)$. But in the theory with two sources of breaking, $\sigma(m)$ is *enormously* sensitive to the values of the couplings at Λ. With the same cutoff $\Lambda = 10^5 m$, if we very slightly change the input value to $\sigma(\Lambda) = 0.56$ and leave $y(\Lambda)$ as above, evolving down now gives us $\sigma(m) \simeq 10^{-2}$—a $< 2\%$ change in input parameters has resulted in a 1000% change in our low energy observable! It's even more sensitive to the input value of y; a $\sim 1\%$ modification solely to $y(\Lambda) = -4.20$ trickles down to give $\sigma(m) \simeq 2 \times 10^{-2}$, a 2000% change. In this theory σ is now a UV-sensitive

parameter, whose low-energy behavior depends strongly on the high-energy physics. To say the least, it seems difficult to find a natural way to achieve the precise values needed to reproduce the observed low-energy physics in this theory. We'll return to this issue at length in Section 2.2.2.

Technical naturalness and masses

Our understanding of technical naturalness allows us to already see another warning sign of the hierarchy problem. An elementary spin-1 field comes along with a gauge symmetry $A^\mu \to A^\mu - \partial^\mu \alpha(x)$ which is broken by a mass term $m_A^2 A^\mu A_\mu$. So a mass for a gauge boson is technically natural and one necessarily finds $\delta m_A^2 \propto m_A^2$. Similarly, a massive Dirac fermion $\Psi = (\psi, \chi^\dagger)$ has a $U(1)$ global symmetry under which $\psi \to e^{i\alpha}\psi, \chi \to e^{-i\alpha}\chi$. In the $m \to 0$ limit, the symmetry is enhanced to $U(1)^2$ as arbitrary rephasings of the two Weyl fermions become symmetries, so again $\delta m_\Psi \propto m_\Psi$. But an elementary scalar does not automatically come with any such protective symmetry, and we've already seen in all our examples above that scalar mass corrections indeed get contributions not proportional to the mass itself.

In fact for discussing the technical naturalness of masses there is an even simpler argument: A massless spin-1 boson has two degrees of freedom and a massive one has three. Quantum corrections cannot generate a degree of freedom *ex nihilo*, so a massless gauge boson must be protected. Similarly a massless chiral fermion has two degrees of freedom, but a massive Dirac fermion has four. So for charged spinors and for vectors, it is simply the representation theory of the Lorentz group that is responsible for the stability of their masses. In either of these cases mass must arise from interactions of the field with a scalar as in the Higgs mechanism, which can pair up chiral spinors together and lend vectors another degree of freedom. But a massless scalar and a massive scalar have the same number of degrees of freedom. If we want a scalar mass to be technically

natural, it must come from some symmetry past simply the Lorentz group. We'll see some examples of how to arrange this in Chapter 3.

1.3.2   Spurion Analysis

An important tool for understanding symmetries and their violation is known as 'spurion analysis'. The basic idea is simple: for a theory which respects a symmetry except for the coupling c, this coupling parametrizes the breaking of the symmetry and any effects which violate the symmetry are proportional to c. More concretely, one assigns such couplings spurious transformation properties under the symmetry such that the action becomes invariant under the symmetry. Physically one can imagine that the observed values of such couplings come from the vacuum expectation values of some heavy fields which are above the cutoff of the theory. This can be viewed as imagining a UV completion where the explicit symmetry breaking in the low-energy effective theory comes microscopically from some spontaneous symmetry breaking, but the value of spurions is not dependent upon a particular realization of the UV completion.

We can quickly see the utility of this by looking at a simple example of a complex scalar field ϕ with an interaction which explicitly breaks the $U(1)$ global symmetry $\phi \to \phi e^{i\alpha}$.

$$S = \int \mathrm{d}^4x \left(-\partial_\mu \phi^\dagger \partial^\mu \phi - m^2 \phi^\dagger \phi - \frac{1}{3!}\lambda \phi^3 + \ h.c. \right), \qquad (1.135)$$

where "$+ \ h.c.$" denotes the addition of the Hermitian conjugate of the non-Hermitian interaction term. A naïve effective field theorist would say that our Lagrangian has no symmetries, and so we have no control and should expect that quantum corrections give us any polynomials in $\phi, \phi^\dagger$ at low energies.

However, we may note that if we assign λ a charge of -3 such that $\lambda \to \lambda e^{-3i\alpha}$, then the theory *is* invariant under that $U(1)$ global symmetry. So quantum corrections cannot

violate our spurious symmetry, and as a result we know that we can only generate terms like $\lambda^2\phi^6$, and there are no ϕ^4 or ϕ^5 interactions generated at any order in perturbation theory.

We can also usefully apply this to the example of technical naturalness studied in the preceding section. If ϵ is given the spurious transformation $\mathbb{Z}_2 : \epsilon \to -\epsilon$ then the $\epsilon\phi^3$ term is invariant. Then it's simple to see that no matter what other sorts of interactions we add, so long as they respect the $\mathbb{Z}_2$ symmetry we must have $\delta\epsilon \propto \epsilon$. But having added $y\phi\Phi^2$ we see that this term can also be made invariant with $y \to -y$, and this allows $\delta\epsilon \propto y$ as well.

An important real-world example where spurion analysis is useful is in understanding the flavor structure of the SM. With all masses turned off, the SM has a large global symmetry group $U(3)^5 = U(3)_Q \times U(3)_u \times U(3)_d \times U(3)_L \times U(3)_e = SU(3)^5 \times U(1)^5$, corresponding to arbitrary unitary reshufflings of the three generations of each fermion representation. These symmetries are explicitly broken by the Yukawa matrices which generate hierarchically different masses for the three generations.

$$\mathcal{L} \supset -Y_d^{ij} Q_i H d_j - Y_u^{ij} Q_i H^\dagger u_j - Y_e^{ij} L_i H e_j \tag{1.136}$$

where $i, j = 1, 2, 3$ are generation indices, and the Yukawa couplings are matrices in this generation space.

We don't understand why these hierarchies are present, but we can carry out a spurion analysis to see how worried we should be. We see that our theory will be invariant under the full flavor group if we assign the Yukawa matrices the following transformation

properties under the various $SU(3)$ symmetry groups

$$Y_d \sim (3, \bar{3}, 1) \text{ under } SU(3)_Q \times SU(3)_d \times SU(3)_u \tag{1.137}$$

$$Y_u \sim (3, 1, \bar{3}) \text{ under } SU(3)_Q \times SU(3)_d \times SU(3)_u \tag{1.138}$$

$$Y_e \sim (3, \bar{3}) \text{ under } SU(3)_L \times SU(3)_e \tag{1.139}$$

Since these are the *only* flavor-violating couplings in the SM and they are all in distinct spurious flavor representations, this tells us the quantum corrections to these matrices must be proportional to the matrices themselves e.g. $\delta Y_e \propto Y_e$. Thus this pattern of Yukawa couplings is stable under RG evolution to higher scales, and we are justified in thinking that the generation of this pattern may take place at large, currently-inaccessible scales.

This eases our minds about when we need to discover the origin of these flavor hierarchies, but this holds true only as long as these remain the only flavor-violating couplings. Fantastic work in precision flavor measurements and theory has provided lower bounds on the scale at which additional flavor violation can occur. Searches for flavor-violating processes have constrained these violations to take place at scales enormously higher than scales we are able to directly probe at colliders, which poses a puzzle. If there is new physics near the TeV scale, how is it arranged to respect the flavor structure of the SM? A phenomenological approach known as Minimal Flavor Violation [79] demands that *all* flavor violation is proportional to these Yukawa couplings, but no fundamental explanation for this is known. For recent introductions to flavor in the Standard Model, see e.g. [80, 81, 82].

1.3.3   Dimensional Transmutation

Perhaps the most important example of a Dirac natural field theory generating a small number is that of 'dimensional transmutation'. In particular, in quantum chromodynamics (QCD) the theory is 'asymptotically free'—meaning that the interaction strength vanishes in the far UV—and the gauge coupling g grows as one goes to lower energies. We skip the Nobel-worthy calculation (see e.g. Srednicki's chapter 73 [25]) and simply quote the results for the beta function of QCD (here parametrized via $\alpha_s = g^2/4\pi$), which dictates the dependence of the gauge coupling on energy. In $\overline{\text{MS}}$, the calculation finds

$$\beta(\alpha_s) \equiv \frac{\mathrm{d}\alpha_s}{\mathrm{d}\ln\mu} = -\frac{\alpha_s^2}{2\pi}(11 - \frac{2}{3}n_f) + \mathcal{O}(\alpha_s^3) \tag{1.140}$$

where μ is the energy scale of interest and n_f is the number of quarks with masses below μ, which at high energies is $n_f = 6$. Then if we know the gauge coupling at a fundamental scale like M_{pl}, we can follow the procedure discussed in Section 1.2.3 of running and matching to sequentially evolve the coupling down to low energies. We end up with a result like

$$\frac{1}{\alpha(M_{pl})} - \frac{1}{\alpha(\mu)} = \frac{b}{2\pi} \ln \frac{M_{pl}}{\mu} \tag{1.141}$$

where b is somewhere between the $11 - 6 \times 2/3 = 7$ value it takes above the top quark mass and the $11 - 3 \times 2/3 = 8$ value it has below the charm quark mass. This tells us that eventually the QCD coupling blows up in the infrared, and the theory becomes strongly coupled—we expect our perturbative understanding of the theory to break down. While there is no proof of the precise effects of this, there is strong evidence that this is responsible for the observed phenomenon of 'color confinement'— at low energies colored particles form bound states which are color-neutral. The intuition being that the gluon interaction is so strong that trying to pull quarks in a color-singlet apart from each other

requires so much energy that it is energetically favorable for a quark-antiquark pair to be created out of the vacuum and to end up with two color singlets. We may define a new scale Λ_{QCD} as being the energy at which α diverges

$$\Lambda_{QCD} \equiv M_{pl} e^{-\frac{2\pi}{b} \frac{1}{\alpha(M_{pl})}} \tag{1.142}$$

So for some reasonable fundamental coupling $\alpha(M_{pl})$ at high energies, the theory generates a new scale which is exponentially far removed from the fundamental physics. Since the mass of the proton is mainly from QCD binding energy, $m_p \sim \Lambda_{QCD}$ this explains the huge hierarchy $m_p \lll M_{pl}$. This is an extremely important mechanism and historically one of the first suggestions for how to generate the electroweak scale was by copying this strategy, as we'll discuss in Section 3.3.

Chapter 2

The Hierarchy Problem

2.1   The Higgs in the Standard Model

The physical question of the hierarchy problem is how to get an infrared scale v out of a microscopic theory whose degrees of freedom live at the much higher scale Λ, with $v/\Lambda \ll 1$. The tools introduced in Section 1 have already provided a window into why this can be difficult in a quantum field theory. Our aim in this section is to expand on that notion for the generation of the electroweak symmetry breaking scale in the Standard Model, where this scale is provided by the Higgs mass. In the Standard Model the Higgs mass is not technically natural, so the example discussed in Section 1.3 suggests the issue that may appear.

It is well-known that the Higgs plays a central role in the Standard Model, but the tagline 'it provides mass' doesn't go far enough in underscoring its importance. The Higgs is needed because the fermions of the Standard Model have a *chiral* spectrum: There are no representations with opposite charges under the full SM gauge group. This means that there are no gauge-invariant fermion bilinears, so no fermions can be paired up to form mass terms. If the Standard Model were *not* chiral, we could write mass terms

directly in the Lagrangian. In such a case there's no reason to expect small masses for the fermions, and indeed in the familiar case of right-handed neutrinos—which are SM gauge-neutral themselves, so can have Majorana masses—we generally expect them to be very heavy. In some sense the natural expectation for such 'vector-like' (non-chiral) fermions would be to have Planck-scale masses, as in the absence of other particle physics, this is the only scale.

So macroscopic structure in the universe is solely made possible by the chiral nature of the Standard Model.[10] There may well be other vector-like sectors which indeed contain Planck-scale masses. The fact that the particles which comprise us have a weakly-coupled description where quantum gravitational effects are suppressed — and so notions like locality and Riemannian geometry work well — would not hold in such a vector-like sector.

Thus the fact that the Standard Model is chiral and so requires the Higgs mechanism to provide masses answers a deep and important question about our place in the universe. But it doesn't provide a full answer, as the scale at which the Standard Model sits still needs to be generated somehow. And in the absence of a mechanism to make it light, we must worry once more about losing our macroscopic existence with a mass scale which is again naturally of the only other mass scale, the Planck scale.

So, far more than the hierarchy problem being a small detail to clean up after having empirically verified the structure of the Standard Model, the question of why $m_H \ll M_{\mathrm{pl}}$ has serious physical importance.[11] If we want an answer to why we live in a world with

---

[10]We note that it's also true that the existence of a macroscopic *universe* relies on the smallness of the cosmological constant, which is the other pressing fine-tuning issue present in the Standard Model. This way of viewing the naturalness problems of the Standard Model has been beautifully articulated by Nima Arkani-Hamed in [83] and in many seminars.

[11]It's worth noting that in the complete absence of the Higgs, electroweak symmetry is broken by the QCD chiral condensate [84, 85]. It's interesting to ponder why Nature did not choose to let QCD confinement solely fill the role, but we know empirically that there is EWSB at ~ 100 GeV scales, so we need to understand the generation of that separate scale.

macroscopic structure, we must grapple with the hierarchy problem.

This section is devoted to understanding the technical statement of this physical question in the framework of effective field theory. We pursue this by introducing, discussing, and refuting some common confusions about the hierarchy problem.

2.2 *Non*solutions to the Hierarchy Problem

In this section I will introduce a few common confusions and misconceptions about the hierarchy problem. Discussion and refutation of these arguments provides a natural backdrop for introducing how the hierarchy problem should be properly understood and why it is important.

2.2.1 An End to Reductionism

A first point of confusion is that the Higgs mass is a free input parameter in the Standard Model, so a natural objection is that we should just set $m_H = 125$ GeV and call it a day. Indeed, this hits on a basic and important point: *There is no hierarchy problem in the Standard Model.* The hierarchy problem exists for a more-fundamental theory which *predicts* the Higgs mass—that is, one in which the Higgs mass is an *output* parameter.

We can evince this in a simple toy model of a scalar ϕ interacting with other general fields ψ_i where a tree-level, 'bare' mass term is allowed by the symmetries. This is our toy version of the Standard Model, in which the scalar mass is likewise an input.

$$S = \int d^d x \left[-\frac{1}{2}(\partial_\mu \phi)^2 - \frac{1}{2}m_0^2 \phi^2 - V(\phi) - \phi g \mathcal{O}(\psi_i \psi_j) \right] \tag{2.1}$$

Of course, as is familiar, m_0 itself is not measurable. When one calculates the two-

point function of ϕ perturbatively in couplings[12] one finds quantum corrections $\Gamma^{(2)}(p) = p^2 + m_0^2 + g^2 \left(m_1^2 + \dots \right) + \mathcal{O}(g^4)$, where these are generically large because the mass is not technically natural—there's no symmetry protecting it. When one measures the mass of ϕ with e.g. some scattering experiment, it is $m_{\text{phys}}^2 = m_0^2 + g^2 m_1^2 + \dots$ which one measures. And luckily so, because m_1^2 may well be formally infinite in a continuum quantum field theory, and a similarly infinite bare mass term is necessary to end up with the correct finite physical mass. Indeed, we are justified in this theory in choosing m_0 such that m_{phys} matches the measured value. This is just the familiar procedure of renormalization, stretching back many decades and first understood in the context of quantum electrodynamics. To define QED one needs to input some definitions of the electron mass and the electromagnetic coupling based on experimental data, and these two inputs then determine all other predictions of the theory—e.g. the differential cross section of Coulomb scattering, the lifetime of positronium, and anything else you could hope to measure.

However, consider now a theory which has a global $SU(2)$ symmetry which is spontaneously broken at a high scale M. We want to understand how to get a light scalar degree of freedom out of this theory—that is, we've measured $m_{\text{phys}} \ll M$. This is a toy model of a Grand Unified Theory, where the microscopic physics exists at a high scale M_{GUT}, and indeed it was in this guise that the hierarchy problem was first recognized. Let's say our light scalar degree of freedom ϕ originated from a doublet $\Phi^a = (\varphi, \phi)^\intercal$. Our microscopic theory now does not have a bare mass term for ϕ but rather solely for Φ as a whole, since it must respect the symmetry. A difference in the masses of ϕ and φ can only come from the spontaneous breaking of the symmetry—let's say when another fundamental scalar Σ gets a vev $\nu = M$. This vev is a physical, measurable

---

[12]Perturbative calculations are an expansion in couplings, *not* $\hbar$ [86], though this subtlety is commonly elided.

parameter related to the mass of Σ and its self-interactions. Our action is controlled by the symmetries, as ever.

$$S = \int d^d x \left[-\frac{1}{2}(\partial_\mu \Phi)^\dagger(\partial^\mu \Phi) - \frac{1}{2}M_0^2 \Phi^\dagger \Phi - \lambda_0 \Phi^\dagger \Sigma \Sigma^\dagger \Phi - V(\Phi) \right] \qquad (2.2)$$

where M_0 is the bare mass of the Φ doublet and λ_0 is its bare interaction strength. In the absence of any other scales we generally expect $M_0 \sim M$ and $\lambda \sim \mathcal{O}(1)$. In this theory there is no reason for the values of these parameters to be connected to each other in any way.

When Σ gets a vev, $\langle \Sigma \rangle = (0, \nu)^\intercal$, it breaks the $SU(2)$ symmetry and gives a mass splitting between the two degrees of freedom in Φ, since $\langle \Sigma \rangle^\dagger \Phi = \nu \phi$. We then have the masses $m_\varphi^2 = M_0^2$, $m_\phi^2 = M_0^2 + \lambda_0 \nu^2$. In this theory our tree-level inputs are M_0 and λ_0 (and the interactions controlling the value of ν, which for simplicity we don't write down) and the scalar mass m_ϕ is an *output.* In fact here the hierarchy problem occurs at tree-level, simply as a result of wishing to produce a small mass via splitting a multiplet. If we wish to have, say, $M \simeq 10^{16}$ GeV and $m_\phi \simeq 100$ GeV—the values of the GUT scale and the electroweak scale in the real world—we need to fine-tune λ enormously so it takes a value like $-1.00000000000000000000000000001 \times \frac{M_0^2}{\nu^2}$.

Of course when we look at our theory at loop level there will again be quantum corrections to our tree-level parameters m_0 and λ_0, and again it will be their corrected values which are physical and measurable $M_{\text{phys}}^2 = M_0^2 + gM_1^2 + \dots$, $\lambda_{\text{phys}} = \lambda_0 + g\lambda_1 + \dots$. But if our theory is renormalizable, we know that quantum corrections will merely change the values of these parameters, and not the operators we have. The point is that the quantum corrections are $SU(2)$ invariant, so the masses of both ϕ and φ will receive the same loop contributions. We will then still predict the mass of ϕ as $m_\phi^2 = M_{\text{phys}}^2 + \lambda_{\text{phys}} \nu^2$. Now at the level of inventing the theory we may still tune these parameters to get a small

m_ϕ, but we're tuning *physical, observable* parameters.

In the theory described by Equation 2.1, one might have also said that we needed to fine-tune m_1 against m_0 in order to get a small m_ϕ, especially if we calculated that the quantum correction m_1 was large. But there the fine-tuning was of a different sort, since m_0 and m_1 were only ever observable in the combination m_{phys}. Here the fine-tuning has a much sharper meaning. This tuning translates into a physical demand on our theory that at high energies the strength of the interaction between our two scalars Φ, Σ for some reason has a value *extremely* close to $-M_{\text{phys}}^2/\nu^2$, despite having nothing to do with either of these parameters. The tuning is now a physical feature of our theory and demands explanation.[13]

So we see explicitly that the hierarchy problem is present when the light scalar mass is an *output* of the theory, rather than an input. If one is so inclined, one can say the words that the Higgs mass is simply an input, but this possibility spells the end of scientific reductionism. It is indeed *conceivable* that this is how the universe works. However, we know there is physics beyond the Standard Model at smaller length scales, and our best ideas for what those could be involve theories where the inputs are defined in the ultraviolet and the Higgs mass comes out. Whether the Higgs ultimately originates as a component of a larger multiplet, or a bound state of fermions, or an excitation of a string, we expect that the Higgs mass is a parameter that comes out in the low-energy theory.

---

[13]Let me mention parenthetically a confusion one may encounter if one reads older literature on the hierarchy problem in GUTs. It was common to speak of having to 're-tune the parameters at every order in perturbation theory', as if imagining an algorithmic process where one first computed a tree-level prediction, tuned that to be correct, and then computed the one-loop corrections, re-tuned those parameters to get it right again, etc. This is framed as being 'worse' than just requiring 'one' set of tunings. This is moronic, for the simple physical reason that Nature does not compute via perturbation theory. There is a physical problem, which is how to get the electroweak scale out of other physical parameters in the theory. Whether you compute the predictions in perturbation theory, or on a lattice, or whilst standing on your head is immaterial.

2.2.2   Waiter, there's Philosophy in my Physics

> You can always use the history of physics to illustrate
> any polemic point you want to make.

> Nima Arkani-Hamed

Now having exhibited that getting a light scalar truly does involve some fine-tuning of physical parameters, one may still say '*So?*'. In the real world, we've observed (the analogue of) m_ϕ, but the physics we've discussed at the heavy scale M is new physics, and we don't have experimental measurements of λ telling us the value is *not* that perfect value to get our light scalar. One might thus object that there's no fundamental inconsistency, and as long as our theory can fit the data anything else is *just philosophy*.

But the criterion of it being *not literally impossible* for a theory to fit the data is an incredibly low bar, and scientists always use additional criteria to select theories. While there is ultimately a degree of subjectivity in any notion of 'naturalness', this is really the same subjectivity that one constantly uses in science to decide which of two explanations for some data to accept.

A simple (approximately) historical example of this can be seen in epicyclic theories of the motion of the planets. The Ptolemaic, geocentric model of the universe predicted at first that the heavenly bodies orbited the Earth in circles, but eventually astronomical data was accurate enough to show that the motion of the planets and sun around the earth was not circular. In the Ptolemaic model, this was dealt with by adding an *epicycle*, the suggestion that the heavenly bodies moved on smaller circular orbits about their circular orbit around the Earth. As astronomical observations became more and more detailed over the ensuing centuries, multiple layers of epicyclic motion were needed to explain the data—circles on circles on circles. The description never stopped working though; if

you'd like to, you may describe the orbit of any solar system object via $r(\theta) = \sum\limits_{n=1}^{N} r_n \sin \frac{\theta}{n}$, and for some large but finite N you'll be able to fit any orbit to within observational precision.

So why did sixteenth century physicists favor Kepler's laws and heliocentrism? The discriminating factor is manifestly *not* which model better-fitted the data. Rather, the choice comes down to Occam's razor, to explanatory power, to simplicity and to fine-tuning of parameters. Physicists favor theories which do more with less—theories which explain more about the world while requiring fewer inputs. This is a subjective bias about how we think the universe should work, and it's possible that this philosophy will ultimately fail—but it's been working well thus far.[14]

While keeping the above intuition firmly in the back of our minds, it can be useful to introduce a mathematical classification of this fine-tuning, with the understanding that no such measure is god-given and so what to do with such a measure is up to us. We'll discuss a couple such schemes, the first being a mathematical formalization of the dependence of an output of interest on the values of the inputs. This has the benefit that it is intuitive and simple to compute, and so it is widely used in the particle physics literature. However, it lacks independence under how variables are parametrized and so can lead to misleading conclusions if used without care.

Furthermore, it will assign a measure of fine-tuning to individual points in the parameter space of a model, whereas we'd like to characterize the naturalness of a model

---

[14]As a semi-autobiographical aside, I had the honor and pleasure of being in the inaugural cohort of the Integrated Studies Program for Benjamin Franklin Scholars in the School of Arts & Sciences at the University of Pennsylvania. The program, founded and spearheaded by the classicist Prof. Peter Struck, offered a dedicated interdisciplinary experience wherein, each semester, three diverse fields gave courses offering perspectives around a central topic, which were concurrently collectively compared and contrasted. In my year we studied biology, anthropology, classics, political science, physics, and litera-ture, all taught by preeminent professors in their respective fields. After noticing a pattern, the group kept track of (among other things) how many times each professor mentioned, discussed, or appealed to 'beauty'. The winner in this regard was Prof. Vijay Balasubramanian—lecturing on the way the universe works—by a country mile.

as a whole—if a model only produces predictions that match the real world in a small region of its parameter space, that's another important element of fine-tuning [87]. For example, if new data removes all but a small fraction of the viable parameter space in a given model, we want to regard that model as being less natural afterward. An approach based on Bayesian statistics allows us to incorporate these issues and gives unambiguous comparisons of the relative naturalness of models upon the collection of new data [88], but loses out on simplicity.

A simple and often-used measure was introduced by Giudice and Barbieri [89], who suggested the definition

$$\Delta_X \equiv \left| \frac{\mathrm{d}\ln m^2}{\mathrm{d}\ln X} \right| = \left| \frac{X}{m^2} \frac{\mathrm{d}m^2}{\mathrm{d}X} \right| \tag{2.3}$$

which may be called a measure of the fine-tuning of the input parameter X necessary to get out the correct output parameter m^2. The logarithmic dependence naturally gives a measure of relative sensitivity and removes dependence on overall scale or choice of units. If Δ_X is large, this denotes a large sensitivity of m^2 to the value of X, and so implies that one must choose the value of X very carefully to get out the right physics. In the example in Equation 2.2 above we have $\Delta_\lambda = \frac{\lambda}{m_\phi^2} \nu^2 \simeq \frac{M^2}{m_\phi^2} \simeq 10^{28}$, indeed signaling enormous fine-tuning, and likewise $\Delta_{\nu^2} = \lambda \frac{\nu^2}{m_\phi^2}$. Contrast this with the familiar case of a seesaw mechanism where the light neutrino mass is given by a formula like $m_\nu^2 = m^4/M^2$, with M a heavy mass scale and m a weak scale mass. We can check whether this mechanism requires fine-tuning with $\Delta_{M^2} = \frac{M^2}{m_\nu^2} \frac{m^4}{M^4} = \frac{1}{m_\nu^2} \frac{m^4}{M^2} = 1$, and we find that seesaw mechanisms are natural. So Δ_X matches our intuition here, and can be quite useful.

However, for a model with free parameters a notion of fine-tuning at a single point in parameter space does not capture the full picture, and we should incorporate a notion of the volume of viable parameter space into our naturalness criterion [90]. This necessity

should be intuitive in the context of constraining models of new physics. Models which achieve their aim throughout parameter space are viewed more favorably than models which only work in some small region of their parameter space. For a relevant example, there are still corners of the MSSM parameter space that are natural under the Giudice-Barbieri measure. But the fact that the LHC has ruled out large swathes of this parameter space means that we should surely view weak-scale supersymmetry as less natural now than we did a decade ago. A pointwise measure of fine-tuning misses this.

An approach based on Bayesian statistics can be used to better match what we want from a measure of naturalness, as is discussed well in [78]. The definition of a model in a Bayesian framework requires priors on its free parameters $\{\theta_i\}$, $p(\theta_i|\mathcal{M})$, and different choices of $p(\{\theta_i\}|\mathcal{M})$ should be considered different models. We also require a prior probability $p(\mathcal{M})$ that the model $\mathcal{M}$ is true as a whole. After we receive data d, Bayes' theorem gives us the posterior probability for the model as

$$p(\mathcal{M}|d) = \frac{p(d|\mathcal{M})p(\mathcal{M})}{p(d)} \tag{2.4}$$

where $p(d|\mathcal{M})$ is known as the likelihood. Both $p(\mathcal{M})$ and $p(d)$ are explicitly subjective, but if we take the ratio of the posterior probabilities for two models $\mathcal{M}_1$ and $\mathcal{M}_2$ we find

$$\frac{p(\mathcal{M}_1|d)}{p(\mathcal{M}_2|d)} = \frac{p(d|\mathcal{M}_1)}{p(d|\mathcal{M}_2)}\frac{p(\mathcal{M}_1)}{p(\mathcal{M}_2)} \tag{2.5}$$

This expresses how the ratio of the likelihood of these two models changes after receiving new data. So while different physicists may disagree on the prior and posterior probabilities, the 'Bayes update factor' $B \equiv p(d|\mathcal{M}_1)/p(d|\mathcal{M}_2)$ is unambiguous and shows that the physicists agree on how the relative naturalness of the two models is affected by the new data.

The likelihood in a model with free parameters $\{\theta_i\}$ is calculated as

$$p(d|\mathcal{M}) = \int \left(\prod_i d\theta_i \right) p(d|\mathcal{M}, \theta_i) p(\theta_i|\mathcal{M}) \tag{2.6}$$

This is an integration over parameter space of the likelihood of producing the data in this model, weighted by the prior probability distribution we've placed on our parameters. This balances the competing effects of how well parameter points fit the data with the principle of parsimony—models with large regions of parameter space which don't fit the data are penalized. The need to compare models to define naturalness in a Bayesian formalism is easily seen by the fact that the likelihood for any new physics model decreases monotonically as more data is collected and previously-viable regions of parameter space are ruled out (in the absence of a discovery, of course).

An interesting playground for these ideas is the strong CP problem. In brief, this is the smallness of the so-called 'theta angle' θ in QCD, which controls the amount of CP breaking in the strong sector. While $\theta \in [0, 2\pi)$, empirical measurements now constrain $\theta \lesssim 10^{-10}$. Although θ is not technically natural, in the Standard Model it runs very slowly—since the other source of CP breaking is from the CKM matrix—such that if one sets it tiny in the UV, it stays small down to the IR. The Guidice-Barbieri measure would thus produce $\Delta_{\theta_{\mathrm{UV}}} \simeq 1$, as the measured value is insensitive to small changes in the UV value. And yet, the small theta angle is regarded as a naturalness problem, which can be justified in a Bayesian approach. The necessity of a prior does help quantify one's surprise at the small value of θ in the context of the Standard Model, but to really think about naturalness we need to have a comparison. We know of simple theories which produce vanishing θ starting from generic values of the parameters—for example, an axion naturally produces $\theta = 0$. Then if we have new data which pushes down the upper bound on θ, axion models receive large Bayesian update factors in relation to the

Standard Model. When we know of a simple model which automatically explains some data, it's puzzling if our current model requires precise choices for free parameters in order to explain the data.

We can see these notions play out for the hierarchy problem by comparing our toy model of a GUT to one with weak-scale supersymmetry. Our toy model in Section 2.2.1 produced a scalar through a cancellation of GUT-scale $\sim 10^{16}$ GeV contributions. A toy model with supersymmetry (to be discussed in Section 3.1) would remove sensitivity to the ultraviolet, and replace the scale of GUT-breaking with the effective scale of supersymmetry-breaking in the SM sector $\tilde{m}$. And while the GUT scale is (more or less) fixed by the running of SM couplings, the SUSY-breaking scale in the SM can be far lower. We use $\tilde{m}$ here as a one-parameter avatar of the scale of superpartners. The general prediction for supersymmetry before the LHC was $\tilde{m} \sim \mathcal{O}(100 \text{ GeV})$, with multiple species of superpartners appearing below the TeV scale. Such a model produces an improvement in Giudice-Barbieri tuning of $\delta\Delta \sim 10^{28}$ over our model without SUSY.

Let's take $p(\tilde{m}|\text{weak scale SUSY})$ to be a logarithmic prior from m_{low} to m_{high}, where an upper limit $m_{\text{high}} \sim 500 \text{ GeV} - 1 \text{ TeV}$ is justified by the requirement that the model give small Higgs mass corrections and a weak-scale dark matter candidate. If we collect collider data for a decade and find that much of the parameter space is ruled out, say with a limit $\tilde{m} \geq \mathfrak{m}$ we should update our thoughts on the naturalness of the model. This data has no effect on our non-SUSY GUT model, as it predicts no new light particles, but there has been a large effect on our SUSY model as its most favored parameter space has been ruled out. So we have a large Bayesian update factor.

$$B \equiv \frac{p(\text{LHC data}|\text{non-SUSY GUT})}{p(\text{LHC data}|\text{weak scale SUSY GUT})} \tag{2.7}$$

$$= \frac{1}{\left(\log \frac{m_{\text{high}}}{m_{\text{low}}}\right)^{-1} \int_{\mathfrak{m}}^{m_{\text{high}}} \frac{\mathrm{d}\tilde{m}}{\tilde{m}}}$$

$$= \log \frac{m_{\text{high}}}{m_{\text{low}}} / \log \frac{m_{\text{high}}}{\mathfrak{m}}$$

Of course supersymmetric extensions of the SM have a large number of parameters, and how to translate to an upper bound $\mathfrak{m}$ on our one-parameter version isn't well-defined, but certainly much of the previously-favored parameter space has been ruled out. The general lesson is that models which don't predict new visible states near the weak scale have received large Bayesian update factors from the LHC. This doesn't give a strict mandate for our overall relative belief, as one may argue there are good reasons to take a large prior for supersymmetry and expect always to find superpartners right around the corner. After all, even if $\tilde{m} \sim 1000$ TeV, it would still have $\delta\Delta \sim 10^{20}$ Giudice-Barbieri tuning better than the GUT without SUSY. But it does motivate further investigation of models which don't succumb to this issue, be they neutral naturalness modules which push the scale of new visible states up by a loop factor or more radical ideas about the origin of the electroweak scale.

Furthermore, we can directly input physics into making a sensible choice of the prior one places on a model, and there has been much discussion of justifying simple choices. The dictum of Dirac naturalness mandates priors peaked at $\mathcal{O}(1)$ values. But with a model which makes some parameter technically natural one, a prior that allows small values is justified, such as a logarithmic prior. An explicit example of this modification of priors by physics can be seen in the application of the Weak Gravity Conjecture to the hierarchy problem [91, 92, 93, 94, 95, 96, 97, 98, 11, 5, 99], which will be discussed in more detail in Section 4.3.1. This mechanism explicitly modifies the prior one should

have on UV theories by linking the notion of the Swampland—that some string vacua don't admit universes like ours—to the allowed range of Higgs masses. This addresses the hierarchy problem by constructing a model where the priors are forced by UV physics to favor a light Higgs.

While these measures of naturalness are useful to help us clarify our expectations, we emphasize again that they must be used sensibly. But it's clear that some notion of naturalness appears solely from the axioms of probability, and is indeed baked-in to the practice of science inquiry. That's not to say that we should be epistemologically committed to the naturalness of the universe, but we can still see it as a useful guide toward new physics which has worked well in the past.

2.2.3 The Lonely Higgs

> All models are wrong, but some are useful.

George E. P. Box [100]

Robustness in the Strategy of

Scientific Model Building (1979)

There is another obvious suggestion that is useful to discuss: perhaps the Higgs does not interact with any physics at higher mass scales, such that despite in principle worries about the hierarchy problem there is no hierarchy problem *in practice*. The physics which can destabilize the Higgs mass must, as seen in the above example, both be heavier than the Higgs and interact with it in order to generate a contribution. Since the Higgs mass in the SM is not technically natural (that is, it is UV sensitive), large mass corrections from such particles are generic, as we saw in Section 1.3. One can explore the idea that perhaps there are no such particles. This faces a number of challenges.

The first difficulty is that we *know* there must be new physics. The Standard Model
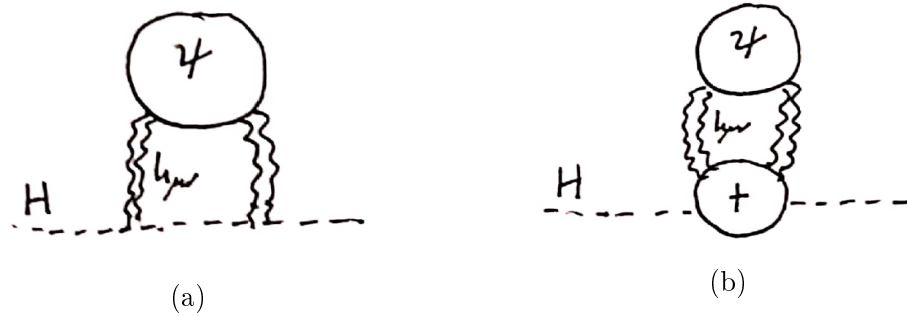
Figure 2.1: In (a), a representative two-loop diagram giving gravitational corrections to the Higgs mass from a new dark fermion ψ. These diagrams do not destabilize the Higgs mass. In (b), a representative three-loop diagram in which the gravitons couple to an off-shell top quark, which is no longer proportional to the Higgs mass.

does not explain neutrino masses nor dark matter, though it is possible that both of these be resolved without introducing new heavy particles. But a deeper issue is that we know the SM cannot be a fundamental theory because the Landau pole in hypercharge makes it inconsistent. This demands that *something* must happen to rid the theory of this pole, and it will interact with the Higgs because the Higgs is hypercharged. This is one motivation for thinking that something like Grand Unification must take place, and of course its breaking introduces a heavy fundamental scale. One can try to get around this by appealing to quantum gravity coming in at scales below that of the Landau pole. But the Higgs certainly interacts gravitationally, so for this program to succeed one needs a quantum gravitational theory which does not introduce any scales. This is interesting to explore, but does not seem to be the way the universe works, though we leave detailed criticism of this idea to the gravity theorists.

Furthermore, even if somehow all heavy particles are neutral under the SM gauge groups, there is no way for them to escape gravitational interactions with the Higgs. This leads to irreducible three-loop corrections to the Higgs mass [101], as depicted in Figure 2.1. Consider a fermion ψ with mass M_ψ. The obvious two-loop diagram where the graviton couples to the Higgs directly gives a correction proportional to m_H^2, as the

graviton coupling to a massless on-shell particle vanishes at zero momentum. However, we can draw a diagram where a ψ loop talks gravitationally to an off-shell top loop contribution to the Higgs two-point function, and integrating this particle out yields a correction

$$\delta m_H^2 \simeq \frac{y_t^2}{(16\pi^2)^3} \frac{M_\psi^4}{M_{pl}^4} M_\psi^2 + \text{ subleading.} \tag{2.8}$$

The powers of M_ψ appear because the only other possible mass scale for the numerator is $m_t = y_t v$, and the top loop power-law correction to the Higgs mass does not vanish in the $v \to 0$ limit. The sensitivity of the Higgs mass is softened by three loop factors as well as by the Planck mass from the gravitational couplings, but insisting that $\delta m_H^2 \lesssim (1 \text{ TeV})^2$ places an upper 'naturalness' limit on such fermions of $\sim 10^{14}$ GeV. While far better than the ~ 1 TeV limit for SM-charged particles, this is still well below the Planck scale and amounts to an enormous constraint on UV physics. In fact the problem is a bit worse than this estimate, as we should sum over all SM loops that couple to the Higgs, but this suffices already to see the problem.

So asking for the Higgs to be lonely enough to cure the hierarchy problem is a humongous requirement on the ultraviolet of the universe, and this approach faces a number of important hurdles. We mention that there is work on interesting theories which touch on some of these points, but as a whole Nature seems not to have taken this approach.

2.2.4   Mass-Independent Regulators

One may hear the statement that the hierarchy problem disappears if you use a mass-independent regularization scheme, for example dimensional regularization. Unlike the prior nonsolutions we've considered, this one is definitively incorrect. The mass scale one introduces in EFT is a stand-in for genuine physical effects of any sort which appear at shorter distance scales. So the cutoff regularization is useful for seeing an avatar of the

Figure 2.2: The one-loop diagram for a scalar quartic interaction contributing to a scalar mass.

hierarchy problem even when one does not know the ultraviolet theory. With a mass-independent scheme, one must instead put in specific short-distance physics to see the problem, but we can easily see the general issue.

As a simple example, take a theory with two real scalars - our light ϕ and a heavier φ. If we impose a $\mathbb{Z}_2$ symmetry for simplicity, the action is

$$S = \int d^4x \left[-\frac{1}{2}(\partial_\mu\phi)^2 - \frac{1}{2}m_0^2\phi^2 - \frac{1}{2}M_0^2\varphi^2 - \frac{\lambda}{4}\phi^2\varphi^2 \right] \tag{2.9}$$

Doing continuum effective field theory with dimensional regularization and the $\overline{\text{MS}}$ renormalization scheme, we must upgrade the masses and couplings to running parameters which depend on the renormalization scale μ, as usual. Since our renormalization scheme is mass-independent, if we want to study physics at an energy scale $\mu \sim m_\phi \ll M_\varphi$, we should implement the decoupling theorem by hand. We integrate out the heavy degree of freedom at the scale $\mu \sim M_\varphi$ and match to a low-energy effective field theory which only contains the low-energy degree of freedom ϕ. So let us go ahead and integrate out the heavy scalar φ. The effect on the ϕ mass comes from the simple diagram of Figure 2.2.

We go to general dimension $d = 4 - \epsilon$ and replace $\lambda \to \lambda\tilde{\mu}^\epsilon$, where $\tilde{\mu}$ is an arbitrary scale which soaks up the mass dimension of λ away from $d = 4$. The resources needed to

compute the integrals for general dimension and to take the limit $\epsilon \to 0$ can be found in Srednicki's textbook [25].

$$
\begin{aligned}
-i\delta m^2 &= \frac{-i\lambda\tilde{\mu}^\epsilon}{2} \int \frac{d^d k}{(2\pi)^d} \frac{-i}{(k^2 + M_\varphi^2)} \\
&= \frac{-i\lambda\mu^\epsilon}{2} \int \frac{d^d \bar{k}}{(2\pi)^d} \frac{1}{(\bar{k}^2 + M_\varphi^2)} \\
&= -i\frac{\lambda\tilde{\mu}^\epsilon}{2(4\pi)^2} \Gamma(-1 + \frac{\epsilon}{2})(4\pi)^{\frac{\epsilon}{2}}(M_\varphi^2)^{1-\frac{\epsilon}{2}} \\
&= i\frac{\lambda}{2(4\pi)^2} M_\varphi^2 \left[\frac{2}{\epsilon} - \gamma_E + 1 + \ln 4\pi - \ln M_\varphi^2 + 2\ln\tilde{\mu} + \mathcal{O}(\epsilon) \right] \\
&= i\frac{\lambda}{(4\pi)^2} M_\varphi^2 \left[\frac{1}{\epsilon} + \frac{1}{2} + \ln \frac{\mu^2}{M_\varphi^2} + \mathcal{O}(\epsilon) \right]
\end{aligned}
\tag{2.10}
$$

Where γ_E is the Euler-Mascheroni constant, and we have Wick rotated $k^0 \to i\bar{k}^d$, integrated in general dimension, expanded in the limit $\epsilon \to 0$ and defined $\mu^2 \equiv 4\pi\tilde{\mu}^2 e^{-\gamma_E}$ to soak up the annoying constants. Indeed, we see that there's no quadratic divergence, which ultimately is due to the fact that scaleless integrals vanish in dimensional regularization $\int \frac{d^d k}{k^n} = 0$, as must be true simply by dimensional analysis.

Now we follow the $\overline{\text{MS}}$ renormalization scheme by adding a counterterm which cancels off the divergent piece and we match at $\mu = M_\varphi$ to ensure that our low-energy EFT produces the same predictions as the UV theory, as discussed in Section 1.2.3. This gives us

$$
S = \int d^d x \left[-\frac{1}{2}(\partial_\mu \phi)^2 - \frac{1}{2}\left(m_0^2 + \frac{\lambda}{(4\pi)^2} M_\varphi^2 \right) \phi^2 - \ldots \right]
\tag{2.11}
$$

We see that the high-energy degree of freedom φ contributes a *threshold correction* when we flow to lower energies and remove it from the spectrum. While there was never a quadratic divergence, we still found a large quadratic correction to the mass of the scalar ϕ which is proportional to the scale of new physics.

This underscores the importance of not getting confused by unphysical features of

renormalization. There is a physical issue, which is the sensitivity of the physical low-energy scalar mass to the physics in the ultraviolet. Indeed if we tame the UV in different ways, we find different avatars of this sensitivity. It's true that a Wilsonian cutoff acts as a stand-in for arbitrary scaleful physics, which is why it's more direct to see the issue in that picture, but the same *physical* problem appears regardless of the regularization.

2.3   The General Approach

After discussing those misconceptions, we may give a one-sentence description of the hierarchy problem:

> In a theory of physics beyond the Standard Model where the Higgs mass is an output, physical parameters must be finely-tuned in order to produce a mass which is far below the scale of new physics, in tension with the principle of parsimony.

With that in hand, we are prepared to delve in to how the hierarchy problem may be solved in the next section. Our discussion below will not take place within a UV complete extension of the Standard Model, so one might worry that we are attacking a problem without knowing its source. While true, the point from Section 2.2.3 is that the sensitivity to UV physics is so general that we expect to need a mechanism which stabilizes the Higgs mass to *whatever* new heavy physics is out there. Our toy calculation of the relative naturalness of SUSY already evinces this point—SUSY tamps UV sensitivity no matter what it is, so the details of the UV completion are immaterial. As a result of this idea, we will mostly worry just about finding a way to produce a light scalar and assume that it can be embedded into whatever UV completion, rather than committing to a particular one. Of course it's possible that interesting mechanisms to produce an IR scale do rely on particular properties of the UV, and we'll discuss this important

idea in Section 4.3. But even there, our goal is solely to produce a light scalar which is compatible with the Standard Model, rather than to write down a full theory of the universe on all scales.

Chapter 3

The Classic Strategies

> Without going out of my door
> I can know all things on earth
> Without looking out of my window
> I could know the ways of heaven
>
> ────────────────────────
>
> George Harrison satirizing the
> past decades of particle theory
> in light of LHC data
> *The Inner Light* (1968) [102]

Fantastic Symmetries and How To Break Them

In large part the story of particle physics over the past decades is the story of attempts to solve the hierarchy problem. Much theoretical effort has been put into understanding interesting symmetries and mechanisms for breaking them, and more generally ways that small numbers can pop out of physical theories; and much experimental effort has focused on locating empirical hints of these ideas. However, the past few years have seen

many practitioners turn their attention toward topics like dark matter, cosmology, and astrophysics. And for good reason—on the experimental side, this is largely where the new data is and will be for the foreseeable future, and at the purely theoretical level the hierarchy problem has become a lot more challenging, as we will argue below. But this has lead to a new generation of particle theorists who are largely unfamiliar with the fantastic and brilliant ideas which drove the field in the prior couple decades.

Despite the fact that we will argue below that these ideas largely appear to not be the way the world works at the weak scale, understanding this prior work can be enormously helpful for inventing novel ideas in the future. It is with this in mind that we introduce below the basics of a variety of interesting ideas and methods in particle physics against the backdrop of their relevance to the hierarchy problem. These are ideas that have not yet made their way into standard textbooks on field theory, but are nevertheless essential topics for students of particle theory to absorb. We will endeavor to explicate the core of these ideas in the simplest models possible, and will largely avoid discussing phenomenological considerations past producing a light scalar. The discussion will not be at the level of depth required for research in the field, but will hopefully be a nice overview of interesting topics for which references to serious introductions and reviews will be provided as well.

So how does one solve the hierarchy problem? The classical solutions may be conceptually divided into two steps. First one introduces some structure above the electroweak scale which protects the Higgs mass from large contributions due to UV physics. This could be something like a new symmetry which forbids a scalar mass term, or a modification to spacetime on small length scales, or the dissolution of a non-fundamental Higgs into component fields.

However, the Higgs is not exactly massless, which is due to the fact that whatever structure we add is not a feature of the low-energy Standard Model. There must thus be

some IR dynamics that break that UV structure at the electroweak scale to ensure that we end up with the Standard Model at low energies. Depending on the UV structure this may be something like spontaneous symmetry breaking or moduli stabilization or dimensional transmutation.

There are two big categories of classical solutions. One is to find a field-theoretic mechanism which prevents contributions to the Higgs mass in the UV. Supersymmetry is the prime example here. The other is to bring the fundamental cutoff of the theory down to the infrared, such that in the UV there's no Higgs to talk about. This is exemplified by composite Higgs theories or theories where the cutoff of quantum gravity is lowered to the weak scale.

To evince these strategies, we'll go through a couple examples of ways to forbid scalar masses and to break those structures. Our aim here is not to construct realistic theories of the Higgs but rather to understand these general principles, so we'll study simple toy models which allow us to appreciate the essential points.

3.1   Supersymmetry

> Superpartners aren't essential
>
> But would have been consequential
>
> Such wasted superpotential
>
> Super once, super twice
>
> Super chicken soup with rice
>
> ――――――――――――――――――――
>
> Maurice Sendak on his disappointment with
>
> the LHC run 1 data
>
> Lost Stanza of *Chicken Soup with Rice* [103]

Supersymmetry exploits a loophole in the classic Coleman-Mandula theorem [67] by introducing *fermionic* generators, which in layman's terms turn bosons into fermions and vice-versa. By the Haag–Łopuszański–Sohnius theorem [104], this is the unique extension to the Poincaré algebra. Since we know that symmetries tend to make physics easier, it is not surprising that supersymmetry is an indispensable tool in high energy theory, regardless of how or whether it is realized in the real world. Some useful general introductions to supersymmetry in $d = 4$ and its application to the real world are Terning's book [105], Martin's periodically-updated lecture notes [106], and Shih's video lectures [107], in roughly increasing order of friendliness to neophytes.

In a supersymmetric theory fields come in multiplets which include particles of different spins (so called 'supermultiplets') all having the same mass and quantum numbers. We add fermionic generators Q_α and $Q_{\dot\alpha}^\dagger$, called supercharges, with the defining

(anti)commutation relations

$$\left\{Q_\alpha, Q_{\dot\alpha}^\dagger\right\} = 2\sigma^\mu_{\alpha\dot\alpha}P_\mu \quad \{Q_\alpha, Q_\beta\} = 0 = \left\{Q_{\dot\beta}^\dagger, Q_{\dot\alpha}^\dagger\right\} \tag{3.1}$$

$$[Q_\alpha, P_\mu] = 0 = \left[P_\mu, Q_{\dot\alpha}^\dagger\right] \tag{3.2}$$

where P_μ is the generator of spacetime translations and $\sigma^\mu = (1, \vec{\sigma^i})$ with σ^i the Pauli matrices. These may be determined simply by writing down all objects with the correct index structure. For later use, recall that spinor indices are raised/lowered with the invariant antisymmetric symbols $\epsilon_{\alpha\beta}, \epsilon_{\dot\alpha\dot\beta}$, as used for example in defining the conjugate invariants $(\overline{\sigma}^\mu)^{\dot\alpha\alpha} = \sigma^\mu_{\beta\dot\beta}\epsilon^{\alpha\beta}\epsilon^{\dot\alpha\dot\beta}$.

We want to find irreducible representations of the supersymmetry algebra, called supermultiplets. Since $P_\mu P^\mu = m^2$ commutes with the generators Q_α, $Q_{\dot\alpha}^\dagger$, the different particles in a supermultiplet will have the same mass. As an example of how to generate supermultiplet states, consider a massive particle. We can go to its rest frame, where it has momentum $P_\mu = (m, 0, 0, 0)$ with m its mass. Then the supersymmetry algebra greatly simplifies to $\left\{Q_\alpha, Q_{\dot\alpha}^\dagger\right\} = 2m\mathbb{1}_{\alpha\dot\alpha}$. Then we see that this is just a Clifford algebra of raising and lowering operators. We define a lowest weight state, or Clifford vacuum $|\Omega_s\rangle$ such that it is annihilated by the undotted generators

$$|\Omega_s\rangle = Q_1 Q_2 |m, s, s_z\rangle \tag{3.3}$$

$$\Rightarrow Q_1 |\Omega_s\rangle = 0 = Q_2 |\Omega_s\rangle \tag{3.4}$$

Now we can use the dotted generators as raising operators to generate the entire multiplet.

$$|\Omega_s\rangle \tag{3.5}$$

$$Q_1^\dagger |\Omega_s\rangle \,, Q_2^\dagger |\Omega_s\rangle \tag{3.6}$$

$$Q_1^\dagger Q_2^\dagger |\Omega_s\rangle \tag{3.7}$$

A single fermionic supersymmetry generator must change the spin of a state by $\frac{1}{2}$. Starting at the top with a spin j particle gives us states of spin $j - \frac{1}{2}$ and $j + \frac{1}{2}$ on the middle line, and another state of spin j on the bottom line. In $d = 4$ the supermultiplet formed from a vacuum state of spin $\frac{1}{2}$ is called a 'vector multiplet' and contains four states with spins $(0, \frac{1}{2}, \frac{1}{2}, 1)$. That formed from spin 0 is called a 'chiral multiplet', and has states with spins $(0, 0, \frac{1}{2})$. Note that this is fewer degrees of freedom, since negative spins are not allowed. Beginning with spins higher than $\frac{1}{2}$ leads to states with spins greater than 1, which will take us into supergravity and will not be necessary for our purposes.

We could repeat this exercise for massless supermultiplets, labelling states by their energy and helicity $|E, \lambda\rangle$. We would find a Clifford algebra with only one set of raising/lowering operators, and find supermultiplets with helicities λ and $\lambda + \frac{1}{2}$ for some starting λ. Then CPT invariance would force us to add states of $-\lambda$ and $-\lambda - \frac{1}{2}$.

Merely from the definition of the symmetry group there are already a few interesting immediate results. For a start, we show that physical states have nonnegative energy in a supersymmetric theory, and the vacuum energy is an order parameter for supersymmetry breaking. First, let's give a simple expression for the Hamiltonian operator of supersymmetry. We act on our anticommutation relation with $(\overline{\sigma}^\nu)^{\dot{\alpha}\alpha}$ and recall various

identities to note that $\sigma^{\mu}_{\alpha\dot{\alpha}}(\overline{\sigma}^{\nu})^{\dot{\alpha}\alpha} = 2\eta^{\mu\nu}$, which gives us

$$4P^{\nu} = (\overline{\sigma}^{\nu})^{\dot{\alpha}\alpha}\left\{Q_{\alpha}, Q^{\dagger}_{\dot{\alpha}}\right\}$$

$$\Rightarrow 4P^0 = 4H = \mathbb{1}^{\dot{\alpha}\alpha}\left\{Q_{\alpha}, Q^{\dagger}_{\dot{\alpha}}\right\}$$

$$4H = Q_1 Q^{\dagger}_{\dot{1}} + Q^{\dagger}_{\dot{1}} Q_1 + Q_2 Q^{\dagger}_{\dot{2}} + Q^{\dagger}_{\dot{2}} Q_2,$$

where we have used the fact that the zeroth component of the generator of spacetime translations is the generator of time translations, which is the Hamiltonian operator. Then we can write the energy of some state state S as

$$\langle S| H |S\rangle = \frac{1}{4}\left(||Q_1|S\rangle||^2 + ||Q^{\dagger}_{\dot{1}}|S\rangle||^2 + ||Q_2|S\rangle||^2 + ||Q^{\dagger}_{\dot{2}}|S\rangle||^2\right) \geq 0, \qquad (3.8)$$

so the energy of S is non-negative. Furthermore, consider a vacuum state $|0\rangle$ of our theory. In a standard QFT, the vacuum energy $\langle 0| H |0\rangle = E_{\text{vac}}$ is non-physical—we can just shift the Hamiltonian arbitrarily to remove it. But here, the supersymmetry algebra gives a preferred frame. If a vacuum state $|0\rangle$ is supersymmetric then it is annihilated by the supercharges $Q_{\alpha}|0\rangle = 0, Q^{\dagger}_{\dot{\alpha}}|0\rangle = 0$, otherwise the vacuum would not be invariant under supersymmetry transforms. This implies that it will have vanishing total energy $\langle 0| H |0\rangle = 0$. Conversely, if the vacuum state is non-supersymmetric, then its energy is strictly positive. We say supersymmetry is broken in such a state. Thus the vacuum energy acts as an order parameter for SUSY breaking.

Connected to that fact is that each supermultiplet contains the same number of fermionic and bosonic degrees of freedom. We can see this by defining an operator F which counts the fermion number of a state, so that bosonic states have eigenvalue 1 under $(-1)^F$, and fermionic states have eigenvalue -1. Since the SUSY generators interchange bosonic and fermionic states, they must anticommute with $(-1)^F$.

Now, for a given supermultiplet consider the states $|a\rangle$ with the same given four-momentum p_μ, $p_0 = E \neq 0$. Since the supercharges commute with P^μ, we know that these must form a complete set of states in this subspace $\sum_a |a\rangle \langle a| = 1$. Now consider the trace of the weighted energy operator $(-1)^F H/4$.

$$
\begin{aligned}
\sum_a \langle a| (-1)^F H |a\rangle &= \sum_a \langle a| (-1)^F Q Q^\dagger |a\rangle + \sum_a \langle a| (-1)^F Q^\dagger Q |a\rangle \\
&= \sum_a \langle a| (-1)^F Q Q^\dagger |a\rangle + \sum_a \sum_b \langle a| (-1)^F Q^\dagger |b\rangle \langle b| Q |a\rangle \\
&= \sum_a \langle a| (-1)^F Q Q^\dagger |a\rangle + \sum_b \langle b| Q (-1)^F Q^\dagger |b\rangle \\
&= \sum_a \langle a| (-1)^F Q Q^\dagger |a\rangle - \sum_b \langle b| (-1)^F Q Q^\dagger |b\rangle \\
&= 0
\end{aligned}
\tag{3.9}
$$

where we have suppressed the contracted spinorial indices. This implies that the number of bosonic degrees of freedom is the same as the number of fermionic degrees of freedom in our supermultiplet, which we found to be true in the example we considered above.

There is a beautiful formalism of 'superspace' which can be used to make super-symmetric theories far more transparent, but introducing this would be too large of a digression for our purposes.[15] We simply want to see the effects of supersymmetry on (in)sensitivity of low-energy physics to the ultraviolet, for which studying a simple the-

---

[15]Martin's notes [106] serve as a good introduction to traditional 'off-shell' superspace for $\mathcal{N} = 1, d = 4$ theories, and Thaler's TASI lecture notes [108] are also a fantastic resource. There is a related but distinct formalism of 'on-shell' superspace, which falls under the heading of the amplitudes/on-shell/S-matrix program. This was first introduced very early on by Nair [109] and was used to great effect by Arkani-Hamed, Cachazo, & Kaplan [110] much later. A pedagogical introduction to on-shell techniques including superspace can be found in the textbook by Elvang & Huang [111]. Until recently, the on-shell program was mostly restricted to massless particles. As it so happens, after Arkani-Hamed, Huang, & Huang [112] introduced a beautiful extension of the formalism to include massive particles, it was Timothy Trott, my undergrad mentee Aidan Herderschee, and myself who formulated an extension of the on-shell superspace formalism for massive particles [9]. The on-shell program is another fascinating line of work that I suggest any aspiring particle or field theorist learn about.

ory of chiral superfields will do. The Wess-Zumino model is the simplest such example which is not free, consisting of a single self-interacting chiral supermultiplet, and was historically the first non-trivial four-dimensional theory proved to be supersymmetric.

We may write down the Wess-Zumino Lagrangian as

$$
\mathcal{L} = \int \mathrm{d}^4 x \left(-\partial_\mu \phi^\star \partial^\mu \phi - m^2 \phi^\star \phi - i\psi^\dagger \bar\sigma^\mu \partial_\mu \psi - \frac{1}{2} m\psi\psi - \frac{1}{2} y\phi\psi\psi - \frac{1}{2} ym\phi^2\phi^\star - \frac{1}{4}|y^2|\phi\phi\phi^\star\phi^\star \right),
\tag{3.10}
$$

where for compactness we've left off the Hermitian conjugate terms. To avoid introducing certain technical complications we eschew the proof that this is indeed invariant under a supersymmetric transformation and instead evince its UV insensitivity. For fun we'll compute without assuming the masses are the same $m_\phi \neq m_\psi$, as can happen in the presence of soft breaking of the symmetry.

Let's look at the vacuum energy, which we'll calculate generally but schematically. A quantum harmonic oscillator has ground state energy $\pm\frac{1}{2}\hbar\omega$ for bosonic and fermionic states respectively, with the sign being familiar from the Casimir effect. If we consider a box of side length $V^{1/D}$, the energy of the fields inside it is

$$
E_0 = \sum_{\vec{k}}^{\text{bosons}} \frac{1}{2}\hbar\omega_{\vec{k}} - \sum_{\vec{k}}^{\text{fermions}} \frac{1}{2}\hbar\omega_{\vec{k}},
\tag{3.11}
$$

where $\vec{k} = (k_1, k_2, \ldots, k_D)(2\pi/V^{1/D})$, $k_i \in \mathbb{Z}$. Now in QFT each mode has energy $\omega_{\vec{k}} = \sqrt{k^2 + m^2}$, and as we make the box bigger $V^{1/D} \to \infty$, the sum turns into an integral

$$
E_0 = V \int \frac{d^d k}{(2\pi)^d} \left(\frac{1}{2}\sqrt{k^2 + m_B^2} - \frac{1}{2}\sqrt{k^2 + m_F^2} \right),
\tag{3.12}
$$

with m_B a boson mass and m_F a fermion mass, where the sum over species is implicit. We also recognize E_0/V as the vacuum energy density, denoted Λ, which we can write

Figure 3.1: The one-loop diagrams contributing to the scalar mass correction in the Wess-Zumino model.

as

$$\Lambda \simeq \frac{1}{2(2\pi)^D} \int d^d k \left(\sqrt{k^2 + m_B^2} - \sqrt{k^2 + m_F^2} \right).$$ (3.13)

Now if we specialize to $D = 4$ and introduce a cutoff k_{max} up to which we're confident that our description of particle physics holds, the schematic form is simply

$$\Lambda \sim k_{max}^4 \left(\overset{\text{bosons}}{\sum} 1 - \overset{\text{fermions}}{\sum} 1 \right) + k_{max}^2 \left(\overset{\text{bosons}}{\sum} m_B^2 - \overset{\text{fermions}}{\sum} m_F^2 \right) + \ldots$$ (3.14)

Now we see quite generally and explicitly that in a supersymmetric state the vacuum energy vanishes, since there are equal numbers of bosonic and fermionic fields with degenerate masses. Furthermore, spontaneous breaking of supersymmetry breaks the degeneracy but does not change the numbers of fields, so softly broken supersymmetry retains protection from the largest contribution.

Let's look now more sharply at the one-loop contributions to the scalar mass in the Wess-Zumino model, regularized with a hard cutoff Λ. Take care that we've written the Lagrangian in terms of two-component spinors, an exhaustive guide to which can be found in [113]. The three diagrams are shown in Figure 3.1, and their evaluation proceeds

as

$$-i\delta m^2 = (-im_\psi y)^2 \int \frac{d^4k}{(2\pi)^4} \frac{(-i)^2}{(k^2+m_\phi^2)^2} + (-iy^2) \int \frac{d^4k}{(2\pi)^4} \frac{(-i)}{(k^2+m_\phi^2)}$$

$$+ (-1)\frac{1}{2}(-iy)^2 \int \frac{d^4k}{(2\pi)^4} \frac{(-i)^2 \mathrm{Tr}\left[\sigma^\mu k_\mu \bar\sigma^\nu k_\nu\right]}{(k^2+m_\psi^2)^2} \tag{3.15}$$

$$= i|y|^2 \int \frac{d^4\bar k}{(2\pi)^4} \left[\frac{m_\psi^2}{(\bar k^2+m_\phi^2)^2} - \frac{1}{(\bar k^2+m_\phi^2)} + \frac{\bar k^2}{(\bar k^2+m_\psi^2)^2} \right] \tag{3.16}$$

$$= i\frac{|y|^2}{16\pi^2} \int_0^\Lambda d\bar k \ \bar k^3 \left[\frac{m_\psi^2-m_\phi^2-\bar k^2}{(\bar k^2+m_\phi^2)^2} + \frac{\bar k^2}{(\bar k^2+m_\psi^2)^2} \right] \tag{3.17}$$

$$= i\frac{|y|^2}{16\pi^2}(m_\phi^2-m_\psi^2)\log\Lambda + \ \text{finite}. \tag{3.18}$$

We see that the UV sensitivity of the scalar mass in this theory has disappeared, even if the two fields have different masses. In the limit of unbroken supersymmetry, the contribution vanishes identically.

This fact of removing the UV sensitivity of the mass of a scalar persists generally, no matter which other superfields are added, so long as supersymmetry is at most softly broken. The connection to the hierarchy problem is clear: If, in the UV, all fields come in supermultiplets, then the Higgs mass is protected from UV contributions.

Of course we do not observe mass-degenerate superpartners, so this soft supersymmetry breaking is a necessary feature of any implementation of supersymmetry to the real world. The Minimal Supersymmetric Standard Model [114] embeds each of our fermions in a chiral supermultiplet and each gauge boson in a vector multiplet. The Higgs sector must be enlarged to two chiral multiplets containing the up and down Yukawas respectively, as is necessary for anomaly cancellation[16].

---

[16]Perhaps the more urgent reason for needing two Higgs multiplets is that the interactions in supersymmetric theories are highly constrained by 'holomorphy', a full explanation of which here would require too much machinery but which leads to the conclusion that the same multiplet cannot have Yukawa interactions with both the up- and down-type quarks. However see [115] for the interesting possibility that at high energies only the up-type Yukawa interactions exist, and the down-type and

The question of supersymmetry-breaking is a very non-trivial one. At the level of a phenomenological accounting of possible soft breaking terms in the MSSM, there are 105 physical parameters [116]. However, constraints on flavor-violating couplings and on CP violation tell us empirically that the soft terms that appear must be very non-generic. In fact, looking at the MSSM in detail it turns out there are no places for supersymmetry-breaking to enter directly, and indeed there are general arguments that such breaking must take place in another, hidden sector and be indirectly communicated to the MSSM fields (see e.g. Martin's Section 7.4 [106]).

The origins of supersymmetry-breaking being a separate sector does force us to expand our model of particle physics, but on the upshot this sequestering means we can explore interesting phenomenology in sectors which are unconstrained. One can write down models where supersymmetry breaking is mediated by supergravity effects [117, 118, 119, 120, 121, 122, 123], communicated to the SM fields by our gauge bosons from a sector with new, massive SM-charged particles [124, 125, 126, 127, 128, 129], or takes place at a physically separate location in an extra dimension [130, 131, 132, 133, 134, 135, 136, 137, 138, 139]. A full discussion of the mechanisms and strategies for models of supersymmetry-breaking is beyond our scope, but we highly recommend Intriligator & Seiberg's lecture notes [140] as a general reference along with Martin's notes [106].

Of course we would like this phase transition to originate as spontaneous symmetry breaking, rather than explicitly putting it in by hand, since we want the far UV to be supersymmetric. Such spontaneous breaking requires the generation of a scale, and so it would be great if such a scale were generated *dynamically*, as in the dimensional transmutation we saw in QCD in Section 1.3.3. This would then be a natural mechanism for SUSY breaking. This phenomenological prospect lead to and benefited from a fantastic body of work understanding the details of supersymmetric gauge theories e.g.

---

charged lepton masses are induced by supersymmetry-breaking.

[141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151], which can be found reviewed in textbooks by Terning [105] and Shifman [152] and in TASI notes by Strassler [153].

We stated above that fields in a given supermultiplet share all the same quantum numbers, but there is in fact one exception. The supersymmetry algebra in Equation 3.1 is invariant under opposite rephasings of the supercharges $Q_\alpha \to Q_\alpha e^{-i\alpha}$, $Q_{\dot\alpha}^\dagger \to Q_{\dot\alpha}^\dagger e^{i\alpha}$. So there is a generator of a global internal symmetry that we may add that has nontrivial commutation relations with the supercharges:

$$[R, Q] = -Q, \qquad \left[R, Q^\dagger\right] = Q^\dagger. \tag{3.19}$$

This generator is known as an R-symmetry generator, and if the theory is invariant under an R-symmetry that means that each supermultiplet Φ can be assigned an R-charge r_Φ and the theory is invariant under transformations of each multiplet $\Phi \to \Phi e^{i r_\Phi \alpha}$, schematically, where Φ is the collection of fields in that multiplet. Since R does not commute with the supercharges, the different fields in the supermultiplet have different R-charges. For example if Φ is a chiral superfield consisting of (ϕ, ψ) then under a global rotation by α they transform as

$$\phi \to \phi e^{i r_\Phi \alpha}, \qquad \psi \to \psi e^{i(r_\Phi - 1)\alpha}, \tag{3.20}$$

which is simply because $|\psi\rangle \sim Q |\phi\rangle$.

It's important to emphasize that this R-symmetry is not part of the supersymmetry algebra, so one may have supersymmetric theories which do or do not implement R-symmetry. Nelson & Seiberg showed a fascinating connection between R-symmetry and supersymmetry-breaking [154]. They show roughly that for a low-energy Wess-Zumino model (possibly after having integrated out confined strong dynamics) as long as one has

a 'generic' potential (in the sense that a generic set of n equations in n unknowns has a solution), then a vacuum spontaneously breaks supersymmetry if and only if it spontaneously breaks R-symmetry. The reasoning is simply that such a symmetry imposes an additional constraint on the potential minimization equations, leading to a solution no longer being generically present.

For later use we mention the possibility of 'extended supersymmetry', where additional supercharges are added

$$\left\{Q_\alpha^A, Q_{\dot\alpha}^{\dagger B}\right\} = 2\sigma_{\alpha\dot\alpha}^\mu P_\mu, \qquad \left\{Q_\alpha^A, Q_\beta^B\right\} = 0 = \left\{Q_{\dot\beta}^{\dagger A}, Q_{\dot\alpha}^{\dagger B}\right\} \tag{3.21}$$

where $A, B = 1..\mathcal{N}$ index the supercharges. The construction of supermultiplets proceeds as before, but there are now more nonvanishing combinations of supercharges to act on the Clifford vacuum, so supermultiplets are enlarged. In four dimensions, the most supersymmetry we can have without gravity is $\mathcal{N} = 4$, which contains enough supercharges to relate the helicity -1 vector all the way to the helicity $+1$ vector; any more supercharges would necessarily yield particles of spins $3/2, 2$. If we are willing to include these degrees of freedom we can only go up to $\mathcal{N} = 8$ 'supergravity' (SUGRA), as more charges would lead to a theory with fundamental particles of spins $s > 2$ which is pathological[17]. The classic references for supergravity are Wess & Bagger [156] and Freedman & Van Proeyen [157]. Ultimately if supersymmetry is a field-theoretic feature of the ultraviolet of our universe we must have SUGRA as well, as it simply results from applying the supercharges to the graviton field, but we won't discuss SUGRA any further.

When we enlarge our superalgebra we also enlarge the (potential) R-symmetry group—the group of symmetries which do not commute with the supercharges—since we can now

---

[17]There's an important exception here, which is a theory which includes particles of *all* spins. This is necessary for string theory to operate, but has also led to the formulation of novel field theories commonly called 'Vasiliev gravity' [155].

shuffle around the supercharges in addition to rephasing them. We'll revisit this in Section 3.2.3 in the context of utilizing non-trivial R-symmetry representations to break supersymmetry.

3.2   Extra Dimensions

> I exist in the hope that these memoirs, in some
> manner, I know not how, may find their way to
> the minds of humanity in Some Dimension,
> and may stir up a race of rebels who shall
> refuse to be confined to limited Dimensionality.
>
> Edwin Abbott Abbott,
> *Flatland*, 1884 [158]

One of the most important ideas in theoretical physics developed in the latter part of the 20th century is that there may be additional spatial dimensions past the familiar three of our everyday experiences. Theories in which additional spatial dimensions are present were first studied in the context of unifying gravity and electromagnetism, first by Nordström [159] (*before* General Relativity!) and then by Kaluza [160] and Klein [161]. These ideas saw a resurgence of interest some half-century later with the advent of string theory, and the vision of all features of the universe being fundamentally geometrized. Against that backdrop, it is clearly prudent to consider the interplay of such theories with the puzzle of the hierarchy problem. As we shall see, extra dimensional theories can produce terribly interesting physics, and the possibilities are *manifold*.

3.2.1   Technology: Kaluza-Klein Reduction

To consider the possibility that there are additional microscopic dimensions, we develop a picture of the effects of fundamentally D-dimensional fields where $D = 4 + d$. On our manifold $M = \mathcal{M}_4 \times K$, where K is compact, we write down an action

$$S = \int_{\mathcal{M}_4} \int_K \sqrt{-\mathbf{g_{(4+d)}}} \mathcal{L}(\boldsymbol{\phi}, \mathbf{A}^M, \mathbf{g}^{MN}) \tag{3.22}$$

where our fields are in irreducible representations of the D-dimensional Poincaré algebra, and the Lagrangian manifestly obeys D-dimensional Lorentz invariance. We'll use boldface for D-dimensional fields and Latin letters for D-dimensional Lorentz indices. However, since K is compact, it places constraints on the mode expansions of our fields. Then when we want to study the effective four-dimensional theory we first need to decompose our fields into irreducible representations of the 4-dimensional Poincaré algebra. The D-dimensional vectors and tensors will become multiplets of 4-dimensional fields. Then we can explicitly integrate over the compact manifold K, which will produce a tower of states.

As an easy, explicit example, take the compact manifold to be the circle S^1 of length $2\pi R$, and consider a single complex scalar. We start with

$$S_5 = \int \mathrm{d}^4 x \int_0^{2\pi R} \mathrm{d}y \left[-\frac{1}{2}\partial_M \boldsymbol{\phi}^* \partial^M \boldsymbol{\phi} - \frac{1}{2}m^2 \boldsymbol{\phi}^* \boldsymbol{\phi} \right] \tag{3.23}$$

From introductory quantum mechanics, we know that the boundary conditions of being single-valued around the circle constrains the mode expansion of $\boldsymbol{\phi}$. We can write

$$\boldsymbol{\phi}(x^\mu, y) = \frac{1}{\sqrt{2\pi R}} \sum_{n=-\infty}^{+\infty} \phi_n(x^\mu) e^{iny/R} \tag{3.24}$$

where the normalization will produce canonically normalized kinetic terms in the $4-$dimensional action, and should generally be $\sqrt{\mathrm{Vol}(K)}$. Plugging this into the action gives

$$S_5 = \frac{1}{2\pi R} \int \mathrm{d}^4 x \sum_{mn} \int_0^{2\pi R} \left[-\frac{1}{2}\partial^\mu \phi_m^* \partial_\mu \phi_n - \frac{1}{2}\left(\frac{-im}{R}\right)\left(\frac{in}{R}\right)\phi_m^*\phi_n - \frac{1}{2}m^2\phi_m^*\phi_n \right] e^{i(n-m)y/R}$$
(3.25)

where the first and second terms come from either the $\mathcal{M}_4$ derivatives or the K derivatives acting on the mode expansion. The integral over the compact direction now gives, using orthogonality $\int_0^{2\pi R} \mathrm{d}y e^{i(n-m)y/R} = 2\pi R \delta_{nm}$, a $4-$dimensional action

$$S_4 = \int \mathrm{d}^4 x \sum_n \left[-\frac{1}{2}\partial^\mu \phi_n^* \partial_\mu \phi_n - \frac{1}{2}\left(m^2 + \frac{n^2}{R^2}\right)\phi_n^*\phi_n \right]$$
(3.26)

We see that we now have a *tower* of $4-$dimensional states that have arisen from the one $5-$dimensional scalar. There is a 'zero-mode' which has a mass given by the $5-$dimensional mass term, and then there are states with larger masses. If R is small, these will be heavy, and so we would only see them at, say, a particle collider. Note that all of these higher levels are doubly-degenerate, since n ranges over all the integers.

If we had included the graviton in this compactification procedure, we would see the $5-$dimensional graviton split up into

$$\mathbf{g}^{MN} = \begin{pmatrix} g^{\mu\nu} & \cdots & g^{\mu 4} \\ \vdots & \ddots & \vdots \\ \hline g^{4\nu} & \cdots & g^{44} \end{pmatrix}$$

so we see the emergence of a 4-dimensional scalar, vector, and two-index symmetric tensor. And it was in this context that compactification was originally studied as a

unification of general relativity and electromagnetism.

There's another very important effect of the compactification. Let's consider the Einstein Hilbert action

$$S = \frac{1}{16\pi G_N^{(4+d)}} \int_{\mathcal{M}_4 \times K} R \sqrt{-\mathbf{g^{(4+d)}}} \mathrm{d}^d x \tag{3.27}$$

Noting that the Ricci scalar $R \sim \partial^2$ always has mass dimension $[R] = 2$, we see that the $(4+d)$-dimensional Newton's constant must have $\left[G_N^{(4+d)} \right] = 2 - D = -2 - d$, so to rewrite the action in natural units we must define a $(4+d)$-dimensional Planck mass as $1/G_N^{(4+d)} = (M_{pl}^{(4+d)})^{2+d}$. If we take the metric to be independent of the K coordinates, then the integration over K just gives a factor of the volume V_K of the compact space

$$S = \frac{(M_{pl}^{(4+d)})^{2+d}}{16\pi} V_K \int_{\mathcal{M}_4} R \sqrt{-g^{(4)}} \mathrm{d}^4 x \tag{3.28}$$

Putting this into the form of the four-dimensional Einstein-Hilbert action, we find $M_{\mathrm{pl}}^2 = (M_{pl}^{(4+d)})^{2+d} V_K$. Taking $V_K \simeq R^d$ for some radius R, we see that if the compact dimensions are not Planck-sized $R \sim 1/M_{\mathrm{pl}}$ but larger for whatever reason, then the effective Planck mass at long distances can be much larger than the fundamental Planck mass.

Now we don't see zero-mode, different spin partners of our particles which fill out representations of the higher-dimensional symmetry, so this simple compactification cannot be the real way the world is. If we have compact dimensions, they must be such that the symmetry of zero-modes is somehow broken for the SM fields. This is important— we don't just want to have different $5d$ fields with zero modes or not, we need different four-dimensional components of them to have or not have zero modes. We'll solve this problem in Section 3.2.3.

3.2.2   Quantum Gravity at the TeV Scale

In the previous section we noticed that extra dimensions dilute the fundamental Planck mass in the higher-dimensional theory to produce a weaker effective four-dimensional Planck mass. So perhaps we can fix the hierarchy between the electroweak scale and Planck scale by lowering the fundamental scale of quantum gravity. Arkani-Hamed, Dimopoulos, and Dvali proposed that the fundamental Planck scale can be $M_{pl}^{(4+d)} \sim$ TeV with the weakness of four-dimensional gravity resulting from the dilution of gravitational flux into the extra d dimensions [162]. Using

$$(1 \text{ TeV})^{(2+d)/2} \simeq (M_{pl}^{(4+d)})^{(2+d)/2} \simeq M_{pl}^{(4)}/R^{d/2} \qquad (3.29)$$

where R is the radius of the extra dimensions, we find that a parsimonious $d = 3$ spherical dimensions of radius $R \simeq 1$ nm suffice to remove the hierarchy problem and accord with Eöt-Wash constraints on the behavior of gravity at scales down to $\sim \mathcal{O}(\mu\text{m})$ [163, 164, 165].

Now a nanometer is tiny compared to human scales, but the associated energy scale is $1/R \simeq 100$ eV, which is a scale we have quite a bit of information about. In particular, if the Standard Model fields were to propagate in all $4 + d$ dimensions then it would be easy to excite the 'winding modes' with momentum in the extra dimensions, and we should have observed many finely-space Kaluza-Klein resonances with the same quantum numbers. This is obviously not how the universe works, so to dilute gravity with large extra dimensions one must trap Standard Model fields to the four-dimensional manifold we know and love.

This can be done by imagining we live on a $(3 + 1)$-dimensional topological defect which is embedded in the larger $(4 + d)$-dimensional space. 'Topological defect' sounds exotic, but these are just (semi-)familiar non-perturbative objects such as the branes of

string theory [166, 167, 168], or the cosmic strings or domain walls that can appear in Higgsed gauge theories and which are introduced well in Shifman's textbook [152]. The original proposal suggests a weak-scale vortex in which zero-modes of our familiar fields are trapped, which is super cool.

Note that the new physics appearing at the TeV scale in this scenario is about as violent as you could imagine: quantum gravity appears at a TeV! This leads to a variety of fascinating signatures and constraints, even in the absence of concrete model of quantum gravity, though embeddings into string theory have also been found [169, 170, 171]. Very generically there are corrections to the Newtonian gravitational laws [172], all sorts of effects on precision observables [173], a loss of flux of high energy particles into the ambient space either astrophysically [174] or at a collider [138], violations of the global symmetries of the SM since quantum gravity does not respect them [175, 176, 177], and production of Kaluza-Klein gravitons [178] and black holes at TeV-scale colliders [179, 180, 181, 182]. This is a fascinating field which is well worth studying in detail, but unfortunately we do not have the space to do it justice. For more detail we refer to the reviews by Rubakov [183] and Maartens & Koyama [184], and the more introductory notes from Csáki [185], Kribs [186], Pérez-Lorenzana [187], Cheng [188], and Csáki, Hubisz, and Meade [189].

However, shrewd readers will be eager to point out that we *haven't* actually solved the hierarchy problem; we've merely traded the $m_H \ll M_{\rm pl}$ hierarchy for a $1/R^{(2+d)/2} \ll M_{\rm pl}^{(4+d)}$ hierarchy. And indeed, it can be difficult to stabilize the size of the extra dimensions in this scheme. But the conceptual leap of considering geometric solutions to the hierarchy problem is incredibly important and leads to many further interesting directions. The next ingredient we need is to control the existence of opposite-spin partners.

3.2.3   Technology: Orbifold Reduction

Let us consider a spacetime manifold $M = \mathcal{M}_4 \times K_G$ with $\mathcal{M}_4$ four-dimensional Minkowski space and K_G a $d-$dimensional compact 'orbifold'. An orbifold is constructed by 'modding out' a manifold K by a discrete symmetry group G. A manifold is a space that looks locally like Euclidean space; an orbifold is one which locally looks like the quotient space of Euclidean space quotiented by a finite group. In layman's terms, quotienting or modding out is just identifying points which are transformed into each other under G—our space becomes the space of equivalence classes of K under the action of elements of G.

As an illustrative example, consider the line $\mathbb{R}$ and the action of the discrete symmetry $\mathbb{Z}_2 : x \mapsto -x$. When we form the quotient space and identify points under this $\mathbb{Z}_2$ action, we find the half line $\mathbb{R}_{\geq 0}$. You'll notice that this space has a boundary at $x = 0$, which was a fixed point of the symmetry group. Mathematicians would say that the $\mathbb{Z}_2$ acts freely except at this point. This is a generic feature, and in fact what makes orbifolds interesting. We could have also imagined 'orbifolding' $\mathbb{R}$ by the translation $T(2\pi R)$, but this would have produced a compact manifold without boundary, S^1, because this symmetry has no fixed points. Part of the power of orbifolds comes because the quotient group structure gives us some information about what happens at fixed points - you could imagine just studying the half-line, but it seems natural to consider smooth structures on the manifold and then look at the effects under the identification. This should be evinced later in our examples.

Compactifying on orbifolds also gives us a way to cure our missing partner ills. Consider quotienting the circle by a $\mathbb{Z}_2$ which folds the circle over onto itself, $y \simeq 2\pi R - y$. This produces a line segment with boundaries at both ends $y = 0, \pi R$. We can alternatively think of this as modding out the real line by both translation and mirroring,
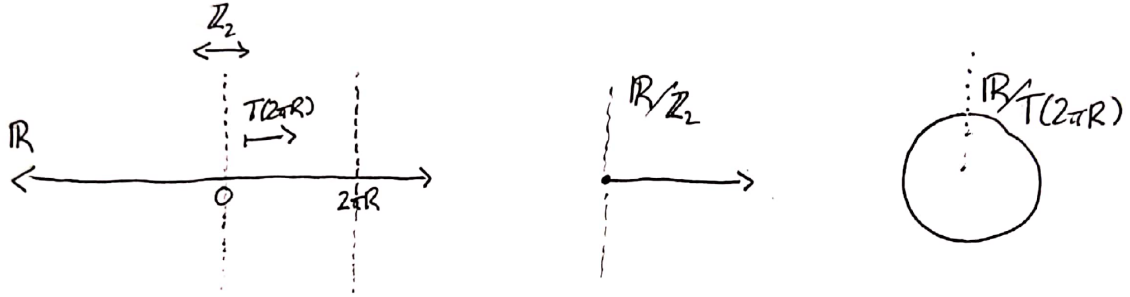
Figure 3.2: Schematic of the construction of an orbifold from the real line. The real line has a $\mathbb{Z}_2(0) : x \mapsto -x$ symmetry and a translational $T(2\pi R) : x \mapsto x + 2\pi R$ symmetry. Modding out by the former yields an orbifold because it has a fixed point, while modding out by the latter returns a regular manifold.
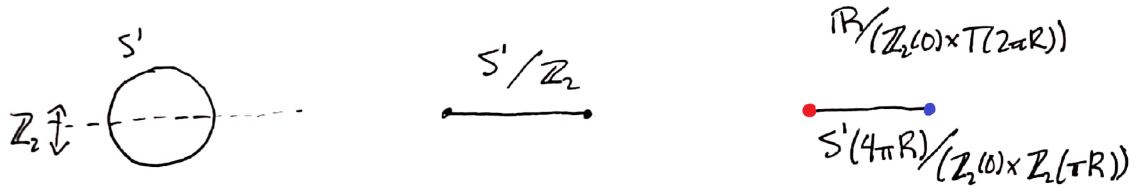
Figure 3.3: Schematic of the construction of an orbifold from a circle. Quotienting by $\mathbb{Z}_2$ corresponds to folding the circle over on itself. We can produce independent fixed points by uplifting first to the real line, or equivalently by folding the circle into quarters.

producing $\mathbb{R}/T(2\pi R) \times \mathbb{Z}_2(0)$. There are then *two* different sorts of fixed points - $y = 0$ is fixed by $\mathbb{Z}_2(0)$, and $y = \pi R$ is fixed by $T(2\pi R)\mathbb{Z}_2(0) \sim \mathbb{Z}_2(\pi R)$. You can then envision this as the circle of length $4\pi R$ with these two discrete identifications, that is $S^1(4\pi R)/\mathbb{Z}_2(0) \times \mathbb{Z}_2(\pi R)$, as depicted in Figure 3.3.

How does this affect the resulting compactification? As a first step should take our results and rewrite them in terms of eigenfunctions of our $\mathbb{Z}_2$ symmetry. We write

$$\phi(x^\mu, y) = \frac{1}{\sqrt{\pi R}} \left[\frac{1}{\sqrt{2}} \phi_0^{(+)}(x^\mu) + \sum_{n=1}^{+\infty} \phi_n^{(+)}(x^\mu) \cos ny/R + \sum_{n=1}^{+\infty} \phi_n^{(-)}(x^\mu) \sin ny/R \right]$$

$$(3.30)$$

where all we have done is rearrange things by defining

$$\phi_0^{(+)} \equiv \phi_0, \qquad \phi_{n>0}^{(+)} \equiv \frac{1}{\sqrt{2}} \left(\phi_n + \phi_{-n} \right), \qquad \phi_{n>0}^{(-)} \equiv \frac{i}{\sqrt{2}} \left(\phi_n - \phi_{-n} \right), \qquad (3.31)$$

where the superscript denotes their eigenvalue under the $\mathbb{Z}_2$. Now the way to change our S^1 compactification to an $S^1/\mathbb{Z}_2$ compactification is to impose in our $5-$dimensional action that ϕ transform with definite parity, which must be the case if the $\mathbb{Z}_2$ is a good symmetry[18]. You'll notice our free action does not demand a particular choice of parity for ϕ, so we are free to choose ϕ either even or odd. But in either case we are forced to get rid of half of our states. If ϕ is even we must set $\phi_n^{(-)} = 0$, and if ϕ is odd we must set $\phi_n^{(+)} = 0$—including, importantly, getting rid of the zero mode. If we wanted to include certain interactions in the higher-dimensional theory, that could dictate the transformation of ϕ—for example, the interaction term ϕ^3 would necessitate an even ϕ. This is just the same as we're familiar with in four dimensions.

More interestingly, let us consider the effect of the orbifold on larger Lorentz representations. Imagine a 5-dimensional gauge field with free action

$$S_5 = -\frac{1}{4} \int_M \mathbf{F}^{MN} \mathbf{F}_{MN} \qquad (3.32)$$

where $\mathbf{F}^{MN} = \partial^M \mathbf{A}^N - \partial^N \mathbf{A}^M$ is the field strength. The mixed terms here read $\mathbf{F}^{\mu 4} = \partial^\mu \mathbf{A}^4 - \partial^4 \mathbf{A}^\mu$, and this must transform coherently under the symmetry in order for the kinetic term to be invariant. But notice that $\mathbb{Z}_2 : \partial_4 \mapsto -\partial_4$. So $\mathbf{A}^4$ and $\mathbf{A}^\mu$ are forced to have opposite transformations under the reflection symmetry! Thus in this example one and only one of the four-dimensional vector and scalar multiplet has a massless

---

[18]Note that since we're getting our eigenmodes through this orbifold reduction our reduced action will still need to be produced by integrating from $y = 0$ to $y = 2\pi R$ because it's this domain over which our eigenfunctions are orthonormal.

zero mode; so orbifold compactifications generally enable us to have light fields without partners.

For another example of the usefulness of orbifolding, consider a five dimensional theory with the minimal amount of supersymmetry and we want to ensure our four-dimensional theory has only $\mathcal{N} = 1$ supersymmetry. First let's recall why we cannot have more than $\mathcal{N} = 1$ in four dimensions. As emphasized above, the SM is a chiral theory, wherein the different chiral components of its Dirac fields are in different gauge symmetry representations. This means that $\mathcal{N} = 2$ supersymmetry in four dimensions is too much for us, since the irreducible representations of super-Poincaré in that case don't allow for chiral matter. In particular, the $\mathcal{N} = 2$ vector multiplet must transform in the adjoint, and the $\mathcal{N} = 2$ hypermultiplet must transform in a real representation in order for it to be CPT self-conjugate, which means the two Weyl fermions must transform in conjugate representations. Thus if we want to imagine that the world came from a higher-dimensional supersymmetric theory, orbifold compactification is necessary.

The same problem appears just for five dimensional spinor fields, since there is no such thing as chirality in odd spacetime dimensions. So under dimensional reduction, one five-dimensional spinor breaks into two conjugate four-dimensional spinors, and you cannot get chiral matter. This is really the same problem as the above, since the possible supersymmety comes exactly from the possible spinor representations.

Compactifying a theory on an orbifold, rather than a manifold, will allow us to solve both of these problems—obtaining chiral matter, and reducing the amount of supersymmetry we have. First note that we can build $\mathcal{N} = 2$ supermultiplets out of two $\mathcal{N} = 1$ supermultiplets, which is really just thinking about a particular ordering for the construction of the supermultiplet by acting with supercharges. So for each $\mathcal{N} = 1$ superfield we want to have, we need to arrange for it to be even under the parity, and so have a zero-mode, and its partner superfield to be odd under the parity, and so only have

$n > 1$ Kaluza-Klein modes. In this way we get a set of zero modes which are chiral and $\mathcal{N} = 1$ supersymmetric, while the towers reflect the full $\mathcal{N} = 2$ supersymmetry and are non-chiral.

As an explicit example, consider an $\mathcal{N} = 2$ vector multiplet, which consists of a real scalar Σ, two fermions ψ^a with $a = 1, 2$, and a vector A_M. Under the $SU(2)_R$ symmetry, the scalar and vector are singlets and the fermions form a doublet. As in the example above, the $\mathcal{N} = 2$ Lagrangian for this multiplet dictates that the two $\mathcal{N} = 1$ multiplets living inside it transform differently under the $\mathbb{Z}_2$. So on the four-dimensional boundary one gets a zero-mode for either (A_μ, ψ^1) or $(\psi^2, A_4 + i\Sigma)$, corresponding to a vector multiplet or a chiral multiplet respectively. In slightly more group-theoretic language we can say that we've embedded the $\mathbb{Z}_2$ in the $SU(2)_R$, and have been forced by the physics to put the doublet in a two-dimensional representation

$$\mathbb{Z}_2 : \psi^a \to \sigma^a_{z\ b}\psi^b, \qquad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{3.33}$$

and to pair each of these fermions with two bosonic degrees of freedom, $\mathbb{Z}_2 : A_\mu \to \pm A_\mu, A_4 \to \mp A_4, \Sigma \to \mp\Sigma$. This gives us either an $\mathcal{N} = 1$ chiral multiplet or an $\mathcal{N} = 1$ vector multiplet on the boundary. More discussion and details can be found in Quirós' TASI notes [190], and in e.g. [191, 192, 193].

In fact orbifolding can do even more symmetry-breaking for us. As mentioned above, we can alternatively consider compactification on an interval as the orbifold $S^1(4\pi R)/(\mathbb{Z}_2(0) \times \mathbb{Z}_2(\pi R))$, which then means the two boundaries are fixed points of *independent* $\mathbb{Z}_2$ symmetries. In particular this means we can choose different embeddings of the $\mathbb{Z}_2$ in the full symmetry on the two ends $y = 0, \pi R$ [194, 195]. Now let's apply this technology to the $\mathcal{N} = 2$ case we considered above. We saw we had two

different choices for nontrivial embeddings of the $\mathcal{N} = 2$ into the R-symmetry to break the supersymmetry down to $\mathcal{N} = 1$ at the boundary. These correspond to choosing which half of the fields are even under the $\mathbb{Z}_2$ and so get zero-modes. Now that we have independent $\mathbb{Z}_2$s on either end, we can choose these to leave different $\mathcal{N} = 1$ symmetries unbroken.

In a microscopic picture, what we end up with is a theory on $\mathbb{R}^{3,1} \times I[0, \pi R]$ where the bulk is ($\mathcal{N} = 2$)-supersymmetric and the boundaries respect different $\mathcal{N} = 1$ supersymmetries. When we look at the effective four-dimensional theory at distances much larger than R, we have fully broken supersymmetry with the breaking being a nonlocal effect—one must be sensitive to physics on both boundaries in order to see the full breaking. As a result of this nonlocality, supersymmetry-breaking is guaranteed to be 'soft' the effects cannot depend on positive powers of UV scales, as we will discuss further momentarily. This leads to fantastically predictive models of BSM physics from around the turn of the century which really do read like they have it all figured out [196, 191, 197, 198, 199, 192, 200, 193]. It's worth understanding these in some detail simply because of how beautiful they are, but they uniformly lead to lots of (thus far) unseen structure near the TeV scale as KK partners become excited.

3.2.4   Nonlocal Symmetry-Breaking

Consider a single extra circular dimension of radius R with gravity and some gauge field. Upon restriction to the four-dimensional Lorentz group, those representations break up as $\mathrm{grav}^{(5)} \rightarrow \mathrm{grav}^{(4)} + \mathrm{vec}^{(4)} + \mathrm{scalar}$ and $\mathrm{vec}^{(5)} \rightarrow \mathrm{vec}^{(4)} + \mathrm{scalar}$. Each of these must be massless by five-dimensional gauge-invariance above the scale $1/R$, but below that they can pick up mass corrections up to that cutoff. Either of these possibilities then amounts to a mechanism for UV protection of a scalar mass. In implementing this to the

SM, the first possibility is known as the Higgs being a radion—the scalar which controls fluctuations of the size of the fifth dimension, $h - \delta h = \langle h \rangle = 1/R$, and the latter is denoted 'gauge-Higgs unification' for obvious reasons. A particular motivation for gauge-Higgs unification is as an extension of the strategy of grand unification. As mentioned in Section 1, the SM gauge bosons and fermions beautifully fit into representations of larger gauge groups, but in 4d GUTs the Higgs is left out in the cold as an extra puzzle piece. But 'grand gauge-Higgs unification' in 6d may allow even further frugality of ingredients.

A very interesting feature of this sort of construction is that the symmetry-breaking which is responsible for producing the light scalar is *nonlocal*—one must traverse around the fifth dimension to see the effects of the breaking. This should be intuitively clear, as at distances small compared to R the theory looks like five-dimensional Minkowski space. If it isn't obvious, I recommend musing by analogy on how we could tell whether or not the universe is a sphere with radius far, far larger than the Hubble scale $R \gg 1/H_0$.

As a result of this nonlocal symmetry-breaking, the scalar mass must be finite and calculable in the low-energy theory below $1/R$. There cannot be any sensitivity to ultraviolet energy scales Λ_{UV}, as this corresponds to a 'local counterterm', but in the theory above $1/R$ we know that the mass vanishes identically by gauge invariance, so such a counterterm cannot occur (recall our discussion in Section 1.2.1). This is powerful because symmetry-breaking generally leaves residual logarithmic dependence on large scales even when quadratic dependence has been eliminated, as we saw in the example of supersymmetry above.

Of course the theory can generate finite corrections to the mass of such a scalar at and below the scale $1/R$. If we look only at the low-energy effective theory then these look divergent, but we know they must get cut off at $1/R$. Since we know the high-energy theory we can ask how the scalar gets a mass at all, which seems to be impossible from gauge invariance. In fact the scalar mass comes from the Wilson loop wrapping the non-

trivial cycle around the fifth dimension [201, 202, 203, 204, 205, 206, 207].In any gauge theory there is a gauge-invariant operator called a 'Wilson loop',

$$W_C = \mathcal{P} \exp ig \oint_C \mathbf{A}_M \mathrm{d}x^M, \tag{3.34}$$

where g is the gauge coupling, C is some closed path through spacetime, and $\mathcal{P}$ denotes 'path ordering' of the operators along C in similarity to the time ordering in the definition of the Feynman propagator. In an Abelian theory, the gauge field transforms as $A_\mu \to A_\mu - \partial_\mu \Gamma(x)$ and we see that $\oint_C A_\mu \mathrm{d}x^\mu \to \oint_C A_\mu \mathrm{d}x^\mu - \oint_C \partial_\mu \Gamma \mathrm{d}x^\mu$. Integration by parts leaves us only with gauge transformations at the endpoints of C, of which there are none if C is a loop (equivalently if we started more generally looking at a 'Wilson line', we can say that for a loop the endpoints are connected and so their gauge transformations cancel each other). This remains true in non-Abelian gauge theories, though we do not go through the proof.

In our case we can consider a path which goes solely around the fifth dimension, in which case the Wilson loop contains only our 4d scalar $A_M \mathrm{d}x^M \to A_5 \mathrm{d}x^5$ and yet is fully gauge-invariant. Quantum corrections can thus generate the operator

$$\mathcal{L} \supset \sigma^4 \mathrm{Tr}\, \mathcal{P} \exp ig_5 \oint \mathbf{A}_5(x_\mu, x_5) \mathrm{d}x^5 + \text{ h.c.} \tag{3.35}$$

We note that on a circle the four- and five-dimensional gauge couplings are related as $g_5^2 = Rg_4^2$. Proceeding naïvely and expanding $\mathbf{A}_5(x_\mu, x_5) = \frac{1}{\sqrt{2\pi R}} \sum_{n=-\infty}^{+\infty} \phi^{(n)}(x_\mu)e^{inx_5/R}$, we see that this operator includes a four-dimensional mass for the zero-mode,

$$\mathcal{L} \supset \sigma^4 \mathrm{Tr}\mathcal{P} \exp ig_5 \oint \phi^{(0)}(x_\mu)\mathrm{d}x^5 + \text{ h.c.} \simeq \sigma^4 \cos\left[g_5\sqrt{2\pi R}\phi^{(0)}\right] \supset \sigma^4 g_5^2 R^2 \left(\phi^{(0)}\right)^2, \tag{3.36}$$

where we note that the natural size for the Wilson coefficient $\sigma \simeq 1/R$ yields a scalar mass expectation which is $m_\phi^2 \simeq g_4^2/R^2$, unsurprisingly as R is the only scale in the problem.

For a successful such model of gauge-Higgs unification, we need not only to get a light scalar but also to endow that scalar with dynamics pushing it to break electroweak symmetry[208, 196, 209, 210, 211, 212, 213, 214, 215, 216]. There is far too much rich physics involved here to mention, but discussions of the calculation of the one-loop effective potential in these models can be found in [217, 218, 219]. Gauge symmetry-breaking by a higher-dimensional gauge field component getting a vev is sometimes referred to as the 'Hosotani mechanism' [205, 206, 207].

As opposed to the Large Extra Dimensional model of the previous section, in these models the extra dimensions are 'universal'—all of the SM fields can propagate around the small dimension, not just gravitons. This means that all of our familiar fields are the zero modes of KK towers with spacing $\sim 1/R \sim m_H$. These KK partners have the same gauge charges, and so should be produced copiously in interactions with $\sqrt{s} \gg 1/R$, yet none have been observed at the LHC (see e.g. [220, 221, 222, 223] for some constraints). As a result, the lower bound on the KK scale is now far above the Higgs mass, which makes all of these sorts of models increasingly less attractive as solutions to the hierarchy problem.

3.3 Compositeness

We have saved for last what is, in some sense, the most obvious strategy to pursue. As discussed in Section 1.3, the Standard Model already breaks a symmetry and generates a mass scale in a natural manner with the chiral condensate of QCD. Perhaps Nature has only this one trick, and repeats the same mechanism to generate the weak scale. After

all, we mentioned above that the QCD condensate does in fact break the electroweak symmetry, just not at the right scale.

The idea is to introduce a new gauge sector which is asymptotically free and so becomes strong and confines at the electroweak scale. If the condensate has electroweak quantum numbers, then it breaks electroweak symmetry just as a Higgs fields would, but now without any Higgs. This strategy is known as 'technicolor', and was proposed in its simplest form by Weinberg [224] and Susskind [225]. A modern pedagogical introduction can be found in TASI lectures from Chivukula [226] or Contino [227], which has heavily influenced this discussion, and a more detailed classic review is from Hill & Simmons [228]. We introduce a technicolor sector which is an $SU(N_{TC})$ gauge group with N_D technicolor fundamentals which are electroweak doublets and their singlet partners, together enjoying a global $SU(2)_L \times SU(2)_R$ symmetry which is broken in the infrared to $SU(2)_V$ by confinement in the $SU(N_{TC})$ sector at $\Lambda_{TC} \sim v$. This structure exactly matches that of the SM QCD sector (save for the values of N_{TC}, N_D), so we expect all the same phenomenology, for example with composite technipions appearing close to the electroweak scale. But in this case since the technicolor chiral condensate is the leading breaking of the gauged $SU(2)_L$, the technipions will be predominantly 'eaten' and appear as the longitudinal polarizations of our W, Z bosons. Technicolor can 'Higgs' the electroweak gauge symmetry just like a fundamental Higgs field would.

However, the SM Higgs not only breaks electroweak symmetry but also gives mass to the quarks, and thus far we haven't introduced any coupling between the quarks and the technicolor sector. To get the appropriate couplings we can embed both of these gauge groups in a larger 'extended technicolor' group, $SU(N_{ETC}) \supset SU(3)_c \times SU(N_{TC})$. After this extended group undergoes spontaneous symmetry breaking at Λ_{ETC}, the broken

gauge bosons generate the appropriate four-Fermi interactions

$$\mathcal{L} \supset \frac{g_{ETC}^2}{\Lambda_{ETC}^2}(\bar{q}q)(\bar{\psi}_{TC}\psi_{TC}) + \mathcal{O}(\frac{\partial^2}{\Lambda_{ETC}^4}), \tag{3.37}$$

and then when the technicolor group confines at a scale Λ_{TC}, we see the emergence of quark masses

$$m_q \simeq \frac{g_{ETC}^2}{\Lambda_{ETC}^2}\left\langle \bar{\psi}_{TC}\psi_{TC} \right\rangle \sim \Lambda_{TC}\left(\frac{\Lambda_{TC}^2}{\Lambda_{ETC}^2} \right). \tag{3.38}$$

But this clearly suffices solely to generate a single quark mass scale, since the Yukawa coupling is originating from a single gauge coupling, so it was quickly realized that accounting for flavor physics required far more structure and significantly larger gauge groups [229, 230, 231, 232]. To generate the variety of quark mass scales with this mechanism would require a cascade of breakings from $SU(N_{ETC})$ down to $SU(3)_c \times SU(N_{TC})$. Furthermore the same massive gauge boson exchanges which generate those needed four-Fermi operators also generate four-quark interactions $(\bar{q}q)(\bar{q}q)$ which can lead to large flavor violation.

While there were insights on how various aspects of this could be tackled, the death-knell for technicolor came with Peskin & Takeuchi's parametrization of oblique corrections to the two-point functions of the electroweak vector bosons from BSM physics [233]. These efficiently parametrize deviations from the tree-level form of the vector boson propagators, and can be simply connected with experiment. Ensuing estimates for the sizes of these parameters in strongly-interacting models were very far from empirical measurements [234]. The program of technicolor lives on with 'walking technicolor', the idea that the confining dynamics may be due to a strongly-coupled gauge theory which behaves very differently from QCD [235, 236, 237, 238, 239, 240, 241]. This is too large a digression for us to introduce, but we mention that there is interesting recent work

relating the existence of 'walking' dynamics to proximity of the theory to a fixed point at complex value of the coupling [242, 243, 244, 245, 246].

However, there is another way that compositeness can be useful for us in securing a light electroweak-symmetry-breaking scalar: It will allow us to realize the dream of a pseudo-Nambu-Goldstone Higgs. Recall that whenever a continuous global symmetry is spontaneously broken, there appear scalar Goldstone bosons $\pi^i(x)$ which parametrize excitations about the vacuum state in the direction of the broken generators. Since the theory had a global symmetry, the potential along these directions is flat, and thus the Goldstones are massless. More strongly, they contain a shift symmetry:

$$\mathcal{L}(\pi^i(x)) = \mathcal{L}(\pi^i(x) + \xi) \tag{3.39}$$

where ξ is independent of x. This means they may only be 'derivatively coupled'; they may appear in the Lagrangian solely as $\partial_\mu \pi^i(x)$. Thus, a non-zero mass for such a scalar is technically natural—if such an operator is present, π^i is said to be a *pseudo*-Goldstone and corresponds to the breaking of an approximate symmetry. This is a familiar story in the context of the QCD condensate breaking the approximate chiral symmetry of the quarks, leading to light but not massless pions.

In the case of technicolor, the phenomenon of confinement of a strongly interacting sector is directly responsible for breaking the gauged electroweak symmetry, leading to no separation between the two scales Λ_{TC} and v. This means the technipions are immediately eaten, and there is no scale at which it looks like a scalar field is responsible for symmetry-breaking, which leads to large electroweak precision constraints. In a general composite Higgs scenario, we'll attempt to arrange for separation between the scale of confinement and the scale of electroweak symmetry breaking by having confinement break an enhanced *global* symmetry, leading to pseudo-Nambu Goldstone bosons whose masses are protected.

We then want to radiatively generate a potential for these pNGBs, leading to them getting a vev and breaking electroweak symmetry at a lower scale. A degree of separation between confinement at the scale f and EWSB at v will reduce the difficulties with electroweak precision constraints, as this scenario returns to the SM elementary Higgs sector in the limit $v/f \to 0$. Even more pressing, now that we have gained experimental access to energies close to v it's even more clear that scale separation is needed—confining dynamics lead generically to *lots* of resonances at the scale f, which would be seen in all sorts of ways.

The general setup is a group $\mathcal{G}$ of (approximate) global symmetries, of which a subgroup $\mathcal{H}_0$ is gauged. We will have strong dynamics at the scale f break the global symmetry to $\mathcal{H}_1$, and in full generality allow for some part of the gauge symmetries to be broken, such that the unbroken gauge symmetry is $\mathcal{H} = \mathcal{H}_0 \cap \mathcal{H}_1$. This leads to $[\mathcal{G}] - [\mathcal{H}_1]$ Goldstone bosons (where $[\cdot]$ is the dimension of a group) of which $[\mathcal{H}_0] - [\mathcal{H}]$ are eaten by the broken gauge bosons. The uneaten, light pNGBs transform non-trivially under the remaining gauge symmetry $\mathcal{H} \supset SU(2)_L \times U(1)_Y$, so we can hope to arrange for them to break this symmetry. In a realistic minimal model, we can have the strong dynamics not break any gauge symmetry, $\mathcal{H}_0 \subset \mathcal{H}_1 \Rightarrow \mathcal{H} = \mathcal{H}_0$. Such a minimal model may be constructed with $\mathcal{G} = SO(5) \times U(1) \to \mathcal{H}_1 = SO(4) \times U(1) \supset \mathcal{H}_0 = SU(2)_L \times U(1)_Y$ [247].

To evince the ideas in the simplest scenario possible, we'll discuss an even simpler model for a composite pNGB which then breaks a $U(1)$ gauge symmetry. This is just a toy model to understand the features, which has already been kindly worked out in the extensive review from Panico & Wulzer [248], and which we'll call a 'composite Abelian Higgs' model. To find the most minimal model we'll ask for a composite pNGB which breaks the smallest continuous gauge symmetry, $U(1)$, for which we simply need two uneaten degrees of freedom (since a charged scalar is necessarily complex). We can

make the even more minimal choice $\mathcal{H}_0 = \mathcal{H}_1 = \mathcal{H}$, as we're not worried about having additional unbroken global symmetries. Since $[U(1)] = 1$, we then just need to choose a group $\mathcal{G}$ with at least $[\mathcal{G}] = 3$ to get the right number of pNGBs. We'll study the breaking $SO(3) \to SO(2) \simeq U(1)$, which is especially nice because we have geometric intuition for the Lie algebras of these groups.

We'll study this using a 'linear sigma model', of which the 'chiral Lagrangian' describing the QCD pions is the most familiar example. Much of the general technology was developed by Callan, Coleman, Wess & Zumino [249, 250], and some modern introductions to this technology can be found in Schwartz' textbook [26], in a pedagogical review of Little Higgs models by Schmaltz & Tucker-Smith [251], and in exhaustive detail in the review by Scherer [252]. The big idea is one of bottom-up effective field theory: Given knowledge of the symmetry-breaking structure, we can cleverly parametrize our fields to easily see the structure of the Lagrangian which is demanded both before and after symmetry-breaking.

In our case we must start with an $SO(3)$-invariant Lagrangian of a fundamental field $\vec{\Phi}$, which is a familiar $3d$ vector.

$$-\mathcal{L} = \frac{1}{2}\partial_\mu\vec{\Phi}^\intercal\partial^\mu\vec{\Phi} + \frac{g_\star^2}{8}\left(\vec{\Phi}^\intercal\vec{\Phi} - f^2\right)^2, \tag{3.40}$$

where $SO(3)$ rotations act as $\vec{\Phi} \to g \cdot \vec{\Phi}$ with $g = \exp i\alpha_A T^A$, where T^A, $A = 1..3$ are the generators of $SO(3)$. The potential of $\vec{\Phi}$ is minimized for $\left\langle \left\|\vec{\Phi}\right\| \right\rangle = f^2$, so $\vec{\Phi}$ gets a nonzero vev and the symmetry is broken down to rotations keeping $\left\langle \vec{\Phi} \right\rangle$ fixed, which is simply $U(1)$. There are a two-sphere worth of vacua corresponding to possible angles for $\left\langle \vec{\Phi} \right\rangle$, which parametrize the Goldstone directions. We can make the split between

broken and unbroken generators explicit by parametrizing our field as

$$
\vec{\Phi} = \exp\left(i\frac{\sqrt{2}}{f}\Pi_i(x)\hat{T}^i\right)\begin{bmatrix} 0 \\ 0 \\ f+\sigma(x) \end{bmatrix} = (f+\sigma)\begin{bmatrix} \sin\left(\frac{\Pi}{f}\right)\frac{\vec{\Pi}}{\Pi} \\ \cos\left(\frac{\Pi}{f}\right) \end{bmatrix} \tag{3.41}
$$

where in the first equality $\hat{T}^i$ are the two broken generators, $\Pi_i(x)$ are the massless Goldstones corresponding to fluctuations along the vacuum manifold, and $\sigma(x)$ is the massive 'radial mode' giving fluctuations about the vev. We eschew writing down the generators explicitly and assert that in this case one finds the compact latter expression, with $\Pi = \sqrt{\vec{\Pi}^{\intercal}\vec{\Pi}}$. We can find an explicit expression for the interaction of the Goldstones and the radial modes by simply plugging this parametrization into the Lagrangian above. We indeed find a mass for σ of $m_\sigma = g_\star f$, massless pions $\vec{\Pi}$ and a tower of all possible interactions between these fields which are consistent with the symmetries.

We now have a theory of massless scalars transforming under an unbroken symmetry; our pions form a doublet of $SO(2)$ transforming as $\vec{\Pi} \to \exp(i\alpha\sigma_2)\vec{\Pi}$, corresponding to rotations about the unbroken $SO(3)$ generator. We can complexify by introducing $H \equiv \frac{1}{\sqrt{2}}(\Pi_1 - i\Pi_2)$ to view $SO(2) \simeq U(1)$ as acting $H \to \exp(i\alpha)H$. As it stands, H is an exact Goldstone boson, so cannot pick up a potential to then itself get a vev and break $U(1)$. However, when we gauge a $U(1)$ subgroup of our original $SO(3)$ global symmetry we're introducing explicit breaking of the symmetry and resultingly H becomes a pNGB and can pick up a mass. The tree-level effect of this gauging is simply to upgrade derivatives to gauge covariant derivatives, $\partial_\mu\vec{\Pi} \to D_\mu\vec{\Pi} = (\partial_\mu - ieA_\mu\sigma_2)\vec{\Pi}$.

Now as a result of $SO(3)$ rotations no longer being an exact symmetry, H is free to pick up a potential, which in general will be radiatively generated as a result of this gauging. It is easy to draw diagrams in which loops of our $U(1)$ gauge bosons generate a
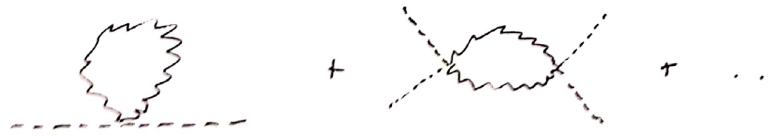
Figure 3.4: Representative diagrams contributing to the one-loop effective potential for a charged scalar.

nonzero mass and quartic for H, as in Figure 3.4, and in a realistic model the corrections from loops of top quarks will be especially important. These loop diagrams come along with an obvious cutoff: Above the compositeness scale f, the fields $\vec{\Phi}$ no longer exist in the spectrum. There is a beautiful general method for deriving the radiatively-generated potential for H by resumming the one-loop diagrams with various numbers of external H legs, known as the Coleman-Weinberg potential [253].

Unfortunately, absent any of other structure, finding $v/f \ll 1$ still requires some degree of cancellation between various contributions to the potential of H. However any amount of $v/f < 1$ will help alleviate the pressure from electroweak precision observables which thus far look empirically as expected for an elementary Higgs field, while still retaining the benefit of forbidding corrections to the Higgs mass above f. But one still can't get away from requiring lots of structure near the electroweak scale, which has not (yet) been observed.

There are a variety of important aspects and interesting directions we do not have the space to discuss. Even past understanding the best sorts of group structures to which to apply this strategy, it is clearly important to understand the sorts of field theories which can confine to break $\mathcal{G} \to \mathcal{H}_1$, as well as the detailed structure of the potential radiatively-generated by SM fields. Then it's important to explore the possibility of a natural structure which dictates $v/f < 1$—while there has been much work on this, we mention in particular the interesting strategy of 'collective' symmetry breaking in which a symmetry is broken only by an interplay between different couplings. This class of

models is known as the 'Little Higgs' [251, 254, 255, 256, 257, 258], and can be seen as a purely four-dimensional application of the strategy of nonlocal symmetry breaking through 'nonlocality in theory space' [259, 260], which is a fascinating topic.

Chapter 4

The Loerarchy Problem

> The great tragedy of science — the slaying of a
> beautiful hypothesis by an ugly fact.

> Thomas Henry Huxley
>
> *Biogenesis and Abiogenesis* (1870) [261]

4.1 The 'Little Hierarchy Problem'

We've seen in Chapter 3 a cadre of theories which can produce a light scalar naturally, and there's one feature all the classic approaches have in common: they predict new states with Standard Model charges close to the mass of the Higgs. This is seemingly inevitable simply from the structure of effective field theory—whatever extra structure protects the Higgs mass at UV scales must be broken close to the electroweak scale to allow the Standard Model, which does not have that extra structure. This feature means that smashing together protons at scales much greater than the electroweak scale would surely reveal the physics of whatever mechanism solves the hierarchy problem. And so the Large Hadron Collider was eagerly awaited to tell us which of these ideas was correct.

127

Yet even years before the LHC turned on, those who could clearly read the tea leaves were realizing that something was amiss with our naturalness expectations (see e.g. the 'LEP paradox' [262], also [263]), and exploring the idea that supersymmetry would not show up to solve the hierarchy problem (e.g. 'split supersymmetry' [264, 265, 266]). Perhaps there was something else present which made weak-scale supersymmetry unnecessary for protecting the Higgs.

Of course the LHC has been a fantastic success. It has confirmed for us the existence of a light Higgs resonance that looks SM-like, and made many great measurements of the SM besides. But rather than revealing to us which TeV-scale new physics kept the Higgs light, we've instead had a march of increasingly powerful constraints on new particles which couple to the Standard Model.

These null results for physics beyond the Standard Model from run 1 of the LHC rapidly popularized the idea that something else might be responsible for stabilizing the Higgs mass at the electroweak scale up to a higher scale where, say, supersymmetry came in. This line of thinking is termed the 'Little Hierarchy Problem'—the idea being that one of those classic solutions would appear at $\Lambda \sim 10$ TeV to solve the 'Big Hierarchy Problem' and explain why the Higgs mass was not at the Planck scale, leaving a smaller hierarchy of a couple orders of magnitude between m_H and Λ unexplained. Perhaps rather than minimal supersymmetry there was another module which provided this last bit of protection. But this has to be a special module to protect the Higgs mass without introducing new colored particles.

4.1.1 The Twin Higgs

The first such proposal in fact appeared before the LHC had even turned on. The mirror twin Higgs (MTH) [267] introduces a second copy of the SM gauge group and

states related to ours by a $\mathbb{Z}_2$ symmetry. Since these 'twin' states are neutral under the SM gauge group, they are subject only to indirect bounds from precision Higgs coupling measurements. The two sectors are connected solely by Higgs portal-type interactions between the two $SU(2)$ doublet scalars.[19] Subject to conditions on the quartic coupling, the Higgs sector enjoys an approximate $SU(4)$ global symmetry, and the breaking of this symmetry leads naturally to a pseudo-Nambu Goldstone boson. Seemingly magically, this structure is accidentally respected by the quadratically-divergent one-loop corrections to the Higgs potential, and the pNGB continues to be protected through one-loop from large corrections to its mass. The Twin Higgs thus allows the postponement of a solution to the 'Big Hierarchy Problem' until scales a loop factor $16\pi^2 \sim \mathcal{O}(100)$ above the Higgs mass.

While the space of Neutral Naturalness models has now been explored more thoroughly and we will discuss some generalities below, the mirror twin Higgs remains perhaps the most aesthetically pleasing of all these approaches and serves as a useful avatar for this general strategy. As a result, in Chapter 5 we consider cosmological signatures of the MTH specifically, so we give here a more-detailed introduction to the twin Higgs in particular.

The scalar potential in this model is best organized in terms of the accidental $SU(4)$ symmetry involving the $SU(2)$ Higgs doublets of the SM and twin sectors, H_A and H_B. The general tree-level twin Higgs potential is given by (see e.g. [268])

$$V(H_A, H_B) = \lambda(|H_A|^2 + |H_B|^2 - f^2/2)^2 + \kappa(|H_A|^4 + |H_B|^4) + \sigma f^2|H_A|^2. \qquad (4.1)$$

The first term respects the accidental $SU(4)$ global symmetry, as can be seen by writing it in terms of $\mathcal{H} = (H_A, H_B)^\mathsf{T}$, which transforms as a complex $SU(4)$ fundamental. The

---

[19]We return in Section 5.3 to the prospect of kinetic mixing between the two $U(1)_Y$ factors, which is also allowed by the symmetries.

second term breaks $SU(4)$ but preserves the $\mathbb{Z}_2$ and the final term softly breaks the $\mathbb{Z}_2$. In order for the $SU(4)$ to be a good symmetry of the potential, we require $\kappa, \sigma \ll \lambda$.

However, the gauging of an $SU(2) \times SU(2)$ subgroup constitutes explicit breaking of the $SU(4)$, so we should worry about whether quantum corrections reintroduce large masses for the would-be Goldstones when $SU(4)$ is broken. But writing down the one-loop corrections reveals a fortuitous accidental symmetry. The one-loop effective scalar potential gets the following leading corrections from the gauge bosons at the quadratic level:

$$V_{1-\text{loop}}(H_A, H_B) \supset \frac{9g_{2,A}^2}{32\pi^2}\Lambda^2|H_A|^2 + \frac{9g_{2,B}^2}{32\pi^2}\Lambda^2|H_B|^2 + \text{ subleading } \underset{\mathbb{Z}_2}{\to} \frac{9g_2^2}{32\pi^2}\Lambda^2|\mathcal{H}|^2, \quad (4.2)$$

where we see explicitly that if the $\mathbb{Z}_2$ symmetry is good at the level of the gauge couplings, then these largest one-loop corrections continue to respect the $SU(4)$. It is easy to see from here by power counting that this holds for all the quadratically-divergent pieces so long as the $\mathbb{Z}_2$ is a good symmetry for the interactions involved.

There is radiative $SU(4)$-breaking at the level of the quartic, since the $\mathbb{Z}_2$ symmetry no longer suffices to form the Higgses into an $SU(4)$ invariant. The coupling κ should naturally be of the order of these corrections, the largest of which comes from the Yukawa interactions with the top/twin top, $\kappa \sim 3y_t^4/(8\pi^2)\log(\Lambda/m_t) \sim 0.1$ for a cut-off $\Lambda \sim 10$ TeV (y_t being the top quark Yukawa coupling and m_t its mass). Requiring $\lambda \gg \kappa$ therefore implies $\lambda \gtrsim 1$. As the SM and twin isospin gauge groups are disjoint subgroups of the $SU(4)$, the spontaneous breaking of the $SU(4)$ coincides with the SM and twin electroweak symmetry breaking. This gives seven Goldstone bosons, six of which are 'eaten' by the $SU(2)$ gauge bosons of the two sectors, which leaves one Goldstone remaining. This will acquire mass through the breaking of the $SU(4)$ that is naturally smaller than the twin scale f. For future reference, it is convenient to define the real scalar degrees

130

of freedom in the gauge basis as $h_A = \frac{1}{\sqrt{2}}\Re(H_A^0) - v_A$ and $h_B = \frac{1}{\sqrt{2}}\Re(H_B^0) - v_B$, where $\langle H_A^0 \rangle = v_A$ and $\langle H_B^0 \rangle = v_B$.

The surviving Goldstone boson should be dominantly composed of the h_A gauge eigenstate in order to be SM-like. The soft $\mathbb{Z}_2$-breaking coupling σ is required to tune the potential so that the vacuum expectation values (vevs) are asymmetric and that the Goldstone is mostly aligned with the h_A field direction. The (unique) minimum of the Twin Higgs potential (4.1) occurs at $v_A \approx \frac{f}{2}\sqrt{\frac{\lambda(\kappa - \sigma) - \kappa\sigma}{\lambda\kappa}}$ and $v_B \approx \frac{f}{2}\sqrt{\frac{\sigma + \kappa}{\kappa}}$. The required alignment of the vacuum in the H_B direction occurs if $\sigma \approx \kappa$, which has been assumed in these expressions for the minimum. The consequences of this are that $v_A \approx v/\sqrt{2}$ and $v_B \approx f/\sqrt{2} \gg v$ (where v is the vev of the SM Higgs, although $v_A \approx 174$ GeV is the vev that determines the SM particle masses and electroweak properties), so that the SM-like Higgs h is identified with the Goldstone mode and is naturally lighter than the other remaining real scalar, a radial mode H whose mass is set by the scale f. The component of h in the h_B gauge eigenstate is $\delta_{hB} \approx v/f$ (to lowest order in v/f). Measurements of the Higgs couplings restrict $f \gtrsim 3v$ [269, 268], and the Giudice-Barbieri tuning of the weak scale associated with this asymmetry is of order $f^2/2v^2$.

The spectrum of states in the broken phase consists of a SM-like pseudo-Goldstone Higgs h of mass $m_h^2 \sim 8\kappa v^2$, a radial twin Higgs mode H of mass $m_H^2 \sim 2\lambda f^2$, a conventional Standard Model sector of gauge bosons and fermions and a corresponding mirror sector. The masses of quarks, gauge bosons, and charged leptons in the twin sector are larger than their Standard Model counterparts by $\sim f/v$, while the twin QCD scale is larger by a factor $\sim (1 + \log(f/v))$ due to the impact of the higher mass scale of heavy twin quarks on the renormalisation group (RG) evolution of the twin strong coupling. The relative mass of twin neutrinos depends on the origin of neutrino masses, some possibilities being $\sim f/v$ for Dirac masses and $\sim f^2/v^2$ for Majorana masses from the Weinberg operator. Mixing in the scalar sector implies that the SM-like Higgs couples

to twin sector matter with an $\mathcal{O}(v/f)$ mixing angle, as does the radial twin Higgs mode to Standard Model matter. These mixings provide the primary portal between the Standard Model and twin sectors.

The Goldstone Higgs is protected from radiative corrections from $\mathbb{Z}_2$-symmetric physics above the scale f. While the mirror Twin Higgs addresses the little hierarchy problem, it does not address the big hierarchy problem, as nothing stabilizes the scale f against radiative corrections. However, the scale f can be stabilized by supersymmetry, compositeness, or perhaps additional copies of the twin mechanism without requiring new states beneath the TeV scale. Minimal supersymmetric UV completions can furthermore remain perturbative up to the GUT scale [270, 271].

As mentioned, the collider constraints on twin Higgs models are very mild and pertain mostly to a lower bound on the soft breaking of the $\mathbb{Z}_2$. In this respect, the Twin Higgs naturally reconciles the observation of a light Higgs with the absence of evidence for new physics thus far at the LHC. The primary challenge to these models comes from cosmology due to the effects of additional light particles on the cosmic microwave background. We will discuss these issues in Chapter 5 and propose a natural resolution.

4.1.2   Neutral Naturalness or The Return Of The Orbifold

More broadly, the twin Higgs is just the simplest example of the more general 'Neutral Naturalness' paradigm in which the states responsible for stabilizing the electroweak scale are not charged under (some of) the Standard Model (SM) gauge symmetries [267, 272, 273, 274, 275, 276, 277, 278, 279, 280, 281, 282, 283], thus explaining the lack of expected signposts of naturalness.

In the symmetry-based solutions to the hierarchy problem discussed in Chapter 3, modifications of the Higgs mass were technically natural as a result of a continuous

symmetry which commuted with the SM gauge groups. This naïvely seems necessary to ensure that the necessary degrees of freedom are present and couple to the Higgs with the right strength to cancel divergences. How is the top quark contribution $\propto N_c y_t$ (where N_c is the number of colors) to be canceled if not by another colored particle with the same coupling to the Higgs, which gives the opposite contribution? Well we saw above that the twin Higgs nevertheless works with a quark charged under a *separate* gauge group. Indeed, at one-loop N_c is really just a 'counting factor' and we are free to get those three opposite-sign contributions in a variety of ways. To some extent the space of Neutral Naturalness models is an exercise in interesting ways to find that color factor. We'll be able to see that picture more clearly in the language of orbifold projection, which will also give us good reason to expect that these models with naïvely strange symmetries in fact do have nice UV completions.

In Section 3.2.3 we saw how orbifolds could be useful dynamically - that is, in affording a higher-dimensional, symmetric theory which at low energies looks like a less symmetric, four-dimensional theory due to boundary conditions imposed by the orbifold discrete symmetries. But these theories had interesting properties in their low-energy behavior below the scale of compactification; we didn't need to make reference to their origins in studying them.

Now we want to understand the variety of low energy theories we can get very generally, but only at the level of the zero-mode spectrum. Rather than decomposing our fields into modes and integrating over the compact manifold and noticing that only those fields invariant under the orbifold symmetry are left with zero-modes, we're going to skip to the answer and look at the spectrum of fields left invariant under our discrete symmetry. We'll find that these theories have enhanced symmetry properties at one loop.

The underlying reason lies in the 'orbifold correspondence' in large N gauge theories [284, 285, 286]. Given a 'mother' field theory, you can create a 'daughter' theory by

embedding some given discrete symmetry in the symmetries of the mother theory and projecting out states which are not invariant under that discrete symmetry; we call this process 'orbifolding'. Then at leading order in large N, the correlation functions of the daughter theory match those of the mother theory. This is nothing short of amazing—a theory with no supersymmetry to speak of can nevertheless 'accidentally' exhibit supersymmetric behavior at leading order.

The general case of the orbifold correspondence and how to construct Neutral Naturalness models is beautiful and I do recommend reading [284, 273, 287], but the group theory required for a full discussion would be too large a detour from our main narrative. Fortunately we can get a good sense for what's going on by considering a few explicit examples, which will not require much mathematical machinery.

Example 1: Folded Supersymmetry

Let's first consider the example of 'Folded Supersymmetry' [273] which was the first Neutral Naturalness model constructed explicitly via orbifolding. The idea is precisely to consider a supersymmetric theory and orbifold project onto a theory with no explicit supersymmetry but in which supersymmetric cancellations still occur at one loop. As a pedagogical example, consider an $\mathcal{N} = 1$ supersymmetric $U(2N)_C$ gauge theory with $2N$ flavors of left and right fundamental chiral superfields $Q = (\tilde{q}, q)$ which enjoy a $U(2N)_{F,L} \times U(2N)_{F,R}$ global flavor symmetry. Let's decree also that the theory respects R-symmetry. We will orbifold by the discrete group $\mathbb{Z}_2$ as before, but we must choose how to embed the $\mathbb{Z}_2$ in each of these symmetry groups. That is, our original 'mother' theory is invariant under a large symmetry group $U(2N)_C \times U(2N)_{F,L} \times U(2N)_{F,R} \times U(1)_R$, which contains *many* $\mathbb{Z}_2$ subgroups, and we must decide precisely which $\mathbb{Z}_2$ we want to

orbifold by. We choose the following embeddings:

$$
\mathbf{C} = \begin{pmatrix} \mathbb{1}_N & 0 \\ 0 & -\mathbb{1}_N \end{pmatrix}, \qquad \mathbf{F} = \begin{pmatrix} \mathbb{1}_N & 0 \\ 0 & -\mathbb{1}_N \end{pmatrix}, \qquad \mathbf{R} = (-1)^F, \tag{4.3}
$$

where the first matrix is in color-space, the second is in flavor-space, and the third transformation is by fermion number. Each of these generates a $\mathbb{Z}_2$ subgroup of one of the symmetry factors of the mother theory. To see which fields are invariant under this $\mathbb{Z}_2$, we must only act them on the fields and see how they transform. Gauge bosons A_μ and their superpartner gauginos λ are in the adjoint of the gauge group. Indexing $A, B = 1..N$ separately over the two halves of the gauge indices, we have

$$
A_\mu = \begin{pmatrix} A_{\mu,AA} & A_{\mu,AB} \\ A_{\mu,BA} & A_{\mu,BB} \end{pmatrix} \to \mathbf{C} \begin{pmatrix} A_{\mu,AA} & A_{\mu,AB} \\ A_{\mu,BA} & A_{\mu,BB} \end{pmatrix} \mathbf{C}^\dagger \mathbb{1}_R = \begin{pmatrix} +A_{\mu,AA} & -A_{\mu,AB} \\ -A_{\mu,BA} & +A_{\mu,BB} \end{pmatrix} \tag{4.4}
$$

$$
\lambda = \begin{pmatrix} \lambda_{AA} & \lambda_{AB} \\ \lambda_{BA} & \lambda_{BB} \end{pmatrix} \to \mathbf{C} \begin{pmatrix} \lambda_{AA} & \lambda_{AB} \\ \lambda_{BA} & \lambda_{BB} \end{pmatrix} \mathbf{C}^\dagger (-\mathbb{1})_R = \begin{pmatrix} -\lambda_{AA} & +\lambda_{AB} \\ +\lambda_{BA} & -\lambda_{BB} \end{pmatrix} \tag{4.5}
$$

where the difference here is because of the different R transformations. We see that if we project down to only those states invariant under this transformation, our gauge group dissolves from $U(2N)_C$ down to disconnected pieces $U(N)_C \times U(N)_C$. We see that embedding the $\mathbb{Z}_2$ non-trivially in the R-symmetry group means the daughter theory will be non-supersymmetric. The superpartners of our gauge fields are no longer present, but rather the gauginos have been twisted into bifundamentals under the two gauge factors. In the matter sector, letting $a, b = 1..N$ similarly index the two halves of the flavor indices

135

(a)

(b)
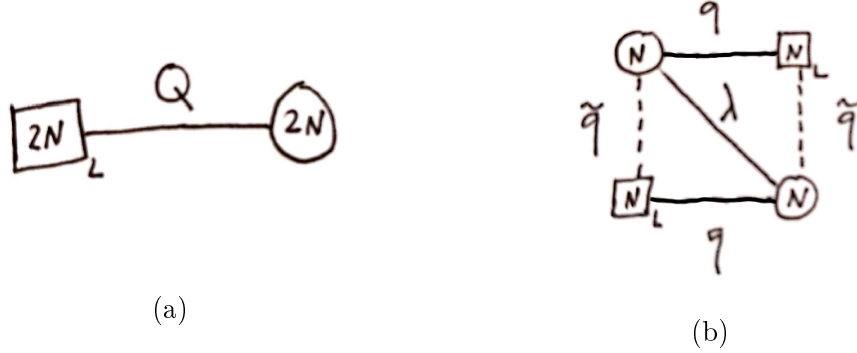
Figure 4.1: In (a), a quiver diagram for the (left-chiral fields in) the mother theory. A circle corresponds to a gauge symmetry and a square to a flavor symmetry. Lines between denote bifundamentals, which here represents a full chiral supermultiplet. In (b), a quiver diagram for the orbifold daughter theory described in the text. The gauge and flavor groups have been divided into two distinct groups each, and the degrees of freedom in the chiral and vector supermultiplet have been shuffled around.

and writing the fields as matrices in a combined color and flavor space, we have

$$
\tilde{q} = \begin{pmatrix} \tilde{q}_{Aa} & \tilde{q}_{Ab} \\ \tilde{q}_{Ba} & \tilde{q}_{Bb} \end{pmatrix} \rightarrow \mathbf{C} \begin{pmatrix} \tilde{q}_{Aa} & \tilde{q}_{Ab} \\ \tilde{q}_{Ba} & \tilde{q}_{Bb} \end{pmatrix} \mathbf{F}^{\dagger} \mathbb{1}_R = \begin{pmatrix} +\tilde{q}_{Aa} & -\tilde{q}_{Ab} \\ -\tilde{q}_{Ba} & +\tilde{q}_{Bb} \end{pmatrix} \tag{4.6}
$$

$$
q = \begin{pmatrix} q_{Aa} & q_{Ab} \\ q_{Ba} & q_{Bb} \end{pmatrix} \rightarrow \mathbf{C} \begin{pmatrix} q_{Aa} & q_{Ab} \\ q_{Ba} & q_{Bb} \end{pmatrix} \mathbf{F}^{\dagger}(-\mathbb{1})_R = \begin{pmatrix} -q_{Aa} & +q_{Ab} \\ +q_{Ba} & -q_{Bb} \end{pmatrix}. \tag{4.7}
$$

Again we see that we have broken supersymmetry. The flavor group has also broken down to $U(N)_{F,L} \times U(N)_{F,L}$ and the invariant states are squarks which are bifundamentals under 'diagonal' combinations of the gauge and flavor groups, and quarks which are bifundamentals under the 'off-diagonal' combinations.

It is not too hard to roughly see the magic of how the orbifold-projected theory continues to protect scalar masses. Draw the one-loop the diagrams in the mother theory which would contribute to a calculation of the mass of, say, $\tilde{q}_{Aa}$, as in Figure 4.2. In the mother theory we know there are no quadratic corrections by supersymmetry. In the
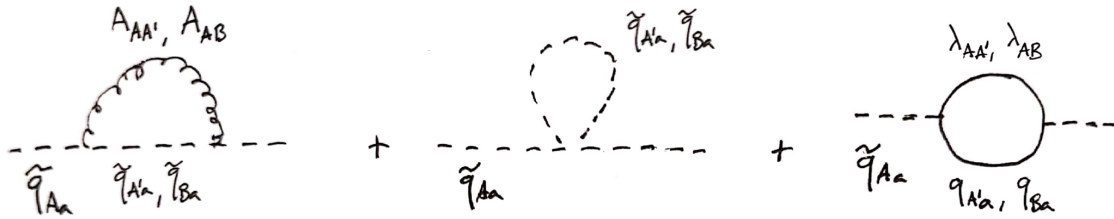
Figure 4.2: Representative diagrams contribution to the one-loop mass correction of a scalar in the supersymmetric mother theory. The daughter theory projects out some of the internal states of each diagram in a pattern that is non-supersymmetric but still enforces cancellations.

daughter theory, half of each sort of diagram will be eliminated by the orbifold projection, so it's clear that there will still be no quadratic divergences. But because it is different internal states that have been eliminated for different classes of diagrams, the daughter theory has no supersymmetry to speak of! Again, as we saw in the example of the twin Higgs, the magic is in that at one-loop we really only need to get the counting right, and so we can use orbifolding to construct theories which do that in clever ways.

The structure of the theory can be succinctly summarized in a 'quiver' or 'moose' diagram where the various symmetry groups correspond to nodes in a graph and the matter is represented by links between these groups which represent their charges, as in Figure 4.1.

Despite the fact that our daughter theory has *no* supersymmetry, the orbifold correspondence guarantees that in the $N \to \infty$ limit the correlation functions have full supersymmetric protection. For finite N the supersymmetric relations are broken, but only by $1/N$ corrections. Going through this explicitly is useful, and pedagogical discussions of it can be found in [284, 276].

Example 2: The Γ-plet Higgs

We can understand the twin Higgs as the simplest example of a Γ-siblings Higgs which is an orbifold projection $SU(3\Gamma) \times SU(2\Gamma)/\mathbb{Z}_\Gamma$. We focus on the sector which contains the Yukawa interaction of the top quark, as the top partners have the most effect on the naturalness of the Higgs. In the mother theory the matter content is:

	$SU(3\Gamma)$	$SU(2\Gamma)$	$SU(\Gamma)$
H	-	$\square$	$\overline{\square}$
Q	$\square$	$\overline{\square}$	-
u	$\overline{\square}$	-	$\overline{\square}$

where the charges allow the operator $y_t H Q u$. The generalization of the twin Higgs is found by embedding the $\mathbb{Z}_\Gamma$ in these groups as the block-diagonal element $\mathbf{S}_n = \text{diag}\left[\mathbb{1}_n, \mathbb{1}_n \exp(2\pi i/\Gamma), \ldots, \mathbb{1}_n \exp(2\pi k i/\Gamma), \ldots, \mathbb{1}_n \exp(2\pi(\Gamma-1)i/\Gamma)\right]$, where $n = 1, 2, 3$ corresponds to which $SU(n\Gamma)$ factor the element belongs to and $\mathbb{1}_n$ is the $n \times n$ identity matrix. This leads to Γ SM-like sectors which are independent except for their scalar sectors, where the Γ-plet Higgses enjoy an accidental $SU(2\Gamma)$ symmetry at one loop.

The general orbifolding approach to constructing models of Neutral Naturalness was fully laid out in [287, 276], where they in particular explore a 'regular representation' embedding of the discrete group into the continuous symmetries of the mother theory.

This approach of orbifolding to find low-energy models with accidental symmetries is useful also because such models come along with guides for how to UV-complete them. When we wrote down the twin Higgs model above, it was perhaps not obvious that there is a nice UV completion of this theory. But now we see that the twin Higgs is an orbifold projection of a $SU(6) \times SU(4)$ gauge theory by $\mathbb{Z}_2$, so we expect we can uplift this to a five-dimensional UV completion where the twin Higgs emerges *dynamically* at low energies from orbifold boundary conditions. So we can confidently study solely the low-

energy effective theory of the zero-modes we've projected out without worrying explicitly about whether a UV completion exists.

4.2   The Loerarchy Problem

Modern particle theorists must now confront a new version of the issue of electroweak naturalness. When originally understood, the pressing problem was understanding what sorts of UV structure could protect a scalar from large mass corrections. Of course this structure needed to be broken to get the SM structure in the IR and there's lots of interesting physics and subtleties on that end as well. But with the structure of the weak scale barely explored, the ways in which this could be done were abundant.

Over the ensuing decades we explored electroweak physics with increasing precision, which has provided invaluable guidance for how the SM structure must appear. Gradually the IR dynamics became more and more constrained to the point where now, as we have emphasized, we have enormous constraints on any appearance of new physics with SM quantum numbers up to mass scales that are often many times the electroweak scale.

We thus have a qualitative change in the electroweak naturalness issue over the past decades. We term the modern, empirical, low-energy puzzle of electroweak naturalness without visible structure around the weak scale as 'The Loerarchy Problem', for obvious reasons.

In this language, the little hierarchy problem is just one approach toward this problem, which assumes that one of the classic solutions is just out of reach and another module is needed to postpone the appearance of SM charged particles. While it's more than worthwhile to continue looking for and exploring those theories, in the face of increasingly powerful LHC data in excellent agreement with the Standard Model it's worth thinking transversely. As intriguing as the Neutral Naturalness models are, these classes of models

rely on one-loop accidents and so do not completely relieve the empirical pressure from a lack of new physics at the energy frontier. Such models are gradually also being constrained by the LHC, so what else are we to do? The *maximalist* interpretation of the LHC data is that Nature may be leading us to the conclusion that *there is no new physics at the weak scale.* With every inverse femtobarn of LHC data without a signal of new physics, the impetus for such a paradigm shift becomes stronger.

But how can we generate a scale *without* additional structure appearing surrounding it? Within the context of effective field theory, decoupling theorems demand that if the RG evolution is to change around a scale M it must be due to fields close to the scale M. This was perhaps clearest in dim reg with $\overline{\text{MS}}$, where the beta functions explicitly change solely at mass thresholds, though the Wilsonian approach is more useful for physical intuition. So how are we to break this feature? This undertaking is the maximalist approach to the Loerarchy Problem.

4.3   Violations of Effective Field Theory

> There are more things in heaven and earth, Horatio,
> Than are dreamt of in your philosophy.

William Shakespeare

Hamlet, c. 1600 [288]

The line of thought we suggest here is that perhaps the apparent violation of EFT expectations at the weak scale is a sign of the breakdown of EFT itself. Depending on how much background in particle physics one has this statement may seem more or less heretical, but the idea is not as radical as it may at first seem—for one reason, the cosmological constant problem has inspired sporadic reexaminations of the validity of

effective field theory in our universe for decades.

The cosmological constant problem is the fine-tuning issue with the *other* dimension-ful parameter in the Standard Model. Just as with the Higgs mass, there is no protective symmetry in the Standard Model for the vacuum energy, and so the natural expectation is $\Lambda \sim M_{\rm pl}^4$, some 120 orders of magnitude higher than observations suggest. However there's an important difference in the severity of these problems—for the Higgs, as emphasized in Section 2.2.1, the worrisome mass corrections are those from new physics. This is why the severity of the problem has only ratcheted up in recent years, as we have seen nothing to protect the Higgs from BSM mass corrections. But for the vacuum energy, there are finite, calculable, physical contributions in the Standard Model itself! For a start, EWSB by the Higgs yields a contribution $\sim -v^4$, and chiral symmetry-breaking yields a contribution $\sim -\Lambda_{QCD}^4$. How is it that these can be nearly perfectly canceled off in the late universe?

There have been important attempts to address the cosmological constant problem with a violation of effective field theory, from Coleman's suggestion [289] that nonlocality induced by wormholes may allow the early universe to be sensitive to late-time requirements to Cohen-Kaplan-Nelson's suggestion [290] that the Bekenstein bound demands an infrared cutoff on the validity of any EFT. From one perspective, our suggestion to extend this philosophy to the hierarchy problem appears natural in light of its apparent need in cosmology. We can point to an even-more-general motivation with the realization that gravity necessarily violates EFT.

4.3.1   Gravity and EFT

The perturbative quantum field theory of the Einstein-Hilbert Lagrangian is quite clearly an effective field theory of a symmetric two-index tensor field which obeys dif-

feomorphism invariance and with power counting in $1/M_{\text{pl}}$ [291], as we can easily see by writing it out and expanding around flat space $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$:

$$-\mathcal{L}_{EH} = \frac{1}{2M_{\text{pl}}^2}\sqrt{g}R \sim \partial^2 h^2 + \frac{1}{M_{\text{pl}}}\partial^2 h^3 + \frac{1}{M_{\text{pl}}^2}\partial^2 h^4 + \dots \tag{4.8}$$

where we have only given the schematic form of the operators in terms of the number of derivatives and linearized gravitational fields they contain, as the full expressions quickly become complicated [292, 293]. This effective field theory will be a good approximation to infrared gravitational physics until we get to energies close to the Planck scale, at which point the higher-dimensional operators are unsuppressed and we need a UV completion. From this bottom-up approach it's not clear why gravity should be particularly special. But we can get some insight at a very basic level by thinking about gravitational scattering.

Let's compare two effective field theories: the four-Fermi theory of the weak interaction below the weak scale $G_F^{-1/2}$ and the perturbative theory of quantum gravity below the Planck mass. If we consider scattering two leptons at $\sqrt{s} \ll G_F^{-1/2}$, we can make very precise predictions by computing in the four-Fermi theory—the possible final states, the differential cross-section; whatever we'd like. And similarly if we scatter two particles at $\sqrt{s} \ll M_{\text{pl}}$, we can compute to high precision in quantum gravitational corrections what will happen.

Now inversely, imagine scattering two leptons at $\sqrt{s} \gg G_F^{-1/2}$ in the four-Fermi theory. At a scale like $10^{15}G_F^{-1/2}$, the EFT has obviously broken down and we can say essentially nothing about what this process will look like—any calculation we tried would be hopelessly divergent, and we have no idea what sorts of states might exist at energies that large.

However, what if we scatter two particles at $\sqrt{s} \gg M_{\text{pl}}$, say a ridiculously trans-

Planckian scale like $10^{15} M_{\text{pl}}$? In this case, in fact, we know what will happen to incredible precision—a black hole will form! This will be a 'big' black hole of mass $\gg M_{\text{pl}}$ with a lifetime of order days. It will be well-described by classical general relativity for some macroscopic time, and then semiclassical GR for an $\mathcal{O}(1)$ fraction of the full lifetime. In fact, unless you are interested in *incredibly* detailed measurements which involve collecting essentially every emitted Hawking quanta (of which there will be $n \sim 10^{30}$ in this case!) and finding their entanglement structure, we know how to describe the evaporation nearly completely.[20]

So something profoundly weird is going on. The key point being that in gravity, the far UV of the theory is controlled by *classical, infrared* physics. This is obviously a feature that we do not see in other EFTs.

This is not a new idea; it has long been known that gravity contains low-energy effects which cannot be understood in the context of EFT. The fact that black holes radiate at temperatures inversely proportional to their masses [295] necessitates some sort of 'UV/IR mixing' in gravity—infrared physics must somehow 'know about' heavy mass scales in violation of a naïve application of decoupling. As a perhaps-more-fundamental *raison d'être* for such behavior, the demand that observables in a theory of quantum gravity must be gauge-(that is, diffeomorphism-)invariant dictates that they must be nonlocal (see e.g. [296, 297, 298, 299, 300]), again a feature which standard EFT techniques do not encapsulate. In view of this, the conventional position is that EFT should remain a valid strategy up to the Planck scale, at which quantum gravitational effects become important. But once locality and decoupling have been given up, how and why are violations of EFT expectations to be sequestered to inaccessible energies? Indeed, the 'firewall' argument [301] evinces tension with EFT expectations in semiclassical gravity

---

[20]I recommend Giddings' Erice lectures [294] for more on the perspective of quantum gravitational behavior as a function of the Mandelstam variables.

around black hole backgrounds at arbitrarily low energies and curvatures, as does recent progress finding the Page curve from semiclassical gravity [302, 303, 304, 305, 306, 307].

That quantum gravitational effects will affect infrared particle physics is likewise not a new idea. This has been the core message of the Swampland program [308], which has been cataloging—to varying degrees of concreteness and certainty—ways in which otherwise allowable EFTs may conjecturally be ruled out by quantum gravitational considerations. These are EFTs which would look perfectly sensible and consistent to an infrared effective field theorist, yet the demand that they be UV-completed to theories which include Einstein gravity reveals a secret inconsistency. While this is powerful information, the extent to which the UV here meddles with the IR is relatively minor— just dictating where one must live in the space of infrared theories. Even so, they have been found to have possible applications to SM puzzles, including the hierarchy problem [91, 92, 93, 94, 95, 96, 97, 98, 11, 5, 99].

Let us review briefly the approach to connect the Weak Gravity Conjecture (WGC) [309] to the hierarchy problem. The WGC is one of the earliest and most well-tested Swampland conjectures, its formulation is relatively easy to understand, and it's emblematic of the way one might try to connect the hierarchy problem to Swampland conjectures in general.

The prime motivation for formulating the WGC was the well-known folklore that quantum gravity does not respect global symmetries. The simple argument for this fact is that 'black holes have no hair'—the only quantum numbers a black hole has correspond to its mass, spin, and gauge charges. This means that if we make a black hole by smashing together a bunch of neutrons, there is no reason why it cannot decay into solely photons, which violates the global $U(1)_{B-L}$ symmetry of the SM but no gauge symmetries. The authors suggested that the fact that Abelian gauge symmetries smoothly become global symmetries as the gauge coupling vanishes means that something must go wrong with

very small gauge couplings as well, so as for the physics to be smooth in this limit.

Another line of thinking for quantum gravity not respecting global symmetries, which connects more closely to the WGC formulation, is based on entropic arguments. If global symmetries could be exact, then we could create a big black hole with an arbitrary global charge, and wait a very very long time while it Hawking evaporates down to the Planck scale. Unlike gauge charges, the global charge does not affect the metric, so the black hole does not shed this charge as it evaporates. We then end up with a Planck-sized black hole with an arbitrarily large global charge, which would necessitate the existence of arbitrarily-many black hole microstates for a fixed-mass black hole. But this is a disaster! In calculating a scattering amplitude one has to sum over all possible intermediate states, in principle including black holes. These effects are surely Boltzmann-suppressed by enormous amounts, but if there are an *infinite* number of possible black holes then any nonzero contribution from a single black hole will lead to a divergence. A cuter (albeit somewhat tongue-in-cheek) 'hand waving argument' is provided by [310]: Were there incredibly large numbers of super-Planckian states which could populate a thermal bath, a vigorous wave of your hand would produce them in Unruh radiation, and you could feel them against your hand.

Now if we have a gauge symmetry, there are no longer Planck-sized black holes with arbitrarily large charge because there is an extremality bound: $g|Q| \leq M/M_{pl}$, with Q the gauge charge and M the mass, in units of the gauge coupling g and the Planck mass. If this were violated, a naked singularity would appear. But for a very tiny gauge coupling, while we no longer have strictly infinitely-many black holes at each mass, there are still an *enormous* number of black hole microstates for a Planck-sized black hole as $g \to 0$. So a worry like the hand-waving argument still applies.

This sort of reasoning motivated Arkani-Hamed, Motl, Nicolis, and Vafa to conjecture that just as quantum gravitational theories must not have exact global symmetries, they

must also suffer some physical malady which disallows the limit where Abelian gauge symmetries become global symmetries. Their conjecture has two forms: The 'magnetic' conjecture dictates that a quantum gravitational theory with a $U(1)$ gauge symmetry must be enlarged into a theory allowing magnetic monopoles by a cutoff $\Lambda \lesssim g M_{pl}$, which connects to such a description never being valid in the limit $g \to 0$. The 'electric' form of the WGC is that such a theory must contain a particle which is 'super-extremal'— it has charge greater than its mass $gqM_{pl} > m$. The existence of such a particle would destabilize the extremal, charged black holes, allowing them to decay.

Well this super-extremality bound should look very interesting to us, as it provides an upper bound on a mass scale. While we cannot apply this directly to the Higgs because it is not charged under any unbroken Abelian gauge symmetries, we know that one of the Higgs' jobs is to provide mass to other particles. So if the weak gravity conjecture bound must apply to some state ϕ with mass $m_\phi = yv$, with v the Higgs vev, then this still amounts to a bound on the electroweak scale.

This was suggested in the context of gauged $U(1)_{B-L}$ with very tiny coupling giving an upper bound on the lightest neutrino mass [91], but the magnetic form of the WGC is difficult to deal with in this context. This can be circumvented by introducing a new dark Abelian gauge group $U(1)_X$ and charged states which get (some of) their mass from the Higgs [5].

The way such a model solves the hierarchy problem is by changing the shape of our prior for the electroweak scale, as mentioned in Section 2.2.2. As naïve effective field theorists with no information about quantum gravity, we assumed some sort of flat prior in $[-M_{pl}, M_{pl}]$. But in fact much of this space is ruled out by quantum gravitational constraints, consisting of theories 'in the Swampland', meaning that our prior should be reshaped to include only values which can actually be produced by a theory of quantum gravity. This is how much of the connection from Swampland conjectures to the real

world has worked schematically: one says that Quantum Gravity demands one live in a subregion of the EFT parameter space.

In theory far more flagrant violations of low-energy expectations are permissible—that is, the extent to which quantum gravitational violation of EFT will affect the infrared of our universe is not at all certain. Of course any proposal to see new effects from a breakdown of EFT must contend with the rampant success of the SM EFT in the IR. Certainly a violation of EFT must both come with good reason and be deftly organized to spoil only those observed EFT puzzles. For the former, the need for quantum gravity is obviously compelling. As to the latter, it is interesting to note that the most pressing mysteries involve the relevant parameters in the SM Lagrangian.

Ultimately, our ability to address the hierarchy problem through quantum gravitational violations of EFT is limited by our incomplete understanding of quantum gravity. This motivates finding non-gravitational toy models that violate EFT expectations on their own, providing a calculable playground in which to better understand the potential consequences of UV/IR mixing. In Chapter 7 we pursue the idea that UV/IR mixing may have more direct effects on the SM by considering noncommutative field theory (NCFT) as such a toy model. These theories model physics on spaces where translations do not commute [311, 312], and have many features amenable to a quantum gravitational interpretation—indeed, noncommutative geometries have been found arising in various limits of string theory [313, 314, 315, 316].

Chapter 5

Neutral Naturalness in the Sky

> In the beginning the Universe was created.
> This has made a lot of people very angry and
> been widely regarded as a bad move.

> Douglas Adams
> *The Restaurant at the End of the Universe*
> (1980) [317]

5.1 Particle Cosmology

It is an amazing and serendipitous fact that the universe started off hot. As a result of the initially high energies and densities, the details of microscopic physics greatly affected the large-scale evolution of the universe. Since the speed of light is finite, by looking out in the sky at enormous distances we can not only learn about the history of the universe but we can use this information to learn about particle physics. While cosmology doesn't give us probes of arbitrarily high temperatures, there's still a humongous amount to be learned—in part due to further serendipity. The fact that the universe transitions from

radiation domination to matter domination shortly before recombination means that the cosmic microwave background (CMB) encodes information both about the light, radiation-like degrees of freedom as well as the matter density in the early universe. Had radiation domination ended far before recombination, it would be far more difficult to use the CMB to constrain light degrees of freedom like extra neutrinos. Had radiation domination ended far after recombination, there would be little evidence of dark matter in the CMB, which is the strongest evidence for particle dark matter instead of, say, a modification of gravity at large distances. In fact such a cosmic coincidence also occurs much later, as there is a very long epoch of matter domination before the universe transitions to dark energy domination. Were there just slightly less dark energy, its effects would be essentially invisible thus far in the history of the universe, and it would be very difficult to measure dark energy at all.

All that is to say that there is enormous value in collaboration between particle physics and cosmology. In this chapter we investigate this connection for the twin Higgs model in particular, though our findings are relevant for general Neutral Naturalness theories as well. In Section 4.1.1 we noted that the energy frontier does not effectively probe these theories. Since they do not introduce new particles with Standard Model charges, it is only precision electroweak measurements made at colliders that constrain them at all. However, such theories are in fact probed very well by cosmology, as they introduce new light degrees of freedom. Despite the fact that these do not directly interact with normal matter, their gravitational effects still contribute to the evolution of the universe, and so the CMB provides a powerful constraint on new light particles.

It is this cosmological effect that provided the biggest obstacle to the original twin Higgs proposal [267], which became an urgent issue after the null results of run 1 of the LHC and the increased interest in models where the lightest states responsible for Higgs naturalness were SM-neutral. The landmark approach taken in [268] was to pare

the model down to a 'minimal' version where only those states necessary for Higgs naturalness appeared in the twin spectrum. This revived the twin Higgs as a solution to the little hierarchy problem, and their 'fraternal' version brought about many interesting phenomenological possibilities.

The 'fraternal twin Higgs' has a twin sector consisting—at energies below its cutoff—solely of the third generation of fermions, and with ungauged twin hypercharge. This brilliantly removes all light particles from the spectrum, so their effects would not cause trouble in the early universe. But this approach leaves perhaps a niggling unpleasant taste for those worried about parsimony. Yes the fraternal twin Higgs introduces fewer new particles than the mirror twin Higgs, so a naïve desire to solve problems with few ingredients might suggest that this is a windfall. However, the mirror twin Higgs really consists of only two 'ingredients': a $\mathbb{Z}_2$ symmetry and some soft breaking to misalign the resulting Higgs vevs—whereas the fraternal twin Higgs has much more structure.

While those statements seem to be of a very subjective sort, we can ground this unease in physics by considering what's needed to UV-complete such a model. At the cutoff of this model $\Lambda \sim 10$ TeV, we need the $\mathbb{Z}_2$ to still be a relatively good symmetry among the largest couplings in the twin sector—that is, the gauge couplings g_2, g_3 and the top Yukawa y_t—such that the cancellation of contributions from the two sectors to the Higgs potential works well. Yet in other parts of the theory we have done great violence to the structure of the theory, having broken twin hypercharge and removed parts of the spectrum. How are we to ensure firstly that the correct degrees of freedom gain large masses, and secondly that this does not radiatively feed into the remaining light degrees of freedom?

In the midst of this digression, I should mention a terminological confusion. In the literature, the phrase 'mirror twin Higgs' often refers to models in which the full collection of twin degrees of freedom are present in the low-energy theory, regardless of how much

$\mathbb{Z}_2$-breaking is present. Judicious introduction of such asymmetries has been used to create models which avoid cosmological issues by making all the twin fermions heavy, while still staying within the technical definition of the mirror twin Higgs. But this is an overreliance on a definition; the fraternal twin Higgs is merely a limit of these theories in which the $\mathbb{Z}_2$-breaking is severe enough to push some degrees of freedom above the cutoff. The real distinction between classes of Neutral Naturalness models should be between those which break the $\mathbb{Z}_2$ only minimally and those which do greater violence to the symmetry. It is this distinction which classifies the difficulty involved in finding a UV completion.

Now let me emphasize that this is not to undercut the value of such a model. After all, the Yukawa interactions in the Standard Model badly break the large global symmetries it would otherwise have. Indeed, the fraternal twin Higgs showcased interesting phenomena, pointed to new general experimental probes, and provided a basis for many intriguing lines of research. In fact we will return to this model in Chapter 6 to study a novel collider search strategy to which it lent credence and which turns out to be a broadly useful probe of many theories of BSM physics. Despite the fact that the vast, vast majority of theory papers written in particle physics will not ultimately be the exact right model of the universe, they still contain value. They may guide experimental searches toward interesting classes of signals to look for, or teach us new things about the range of particle phenomenology or quantum field theories. Regardless, we obviously don't know in advance which model will be correct, so exploring all possible directions is crucial.

Yet when there *is* the possibility for a more parsimonious model, it's certainly worth pursuing that option. This is the philosophy that led to my collaborators and me looking into the prospect of attaining a realistic twin Higgs cosmology that respected the $\mathbb{Z}_2$ symmetry.

5.2   Asymmetric Reheating

5.2.1   Introduction

The primary challenge to the mirror Twin Higgs comes not from LHC data, but from cosmology. An exact $\mathbb{Z}_2$ exchange symmetry predicts mirror copies of light Standard Model neutrinos and photons states, which contribute to the energy density of the early universe. In particular, twin neutrinos and a twin photon provide a new source of dark radiation that is strongly constrained by CMB and BBN measurements [318, 319]. While these constraints could be avoided if the two sectors were at radically different temperatures, the Higgs portal couplings required by naturalness keep the two sectors in thermal equilibrium down to relatively low temperatures.

Constraints on dark radiation in the mirror Twin Higgs have motivated models in which the $\mathbb{Z}_2$ symmetry is approximate [276, 287, 320, 321, 322, 268, 280, 323, 324, 325, 326], in which case the dark radiation component can be made naturally small. These models have proved to be a boon for phenomenology. Among other things, they quite generally motivate looking for Higgs decays to long-lived particles at colliders [327, 328, 329, 330, 12, 331, 10, 332] and contain well motivated dark matter (DM) candidates [333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343, 344]. However, such cosmological fixes come at the cost of minimality, as models with approximate $\mathbb{Z}_2$ symmetries require a considerable amount of additional structure near the TeV scale.

In this work we take an alternative approach and investigate ways in which early universe cosmology can reconcile the mirror Twin Higgs with current CMB and BBN observations. In doing so, we find compelling scenarios that transfer the signatures of electroweak naturalness from high-energy colliders to cosmology. We consider several possibilities in which the energy density of the light particles in the twin sector is diluted by the out-of-equilibrium decay of a new particle after the two sectors have thermally

decoupled. Crucially, the new physics in the early universe respects the exact (albeit spontaneously broken) $\mathbb{Z}_2$ exchange symmetry of the mirror Twin Higgs. This symmetry may be used to classify representations of the particle responsible for this dilution. We concentrate on two minimal cases: In the first, the long-lived particle is $\mathbb{Z}_2$-even and the asymmetry is naturally induced by kinematics. In the second, there is a pair of particles which are exchanged by the $\mathbb{Z}_2$ symmetry and which may be responsible for inflation.[21] Moreover, in these cases the new physics does not merely reconcile the existence of a mirror twin sector with cosmological constraints, but predicts contributions to cosmological observables that may be probed in current and future CMB experiments. This raises the prospect of discovering evidence of electroweak naturalness first through cosmology, rather than colliders, and provides natural targets for future cosmological constraints on minimal realizations of neutral naturalness.

The next sections are organized as follows: In Section 5.2.2 we discuss the thermal history of the mirror Twin Higgs, with a particular attention to the interactions keeping the Standard Model and twin sector in thermal equilibrium and the cosmological constraints on light degrees of freedom. In Section 5.2.3 we present a simple model where the out-of-equilibrium decay of a particle with symmetric couplings to the Standard Model and twin sector leads to a temperature difference between the two sectors after they decouple. We turn to inflation in Section 5.2.4, constructing a model of "twinflation" in which the softly broken $\mathbb{Z}_2$-symmetry extends to the inflationary sector and leads to two periods of inflation. The first primarily reheats the twin sector, while the second primarily reheats the Standard Model sector. We conclude in Section 5.2.5.

---

[21]A third case exists, in which the particle is $\mathbb{Z}_2$-odd. This may additionally be related to the spontaneous $\mathbb{Z}_2$-breaking in the Higgs potential, although we find that a realisation of such a scenario is dependent upon the UV completion of the model.

5.2.2   Thermal History of the Mirror Twin

The primary challenge to the mirror Twin Higgs comes from cosmology, rather than collider physics. The mirror Twin contains not only states responsible for protecting the Higgs against radiative corrections (such as the twin top), but also a plethora of extra states due to the $\mathbb{Z}_2$ symmetry that are irrelevant to naturalness. The lightest of these, namely the twin photon and twin neutrinos, contribute significantly to the energy density of the early universe around the era of matter-radiation equality, since they have a temperature comparable to that of the Standard Model plasma at all times. This is because the same Higgs portal coupling that makes the Higgs natural also keeps the two sectors in thermal equilibrium down to $\mathcal{O}(\mathrm{GeV})$ temperatures. Then the identical particle content in the twin and Standard Model sectors guarantees that they remain at comparable temperatures even after they decouple - for every massive Standard Model species that becomes non-relativistic and transfers its entropy to the rest of the plasma, its twin counterpart does the same within a factor of f/v in temperature.

In this section we undertake a detailed study of the decoupling between the Standard Model and twin sectors as well as the constraints from precision cosmology.

Twin Degrees of Freedom

In thermal equilibrium, each relativistic degree of freedom has roughly the same energy density. In general, we express the energy density of the universe ρ during the radiation-dominated era as $\rho \equiv g_\star \frac{\pi^2}{30} T^4$, where we define $g_\star$ through this relation as the effective number of relativistic degrees of freedom and T the temperature of the SM photons. This then determines the evolution of the scale factor through the first

Friedmann equation

$$H = \frac{1}{M_{\text{pl}}} \left[\frac{\pi^2}{90} g_\star T^4 \right]^{1/2} \tag{5.1}$$

(assuming spatial flatness), where M_{pl} is the reduced Planck mass. In general, the energy density of a particular species i may be computed from $\rho_i = g_i \int \frac{d^3 p}{(2\pi)^3} f_i(p, T_i) E(p)$, where g_i are the number of internal degrees of freedom, $E(p)$ is the energy as a function of momentum p, while $f_i(p, T_i)$ is the phase-space number density and is a Bose-Einstein or Fermi-Dirac distribution if the species is in equilibrium at temperature T_i. The number of effective relativistic degrees of freedom may then be defined for each sector separately as $g_\star^{\text{SM}}(T)$ and $g_\star^{\text{t}}(T)$ satisfying $\rho_{\text{SM}}(T) = \frac{\pi^2}{30} g_\star^{\text{SM}}(T) T^4$ and $\rho_{\text{t}}(T) = \frac{\pi^2}{30} g_\star^{\text{t}}(T) T^4$, respectively, where $\rho_{\text{SM}}(T)$ and $\rho_{\text{t}}(T)$ are the total energy densities of SM and twin particles. The values of $g_\star(T)$ for the SM and twin sectors are shown in Figure 5.1, where all species within each sector are in thermal equilibrium. These can then be used to calculate the total number $g_\star$ as a function of temperature, by weighting twin sector energy density by its temperature: $g_\star(T) = g_\star^{\text{SM}}(T) + g_\star^{\text{t}}(\hat{T})(\hat{T}/T)^4$, where $\hat{T}$ is the twin sector photon temperature when the SM photon temperature is T.

Likewise, entropy densities for each sector i are defined as $s_i(T) = \frac{2\pi^2}{45} g_\star^{\text{i}}(T) T^3$. We neglect the small differences between the number of relativistic degrees of freedom defined from energy and entropy densities, which are not significant over the range of temperatures of interest here.

Decoupling

In the early universe, the two sectors are thermally linked by interactions mediated by the Higgs, which, through mixing with both h_A and h_B components, allows for SM fermions and weak bosons to scatter off or annihilate into their twin counterparts.
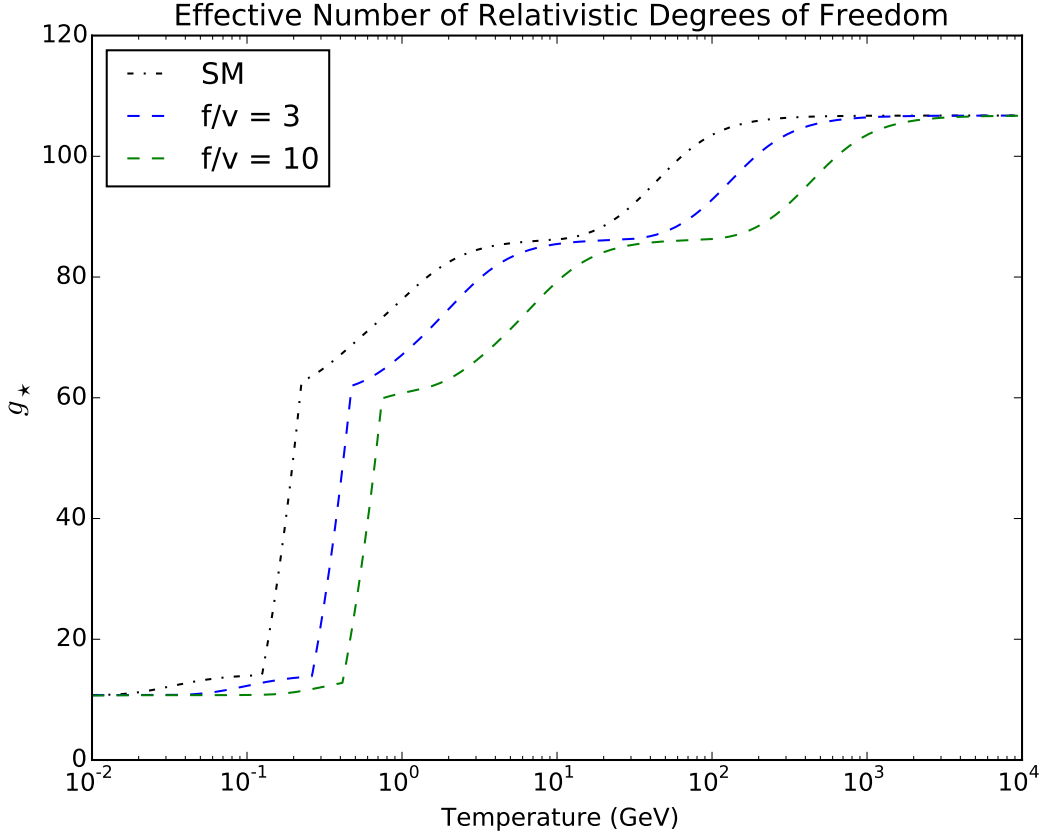
Figure 5.1: The effective number of relativistic degrees of freedom for mirror Twin Higgs models for different values of f/v. The dash-dotted line is the for the Standard Model $g_\star^{\mathrm{SM}}(T)$ and the dashed lines are the twin sector degrees of freedom $g_\star^{\mathrm{t}}(T)$. The evolution of $g_\star$ during the QCD phase transition (QCDPT) is not well-understood, so we assign the SM QCDPT a central value of 175 MeV and a width of 50 MeV and interpolate linearly between the values of $g_\star$ at 225 MeV for free partons and at 125 MeV for pions. Further discussion may be found in [345]. For the twin sector we use a central value and width which are $(1 + \log(\frac{f}{v}))$ times larger than the SM values. Note that new mass thresholds, expected to appear at energies ~ 10 TeV in UV completions of the twin Higgs, have not been included.

However, once the temperature drops sufficiently for this Higgs-mediated interaction to become rare on the expansion time-scale, the sectors decouple and thereafter thermally evolve independently. More precisely, thermal decoupling will occur once the rate at which energy can be exchanged between SM and twin particles (through the Higgs) falls

below the Hubble rate.

Thermal decoupling is traditionally formulated from the Boltzmann equations describing the evolution of single-particle phase space number densities, wherein collisions induce instantaneous changes to the shape of these distributions. When the collisions occur faster than the expansion rate, the phase space probability density functions of the interacting species are expected to relax to an equilibrium distribution (Boltzmann, neglecting quantum statistics, will be applicable to our case). However, once the rate of collisions falls below the expansion rate, collisions become rare on cosmological time scales and the phase space distributions depart from equilibrium. The decoupling temperature is determined as that at which the scattering rate of a participating particle, Γ, drops below the Hubble rate, assuming that this occurs instantaneously across the entire phase space where the number density is significant. This formulation can be used to determine the time at which a particular species of particle will cease to scatter off twin particles on cosmological time scales.

In the case of interest here, however, both sectors of particles remain thermalised within themselves while the interactions between sectors freeze-out. This implies that the phase space number densities are still Boltzmann distributions throughout decoupling, with a different temperature for each sector. As it is the twin sector temperature that ultimately determines the impact of the light twin degrees of freedom on the cosmological observables (discussed below in Section 5.2.2), we wish to describe the thermal evolution of the two sectors by that of their entire energy or entropy content and the bulk heat flows between them. They may then be identified as thermally decoupled once the rate at which they exchange energy falls below the expansion rate.

If the SM and twin sector plasmas have temperatures T and $\hat{T}$ respectively, then calling q the net heat flow density from the SM to the twin sector, the rate at which the

twin entropy densities s_t and s_SM evolve is determined by

$$\frac{ds_\mathrm{t}}{dt} + 3H s_\mathrm{t} = \frac{1}{\hat{T}}\frac{dq}{dt} = \frac{1}{\hat{T}}\left(\frac{dq_\mathrm{in}}{dt} - \frac{dq_\mathrm{out}}{dt}\right) \tag{5.2}$$

$$\frac{ds_\mathrm{SM}}{dt} + 3H s_\mathrm{SM} = \frac{-1}{T}\frac{dq}{dt} = -\frac{1}{T}\left(\frac{dq_\mathrm{in}}{dt} - \frac{dq_\mathrm{out}}{dt}\right). \tag{5.3}$$

Here, H is the Hubble rate. The heat flow rate has been decomposed into the sum of the energy transferred into and out of the twin sector by collisions in the second equality in each line, where $\frac{dq_\mathrm{in}}{dt}$ and $\frac{dq_\mathrm{out}}{dt}$ are both positive.

The rate of heat flow q may be calculated by performing a phase space average of the rate that energy is transferred from the SM to the twin sector through particle interactions. Since the decay rates of top quarks or weak bosons are fast compared to their scattering rate and the Hubble rate, energy transferred to them is instantaneously transferred to the rest of the plasma. Similarly, the scattering rate of lighter fermions off other particles of the same sector (such as photons or gluons) is much faster than their interaction rate with twin fermions. Energy transferred to the lighter fermions therefore quickly diffuses throughout their respective plasmas. The rate of heat flow between sectors may therefore be well approximated by the rate at which energy is transferred from SM particles to twin particles in Higgs mediated interactions. This may occur through elastic scattering of SM particles off twin particles or annihilations of SM particle/antiparticle pairs into twin particles (or the reverse). The energy density transferred to twin particle i from SM particle j in scattering is given by

$$\frac{dq_{ij\to ij}}{dt} = \frac{g_i g_j}{(2\pi)^6}\int\int \frac{d^3k}{2E_i(k)}\frac{d^3h}{2E_j(h)}f_i(k,\hat{T})f_j(h,T)$$
$$4E_i(k)E_j(h)\int v_{rel}(E_i(p) - E_i(k))\frac{d\sigma_{ij\to ij}}{d\Omega}d\Omega, \tag{5.4}$$

where p is the outgoing 4-momentum of particle i. In the cosmic comoving frame, the

phase space number densities f_i and f_j are just Boltzmann factors, although evaluated at the different temperatures of each sector. The factor g_i is the number of internal degrees of freedom of particle i, which here includes colour (the cross section should not be colour averaged, as each colour of quark is present in the plasma in equal abundances and each mediates the exchange of energy, so have their contributions summed). Finally, $E_i(k)$ is the on-shell energy of particle i with momentum k, while $\frac{d\sigma_{ij \to ij}}{d\Omega}$ is the differential scattering cross section for species i scattering off j per solid angle Ω and v_{rel} is the usual relative speed of the incoming particles. As described in [?], the factor in the integrand giving the energy transferred per reaction is simply a component of a 4-vector,

$$X = 4E_i(k)E_j(h) \int (p - k)v_{rel} \frac{d\sigma_{ij \to ij}}{d\Omega} d\Omega. \tag{5.5}$$

This may be calculated in the centre-of-mass frame and then boosted back into the cosmic comoving frame where the integrals in (5.4) can be evaluated, similarly to the thermal averaging procedure described in [346].

The integral (5.4) may be decomposed into two terms giving the positive and negative energy changes of the twin particle, which respectively contribute to $\frac{dq_{\mathrm{in}}}{dt}$ and $\frac{dq_{\mathrm{out}}}{dt}$. When evaluated in the centre-of-mass frame, these terms correspond to the cases where the scattering angle of the twin particle is respectively less than and greater than the angle between its initial momentum and the total momentum of the system. However, when $T \neq \hat{T}$, we find the integrals involved in this decomposition substantially more arduous than when they are evaluated together.

Energy transferred through annihilations may be similarly calculated as

$$
\begin{aligned}
\frac{dq_{j\bar{j}\to i\bar{i}}}{dt} = {} & \frac{g_j^2}{(2\pi)^6} \int\int \frac{d^3k}{2E_j(k)}\frac{d^3h}{2E_j(h)} f_j(k)f_j(h) \\
& \qquad 4E_j(k)E_j(h)\int v_{rel}(E_j(h)+E_j(k))\frac{d\sigma_{j\bar{j}\to i\bar{i}}}{d\Omega}d\Omega \\
& -\frac{g_i^2}{(2\pi)^6}\int\int \frac{d^3k}{2E_i(k)}\frac{d^3h}{2E_i(h)} f_i(k)f_i(h) \\
& \qquad 4E_i(k)E_i(h)\int v_{rel}(E_i(h)+E_i(k))\frac{d\sigma_{i\bar{i}\to j\bar{j}}}{d\Omega}d\Omega,
\end{aligned}
\tag{5.6}
$$

where $\frac{d\sigma_{j\bar{j}\to i\bar{i}}}{d\Omega}$ is now the differential annihilation cross section. This rate may be evaluated as described above and is more directly amenable to the factorisation of the integrals observed in [346]. See also [347] for further details of similar calculations. The first term of (5.6) is the energy transferred from the SM to the twin sector and contributes to $\frac{dq_{\mathrm{in}}}{dt}$ in (5.2), while the second term is the energy transferred from the twin sector to the SM and contributes to $\frac{dq_{\mathrm{out}}}{dt}$.

In thermal equilibrium, the rate of energy transferred through collisions into one sector will be balanced by that of energy transferred out of it so that there is negligible net heat flow. This state will be rapidly attained (compared to the age of the universe) if $\frac{dq_{\mathrm{in,out}}}{dt} \gg 3H\hat{T}s_{\mathrm{t}}$. However, as the universe expands and the plasma cools, the energy transfer rates fall faster than the Hubble rate. This is demonstrated in the Figure 5.2 below. Once they drop below the Hubble rate, energy exchange ceases on cosmological time scales and the sectors thermally decouple, thereafter thermodynamically evolving independently.

To determine the decoupling temperature of the sectors, we calculate the rates of positive energy exchange for the twin particles interacting with the SM particles. The cross sections are calculated using a tree-level effective fermion-twin fermion contact interaction that, in the full twin Higgs model, would be UV completed by a SM Higgs exchange

(the heavier mass of the radial mode would make its exchange subdominant). The interaction strength is determined by the masses of the fermions through their Yukawa couplings, as well as the mixing angle of the SM-like mass state h with the gauge eigenstate h_B, giving a 4-fermion coupling of strength $\frac{m_f m_{\hat{f}}}{m_h^2 f^2}$ (here m_f and $m_{\hat{f}}$ are the masses of fermions f and $\hat{f}$). See [270], [323] for a more detailed discussion of the cross sections. This effective interaction is appropriate for the temperatures of interest here and helps to simplify the integrals of (5.4). In order to further simplify the integrations of (5.4) when it is to be decomposed into terms in which the energy exchange is positive and negative, we calculate $\frac{dq_{\mathrm{in}}}{dt}$ under the assumption that the sectors have the same temperature (this ensures that the rate $\frac{dq_{\mathrm{out}}}{dt}$ is identical). This is then combined with the rate of energy transferred from annihilation. A similar calculation of these rates was recently performed in [323], for cases where the Yukawa couplings do not respect the $\mathbb{Z}_2$ twin symmetry.

In Figure 5.2 we compare the energy transfer rate to the Hubble rate in order to determine when decoupling occurs. As long as the energy exchange rate exceeds the expansion rate, the sectors will be thermalised and have the same temperature. Decoupling then occurs once this rate drops below the Hubble rate. From Figure 5.2, this occurs at a temperature ~ 2 GeV. However, even after the energy exchange rate drops below the Hubble rate, the sectors will remain at the same temperature unless some event that either injects or redistributes entropy occurs within a sector (such as the temperature dropping below a mass threshold). As the heavy quark masses roughly coincide with the decoupling temperature, these do cause the twin sector to be mildly reheated with respect to the SM below decoupling. However, the resulting temperature difference is small and the energy exchange rates are expected to continue to be well-approximated by the rates presented in Figure 5.2 beyond decoupling.

The lower plot of Figure 5.2 illustrates the decomposition of the energy exchange rates into contributions from interactions involving different SM quarks. The interaction

cross sections are proportional to the Yukawa couplings of the interacting fermions. The greatest heat exchange is therefore expected to be mediated by the most massive particles, provided that their abundances are not too Boltzmann suppressed. As expected, at temperatures ~ 1 GeV, the bottom quark is the best conduit of thermal equilibration, followed by the charm quark and then the τ (with colour factors enhancing the former two with respect to the latter). The rate of heat flow that the top quarks and weak bosons can mediate at these temperatures (or below) is negligible because of Boltzmann suppression. The bend in the curves at temperatures ~ 5 GeV in the lower plot corresponds to a transition from temperatures where the dominant energy exchange rate is through scatterings to those where it occurs through annihilations, as can be seen in the upper plot. The annihilation rate into twin bottom quarks is the dominant component at high enough energies (again because of the larger Yukawa coupling), but this becomes rapidly threshold suppressed as the temperature drops. As can also be inferred in the upper plot, the energy exchange rate through annihilations involving the twin charmed quarks and tau leptons overtakes that of twin bottom quarks at similar temperature, but are still subdominant to scatterings.

The decoupling temperature depends upon f/v, which sets both the mass scale of the twin sector and the strength of the Higgs-mediated coupling. As f/v is increased, decoupling occurs earlier because of the greater Boltzmann suppression, although this is only a relatively small effect that, for $f/v = 10$, increases the decoupling temperature by only 4 GeV.

When the twin sector is colder than the SM (which will be important for much of what follows) the heat flow is typically dominated by annihilations of SM into twin particles. However, the energy exchange from elastic scattering can be comparable to that from annihilations, as illustrated in Figure 5.2. Although the energy exchange in an annihilation will generally exceed that of a scattering because all of the energy involved
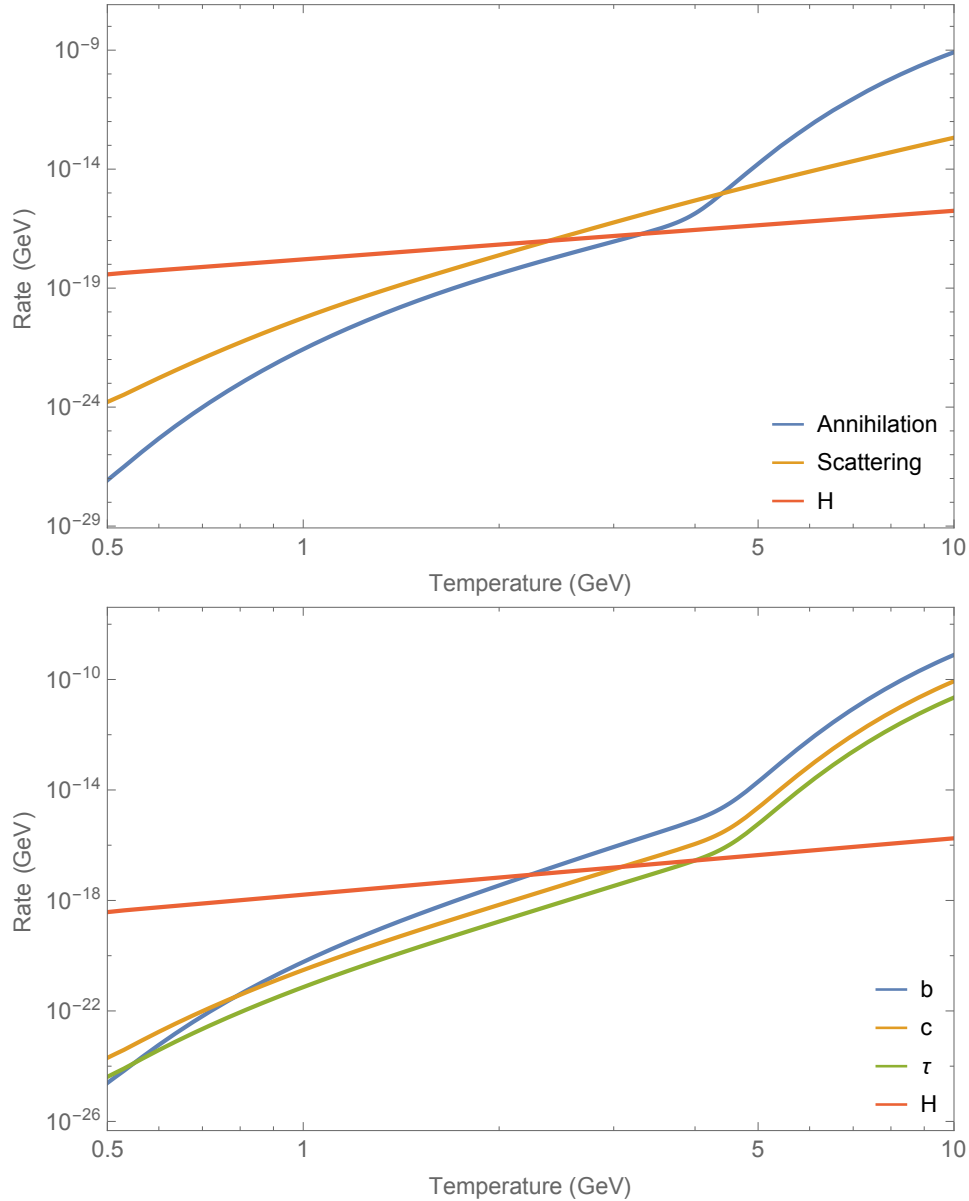
Figure 5.2: Rates of energy density exchange per twin entropy density $\left(\frac{1}{3s_{\mathrm{t}}\hat{T}}\frac{dq_{\mathrm{in}}}{dt}\right)$ decomposed into contributions from scattering and annihilation (top) and for interactions involving different species of SM fermions (bottom), along with the Hubble parameter, for $f/v = 4$. The decoupling temperature is that where the sum of the energy exchange rates equals the Hubble rate, which occurs at $T_{\mathrm{decoup}} \approx 2$ GeV.

in the process must be transferred, the annihilation rate also becomes more Boltzmann

or threshold suppressed when the temperature drops below the mass of the heavier twin

particles. It is therefore not always clear that energy transfer through annihilations dominates.

Decoupling is not exactly instantaneous and there is some range of temperatures over which the rate of heat flow freezes-out. The net heat flow rate $\frac{dq}{dt}$ is greater for larger temperature differences between sectors. The generation of a potentially large temperature difference within this brief epoch of sector decoupling, such as those discussed below in Section 5.2.3, may be cut off when the heat flow rate becomes comparable to the Hubble rate. For a given SM temperature T, the minimum twin-sector temperature $\hat{T}_{\min}$ during the decoupling period may be roughly estimated as that which satisfies

$$H \sim \frac{1}{3s_{\mathrm{t}}\hat{T}} \frac{dq}{dt}\bigg|_{\hat{T}=\hat{T}_{\min}}. \tag{5.7}$$

Twin temperatures colder than $\hat{T}_{\min}$ will partially thermalise back to this value. As the participating fermions are not non-relativistic, instantaneous decoupling is not as accurate an approximation as it is, for example, for chemical decoupling of a WIMP, although it is still reliable.

In Figure 5.3, we show the minimum temperature that the twin sector may have as a function of SM temperature for heat flow to freeze out, estimated using (5.7). Only annihilations have been included in the determination of the minimum temperature, although we have verified that, for these temperatures, the scatterings contribute only $\lesssim 10\%$ to the heat flow. Note that while the energy exchange rate, such as $\frac{1}{\hat{T}} \frac{dq_{\mathrm{in}}}{dt}$ in (5.2), in scattering processes may be faster, the net energy flow rate, or heat flow ($\frac{1}{\hat{T}} \frac{dq}{dt}$ in (5.2)), which is the difference between energy exchange rates into and out of the sector, is actually dominated by annihilations. Generally, we find that decoupling begins at temperatures ~ 4 GeV. The temperature difference can reach an order of magnitude without relaxing once the SM temperature drops to ~ 1 GeV.
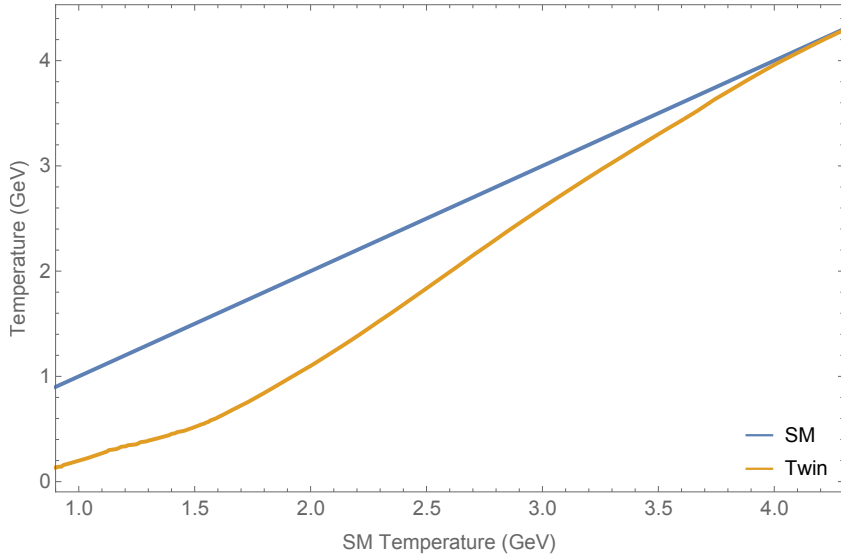
Figure 5.3: Minimum temperature of the twin sector that will not be heated by inter-
actions with a hotter SM plasma, as a function of SM temperature, for $f/v = 4$. Also
shown is the SM temperature, for reference.

While the extent of thermal decoupling is temperature dependent, the maximum
temperature difference that will not relax grows quickly as the SM temperature drops.
Then we may describe the two sectors as being decoupled if, in a given cosmology, all
events that raise the temperature of one sector relative to the other (such as the crossing
of a mass threshold and the resulting entropy redistribution, the most significant of which
is the confinement of colour) induce temperature differences that are too small to partially
relax.

At energies $\lesssim 1$ GeV in Figure 5.2, the reliability of the calculation of the heat flow
rate diminishes because of the strengthening of the strong coupling and the eventual
confinement of colour. Fortunately, for a cooler twin sector, which will be of interest in
subsequent sections, annihilations from the SM dominate other processes over most of
the parameter space. These are the least sensitive to higher order corrections and non-
perturbative effects because of their higher temperature, and hence energy, compared
to the potentially cooler twin sector. The range of temperatures illustrated in Figures

5.2 and 5.3 have been selected to roughly illustrate the duration of decoupling, but may extend below the range where the perturbative calculation of the heat flow rate is valid. For example, at temperatures below the twin sector QCDPT, which occurs at $\sim \left(1 + \log(\frac{f}{v})\right)$ higher temperatures than in the SM, the partonic calculation of twin quark/anti-quark pair production must be replaced by a hadronic one. Furthermore, the growth of the twin strong coupling necessitates that the quark-Higgs Yukawa couplings be RG evolved to the scale of the energy exchanged, which can induce an $\mathcal{O}(1)$ change to the cross section, although this has only a relatively small effect on the decoupling temperature. It is nevertheless clear that decoupling is mostly complete by then and that these uncertainties are not large enough to affect this conclusion.

In the standard mirror Twin Higgs cosmology, knowing the decoupling temperature tells us how the temperatures of the two sectors will be related at subsequent times. The sectors separately evolve adiabatically after decoupling, though they redshift in the same way and differences in temperature only arise from events that redistribute entropy. Non-minimal cosmological events that could potentially cause the temperatures of each sector to diverge can therefore only be effective if they leave each sector colder than this approximate decoupling temperature.

Cosmological Constraints

Given that the twin and Standard Model sectors remain in thermal equilibrium to $\mathcal{O}(\text{GeV})$ temperatures, the simplest mirror Twin Higgs scenario is cosmologically inviable due to the presence of light twin species (photons and neutrinos) with abundances comparable to those of the SM. The cosmological observables through which evidence of light species may be inferred are typically represented by N_{eff}, the "effective number of neutrino species" in the early universe; their individual masses, which determine their free-streaming distances; and the "effective mass" m_{ν}^{eff}, which parameterises their contri-

bution to the present-day energy density of non-relativistic matter. These observables are probed by both the CMB and large scale structure (LSS).

Effective number of neutrinos   The parameter N_{eff} describes the amount of radiation-like energy density during the evolution of the CMB anisotropies before photon decoupling. It is defined as the effective number of massless neutrinos with temperature as predicted in the standard cosmology that would give equivalent energy density in radiation:

$$\rho_r = \rho_\gamma + \frac{7}{8} \left(\frac{4}{11} \right)^{4/3} N_{\text{eff}} \rho_\gamma, \tag{5.8}$$

where ρ_r is the energy density of radiation and ρ_γ is the energy density of photons (the factor of $\left(\frac{4}{11} \right)^{4/3}$ arises from the relative reheating of the photons from electron/positron annihilation, which occurs after most of the neutrinos have decoupled, and the factor of 7/8 is from the opposite spin statistics). A deviation from the Standard Model prediction of 3.046 [348] is denoted by $\Delta N_{\text{eff}} = N_{\text{eff}} - 3.046$. This definition of radiation, or equivalently, relativistic degrees of freedom, becomes less clear if the new fields have a non-negligible mass, as we discuss further below.

We here review the CMB physics of dark radiation, summarising the discussion in [349]. See also [318] for further review. The angular size and scale of the first acoustic peak is well-measured and this approximately fixes the scale factor at matter-radiation equality a_{eq}. If we imagine fixing all other ΛCDM parameters, extra radiation would delay the epoch of matter-radiation equality. This would have a pronounced effect on the power spectrum in the vicinity of the first acoustic peak through the early Integrated Sachs-Wolfe (eISW) effect. The modes corresponding to this feature entered the horizon close to matter-radiation equality and the evolution of their potentials is highly sensitive to the radiation energy density. However, the impact of a $\Delta N_{\text{eff}} \sim \mathcal{O}(1)$ deviation on the

peak height can be simultaneously balanced by increasing the amount of non-relativistic matter, to the extent to which other observations providing independent constraints upon Ω_c permit (for ΛCDM+N_{eff}, a variation of $\sim 10\%$ in $\Omega_c h^2$ is consistent with present CMB+BAO measurements [318], although these variations must be consistent with other observables). This degeneracy is not expected to be broken by CMB-S4 [350].

Given that a_{eq} is approximately fixed, the utility of N_{eff} arises because, in simple extensions of the ΛCDM model, it approximately corresponds to the suppression of power in the small scale CMB anisotropies that arises from Silk damping. The reason for this is roughly that, although the greater expansion rate induced by the extra radiation reduces the time that CMB photons have to diffuse before decoupling, it also reduces the sound horizon size more severely. As the angular size of the sound horizon is determined by the location of the acoustic peaks and is also well measured, the reduction in the sound horizon must be compensated for by a reduction in the angular diameter distance to the CMB. This effectively raises the angular distance over which photon diffusion proceeds and results in a prediction of smoother temperature anisotropies at small scales. This correspondence with the Silk damping allows N_{eff} to be approximately factorised from other parameters and constrained independently, providing a direct observational avenue for detecting the presence of new, massless fields [349] (see [351] for further implications for model building). This relationship arises because the fixing of a_{eq} implies that N_{eff} effectively determines the energy density of the universe, and hence the Hubble rate, during CMB decoupling. Note, however, that further extensions of ΛCDM may complicate this correspondence, in particular deviations from the standard Big Bang Nucleosynthesis prediction of the primordial helium abundance.

The contribution to N_{eff} (or ΔN_{eff}) in the mirror Twin Higgs arises from two sources: the twin photons, which can be treated as massless dark radiation with an appropriate twin temperature T_{eq}^{t} at the time of matter-radiation equality, and the twin neutrinos,

whose non-zero masses may need to be accounted for. For the twin photons, the contribution to N_{eff} is simple; their equation of state is always $w = 1/3$ and their energy density is given by $g\frac{\pi^2}{30}\left(T^{\mathrm{t}}_{\mathrm{eq}}\right)^4$, where $g = 2$. The twin temperature at matter-radiation equality is found from the SM temperature using comoving entropy conservation,

$$\frac{T^{\mathrm{t}}_{\mathrm{eq}}}{T^{\mathrm{SM}}_{\mathrm{eq}}} = \left(\frac{g^{\mathrm{t}}_{\star}(T_{\mathrm{decoup}})}{g^{\mathrm{SM}}_{\star}(T_{\mathrm{decoup}})}\right)^{1/3}\left(\frac{g^{\mathrm{SM}}_{\star}(T^{\mathrm{SM}}_{\mathrm{eq}})}{g^{\mathrm{t}}_{\star}(T^{\mathrm{t}}_{\mathrm{eq}})}\right)^{1/3}, \tag{5.9}$$

where the two sectors have the same number of thermalized degrees of freedom by this time. Here, $T^{\mathrm{SM}}_{\mathrm{eq}}$ is the SM photon temperature at matter-radiation equality and T_{decoup} is the sector decoupling temperature.

Since neutrinos are massive, their behavior is more complicated. Their equation of state parameter takes on a scale factor dependence which is controlled by their mass. In the Standard Model, this sensitivity is negligible because present CMB bounds imply that neutrinos are ultra-relativistic at a_{eq} to good approximation [318]. However, the factor by which the twin neutrino masses are enhanced may raise them to order T^{t}_{eq} or greater (see Section 4.1.1 for discussion of the scaling of the masses with f/v).

To better describe the impact of the extra twin (semi-)relativistic degrees of freedom on the CMB, we choose to define N_{eff} through the effects of neutrinos at matter-radiation equality, when the impact on the expansion rate of the universe for most of the period relevant for the evolution of the CMB is greatest. Note that, in their presentation of joint exclusion bounds on N_{eff} and $\sum m_{\nu}$ (the sum of SM neutrino masses) or m^{eff}_{ν} (effective mass contributing to the present-day non-relativistic matter density of an extra sterile neutrino), the Planck collaboration define N_{eff} as the value in (5.8) at temperatures sufficiently high that the neutrinos are fully relativistic. Our values cannot be directly compared with their analysis, although we consider ours to be a reasonable rough estimate that is more representative of the CMB constraints. The ensuing correction from the finite

neutrino masses is, in the cases considered in this work, a small effect anyway.

To determine this correction and provide a definition of N_{eff} that better describes the impact of quasi-relativistic particles on the CMB, we first define the epoch of matter-radiation equality as the time at which the average equation of state parameter of the universe is $\bar{w} = 1/6$ (the equation of state is defined as $\rho = \bar{w}P$, where ρ is energy density and P is pressure). We can express this condition as

$$\frac{d\ln H}{d\ln a}\bigg|_{a_{eq}} = -\frac{7}{4},\tag{5.10}$$

as in [352].

Call the quasi-relativistic neutrino energy density $\tilde{\rho}(a)$ with time-evolving equation of state parameter $w(a)$, which is to be balanced against some extra non-relativistic energy density $\Delta\rho_{CDM}(a) \propto a^{-3}$ to keep a_{eq} the same. This amount of non-relativistic energy density $\Delta\rho_{CDM}$ is

$$\Delta\rho_{CDM}(a_{eq}) = \rho_r(a_{eq}) - \rho_m(a_{eq}) - 2a_{eq}\frac{d\tilde{\rho}}{da}\bigg|_{a_{eq}} - 7\tilde{\rho}(a_{eq}),\tag{5.11}$$

where ρ_r and ρ_m are the energy densities of the radiation and non-relativistic matter.

For a perfect fluid, $\frac{d\tilde{\rho}}{da} = -3(1+w(a))\tilde{\rho}/a$ (neglecting the anisotropic stress that is expected only to contribute to a weak phase shift in the CMB [353]), this results in a Hubble parameter of

$$H^2(a_{eq}) = \frac{2}{3M_{\text{pl}}^2}\left[\rho_r(a_{eq}) + 3w(a_{eq})\tilde{\rho}(a_{eq})\right].\tag{5.12}$$

This suggests a definition of the effective number of neutrinos, N_{eff}, via

$$H^2(a_{eq}) \;\; = \;\; \frac{2}{3M_{\text{pl}}^2} \left(\rho_\gamma + N_{\text{eff}} \rho_{\nu,m=0}^{th} \right)\big|_{a_{eq}} \tag{5.13}$$

$$N_{\text{eff}} \;\; \equiv \;\; \sum_i \frac{w_i}{1/3} \frac{\rho_i}{\rho_{\nu,m=0}^{th}}, \tag{5.14}$$

where ρ_i is the contribution to the energy density from some species i with equation of state parameter w_i and $\rho_{\nu,m=0}^{th}$ is the energy density of a massless neutrino with a thermal distribution in the standard cosmology. Then $3w$ gives the 'relativistic fraction' of the energy density. Note that this is simply a ratio of the pressure exerted by the new fields to that of a massless neutrino. The effectiveness of this approximation was discussed in [354] in the context of thermal axions (while effective at keeping a_{eq} fixed, changes to odd peak heights subsequent to the first are imperfectly cancelled and require further changes to H_0 to compensate - see Section 5.2.2 below).

Calling T_ν^i the temperature at which the neutrinos in sector i freeze-out and a_ν^i the corresponding scale factor, then assuming instantaneous decoupling, the phase space number density for scale factor a is given by a redshifted Fermi-Dirac distribution [355]

$$f_\alpha^i(p) \approx \left[1 + e^{pa/\left(a_\nu^i T_\nu^i \right)} \right]^{-1} \tag{5.15}$$

for the α neutrino mass eigenstate in the i sector ($m_\alpha^i \ll T_\nu^i$, so has been dropped). The energy density and pressure are

$$\rho_{\nu_\alpha}^i \;\; = \;\; \frac{g_\alpha}{2\pi^2} \int_0^\infty dp \; p^2 \sqrt{p^2 + (m_\alpha^i)^2} \, f_\alpha^i(p) \tag{5.16}$$

$$P_{\nu_\alpha}^i \;\; = \;\; \frac{g_\alpha}{2\pi^2} \int_0^\infty dp \; \frac{p^4}{3\sqrt{p^2 + (m_\alpha^i)^2}} f_\alpha^i(p), \tag{5.17}$$

where $g_\alpha = 2$ is the number of degrees of freedom for a neutrino species.

171

Since the neutrino decoupling temperature depends on the strength of the weak interaction as $T_\nu \propto G_F^{-2/3}$, while $G_F \propto v^2$, then the twin neutrino decoupling temperature T_ν^{t} is related to the SM neutrino decoupling temperature T_ν^{SM} by

$$T_\nu^{\mathrm{t}} = (f/v)^{4/3} T_\nu^{\mathrm{SM}}. \tag{5.18}$$

We can then simply use (5.16) and (5.17) at matter-radiation equality to find ΔN_{eff} (assuming instantaneous decoupling). We thus obtain

$$H^2(a_{eq}) = \frac{2}{3M_{\mathrm{pl}}^2} \left(\rho_\gamma^{\mathrm{SM}} + 3.046 \rho_{\nu,m=0}^{th} + \rho_\gamma^{\mathrm{t}} + \sum_\alpha 3 w_{\nu_\alpha} \rho_{\nu_\alpha}^{\mathrm{t}} \right) \Bigg|_{a_{eq}} \tag{5.19}$$

and

$$\Delta N_{\mathrm{eff}} = \left(\frac{11}{4} \right)^{4/3} \frac{120}{7\pi^2 \left(T^{\mathrm{SM}} \right)^4} \left(\rho_\gamma^{\mathrm{t}} + \sum_\alpha 3 w_{\nu_\alpha}^{t} \rho_{\nu_\alpha}^{\mathrm{t}} \right), \tag{5.20}$$

where we now have equation of state parameters w_{ν_α} for each neutrino, while $\rho_\gamma^{\mathrm{SM}}$ and ρ_γ^{t} are the SM and twin photon energy densities, $\rho_{\nu,m=0}^{th}$ and $\rho_{\nu_\alpha}^{\mathrm{t}}$ are the neutrino energy densities.

Neutrino masses   Because they are so weakly interacting, the neutrinos have a long free-streaming scale given by the distance travelled in a Hubble time v_ν/H, with $v_\nu \propto m_\nu^{-1}$ the speed of the neutrino once it becomes non-relativistic. This defines a free-streaming momentum scale $k_{fs} = \sqrt{\frac{3}{2}} \frac{aH}{v_\nu} \propto m_\nu$, above which neutrinos do not cluster. Below this scale, perturbations in the matter density consist coherently of neutrinos and other matter, but well above it only non-neutrino matter contributes to density perturbations. This results in a suppression of the matter power spectrum on large scales which is proportional to the fraction of energy density in the free-streaming matter. Since this

172

occurs at late times when neutrinos are non-relativistic, the energy density is simply $\rho_{\nu_\alpha} = n_{\nu_\alpha} m_{\nu_\alpha}$ for each neutrino species α, where n_{ν_α} is the number density. Constraints on the sum of neutrino masses then come from the observations of power on small scales, which is suppressed relative to that expected for massless neutrinos by a factor $\simeq 1 - 8 f_\nu$, where $f_\nu = \Omega_\nu / \Omega_m$ is the fraction of non-relativistic energy in neutrinos at late times [356].

More generally, inferences of the matter power spectrum constrain the present-day energy density fraction of free-streaming species that do not cluster on small scales and have since become non-relativisitic, $\Omega_\nu = (\sum m_\nu + m_\nu^{\mathrm{eff}})/(94.1\,\mathrm{eV})$, where $\sum m_\nu$ is the sum of SM neutrino masses and m_ν^{eff} is the sum of twin neutrino masses weighted by their number density

$$m_\nu^{\mathrm{eff}} = \frac{n_\nu^{\mathrm{t}}}{n_\nu^{\mathrm{SM}}} \sum_\alpha m_{\nu_\alpha}^{\mathrm{t}}. \tag{5.21}$$

Here n_ν^{t} is the number density of a relic twin neutrino flavour and n_ν^{SM} is that for a SM neutrino. It is assumed that the neutrinos have been thermally produced as hot relics.

The relic abundance of a neutrino species is given by its number density when it decoupled, diluted by the factor by which the universe has since expanded. The scale factors at which neutrino decoupling occurs in the two sectors, a_ν^{SM} and a_ν^{t} can be determined from (5.18), the relative temperatures in the two sectors and comoving entropy conservation, to obtain

$$a_\nu^{\mathrm{t}} \;=\; a_\nu^{\mathrm{SM}} \left(\frac{v}{f}\right)^{4/3} \left(\frac{g_\star^{\mathrm{t}}\left(T_{\mathrm{decoup}}\right)}{g_\star^{\mathrm{SM}}\left(T_{\mathrm{decoup}}\right)}\right)^{1/3} \tag{5.22}$$

where the same mass thresholds have been assumed in each sector below their neutrino decoupling temperatures, so that $g_\star^{\mathrm{SM}}\left(T_\nu^{\mathrm{SM}}\right) = g_\star^{\mathrm{t}}\left(T_\nu^{\mathrm{t}}\right)$. The neutrino number densities

are then

$$\frac{n_\nu^{\text{t}}}{n_\nu^{\text{SM}}} = \left(\frac{T_\nu^{\text{t}} a_\nu^{\text{t}}}{T_\nu^{\text{SM}} a_\nu^{\text{SM}}}\right)^3 = \frac{g_\star^{\text{t}}\left(T_{\text{decoup}}\right)}{g_\star^{\text{SM}}\left(T_{\text{decoup}}\right)}. \tag{5.23}$$

For f/v from 3 to 10 and using $T_{\text{decoup}} \sim 2 - 6$ GeV from Section 5.2.2, we find $g_\star^{\text{t}}\left(T_{\text{decoup}}\right) / g_\star^{\text{SM}}\left(T_{\text{decoup}}\right) \sim 0.8$ and thus arrive at

$$m_\nu^{\text{eff}} \approx 0.8 \left(\frac{f}{v}\right)^n \sum_\alpha m_{\nu_\alpha}^{\text{SM}}, \tag{5.24}$$

where $n = 1$ for Dirac masses and $n = 2$ for Majorana masses.

If they are sufficiently light and hot, the twin neutrinos only affect the CMB as dark radiation and their masses may then only be inferred from tests of the matter power spectrum. However, if heavier and colder, they are better described as a hot dark matter component. Their impact on the CMB is discussed in [357], where the shape of the power spectrum can depend upon the individual neutrino kinetic energies through their characteristic free-streaming lengths. The early Integrated Sachs-Wolfe effect (eISW) is also sensitive to the masses if the neutrinos become non-relativistic during decoupling (thereby affecting the radiation energy density and the growth of inhomogeneities) [356].

There is a significant degeneracy in cosmological fits to the CMB between Ω_m and H_0 (the Hubble constant) [?], where raising the non-relativistic matter fraction, such as with nonrelativistic neutrinos, can be accommodated by a decrease in H_0 (or equivalently, the dark energy density), which keeps the angular diameter distance to the CMB approximately fixed. This degeneracy can be broken by measurements of the baryon acoustic oscillations (BAOs), which are sensitive to the expansion rate of the late universe and provide an independent measurement of Ω_m and H_0. It is through combination with these results that bounds from Planck on neutrino masses are strongest [318].

Bounds   The authors are unaware of any specialised analysis of the present and pro-jected future cosmological constraints on scenarios with both massless dark radiation and additional light, semi-relativistic sterile neutrinos. In the absence of this, we use bounds from [318] as a rough indication of the present level of sensitivity to these pa-rameters, which we nevertheless expect to be a reliable indication of the (in)viability of this model. The 95% confidence limits on these parameters are $N_{\text{eff}} = 3.2 \pm 0.5$ and $\sum m_\nu < 0.32$ eV when each are constrained separately with the other fixed. This, of course, overlooks correlations between the impacts of masses and ΔN_{eff} on the CMB and LSS. Bounds on an additional sterile neutrino as the only source of dark radiation are also presented with number density, or equivalently, contribution to ΔN_{eff}, left to float. These are similar to the limit on $\sum m_\nu$. It was found in [358] that, allowing $\sum m_\nu$ and m_ν^{eff} to float independently for a single extra sterile neutrino, the bound mildly relaxes to $m_\nu^{\text{eff}} \lesssim 1$ eV, although the bound may be stronger depending on the combination of data sets chosen (the lensing power spectrum presently prefers higher neutrino masses and raises the combined bounds if included). Other bounds from LSS on $\sum m_\nu$ exist and are potentially stronger than those placed from the CMB, possibly as low as $m_\nu^{\text{eff}} \lesssim 0.05$ eV, again depending on data sets combined (see [359], [360]), although these are subject to greater uncertainties in the inference of the power spectra of dark matter halos from galaxies surveys and the Lyα forest.

It must also be noted that the shape of the CMB temperature anisotropies depends upon both the mass of individual neutrino components (through their free-streaming distance) and their contribution to the energy density of the nonrelativistic matter that does not cluster on small scales. However, it is not expected that improvements in bounds on the former will be made from improved measurements of the primary CMB itself, but rather from weak lensing of the CMB, in conjunction with future measurements from DESI of the BAOs to break degeneracy with Ω_m. The lensing spectrum, like

inferences of the matter power spectrum made in galaxy surveys, is expected to measure the suppression of small scale power and therefore to strengthen constraints upon m_ν^{eff}, rather than the individual neutrino masses. One of the goals of CMB-S4 will be the detection of neutrino masses, given the present lower bound $\sum m_\nu \gtrsim 0.06$ eV from oscillations. Projected bounds are as low as ~ 0.02 eV [350], although this assumes no extra dark radiation or sterile neutrinos. A projection of the joint bound on N_{eff} (from extra massless dark radiation) and m_ν^{eff} combining improved measurements CMB temperature measurements, lensing and BAOs indicates a limit of $m_\nu^{\text{eff}} \lesssim 0.1$ eV at 1σ [350]. Any contribution from additional states to m_ν^{eff} may therefore be testable and bounded by the excess of the neutrino mass inference over the minimum neutrino mass, although laboratory measurements or measurements of ΔN_{eff} will be required to further ascertain the contribution from the new particles.

Constraints on ΔN_{eff} from improved measurements of the damping tail as part of CMB-S4 are projected to be $\sim 0.02 - 0.05$ at 1σ [350]. In the following sections, we use an optimistic estimate of 0.02 for its reach in order to identify as much of the potentially testable parameter space as possible.

To estimate the impact of current and projected CMB limits on the mirror Twin Higgs, we consider two scenarios: the minimal Standard Model neutrino mass spectrum of $m_\nu = [0.0, 0.009\text{ eV}, 0.06\text{ eV}]$ and a degenerate spectrum of $m_\nu = [0.1\text{ eV}, 0.1\text{ eV}, 0.1\text{ eV}]/3$ from [318]. In Figure 5.4 we plot the predictions of the mirror Twin Higgs for ΔN_{eff} and m_ν^{eff} for both types of spectra, as well as for both Dirac and Majorana masses (which scale differently with f/v). As is plainly evident, the mirror Twin Higgs is ruled out cosmologically, no matter the choices of neutrino masses one makes, if only for the presence of the twin photon. In the standard cosmology, the twin sector will have roughly the same temperature as the SM, giving $4.6 \lesssim \Delta N_{\text{eff}} \lesssim 6.3$ for $f/v < 10$, according to the definition of (5.20). This range depends upon f/v through the twin neutrino decoupling temperature
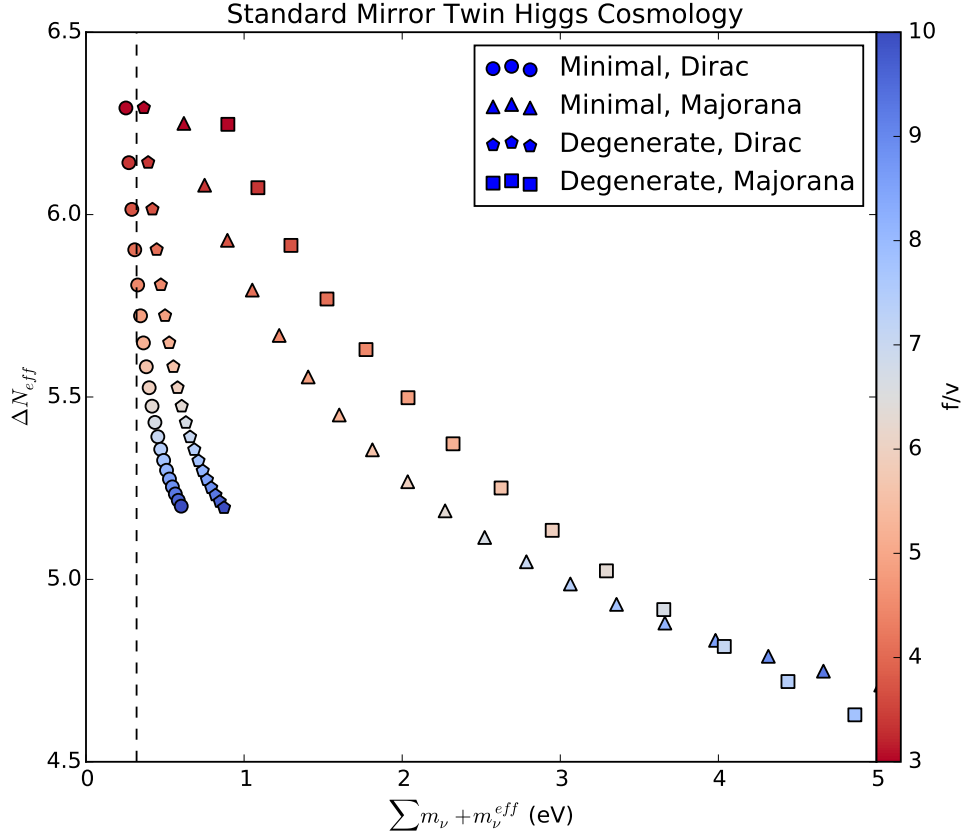
Figure 5.4: Predicted values of ΔN_{eff} and $\sum m_\nu + m_\nu^{\text{eff}}$ for minimal and degenerate neutrino mass spectra with both Dirac and Majorana masses for f/v from 3 to 10. The Planck 2015 constraint[318] is the dashed line; the corresponding N_{eff} upper bound is well below the bottom of the plot. All points are excluded by the combination of bounds on ΔN_{eff} and $\sum m_\nu + m_\nu^{\text{eff}}$.

(5.18), which determines the extent to which the twin photons are reheated relative to the

twin neutrinos after twin electron/positron annihilations. This is sufficiently large that

even the cold dark matter fraction cannot be adjusted to keep matter-radiation equality

fixed, resulting inevitably in changes to the height and shape of the first acoustic peak.

The energy density in neutrinos is predicted to be above the present observational upper

bounds for most neutrino mass configurations, with the exception of the minimal values

permitted by neutrino oscillation measurements with $f/v \lesssim 6$. We therefore discuss

cosmological mechanisms in which the twin radiation is diluted to levels compatible with these observational bounds in the subsequent sections of this paper.

5.2.3   Reheating by the decay of a scalar field

We now turn to simple scenarios that reconcile the mirror Twin Higgs with cosmological bounds, while taking care to respect the softly-broken $\mathbb{Z}_2$ symmetry. We begin with the out-of-equilibrium decay of a particle with symmetric couplings to the Standard Model and twin sectors, in which the desired asymmetry is generated kinematically. That is to say, the dimensionless couplings between the decaying particle and the two sectors are equal, and asymmetric energy deposition into the two sectors is a direct consequence of the asymmetric mass scales. In this respect, the scenario is philosophically similar to Nnaturalness [361], albeit with a parsimonious $N = 2$ sectors. See also [362], [363] and [347] for other recent related ideas of using long-lived particles for the dilution of dark sectors.

For simplicity, here we will focus on the case of a real scalar X coupled symmetrically to the A and B sector Higgs doublets. Due to the difference in masses between the sectors after electroweak symmetry breaking, simple kinematic effects give X a larger branching ratio into the Standard Model. This occurs over a range of X masses within a few decades of the weak scale. If X decays out-of-equilibrium below the decoupling temperature of the two sectors, this injects different amounts of energy into the two sectors, effectively suppressing the temperature of the twin sector relative to the Standard Model. This relative cooling suppresses the contribution of the light degrees of freedom of the mirror Twin Higgs to below cosmological bounds. Insofar as the asymmetry is driven entirely by kinematic effects arising from $v \ll f$, the resulting temperature inequality between the two sectors is proportional to powers of v/f.

178

The requisite suppression of the twin sector temperature relative to the Standard Model temperature necessitates that the X dominate the cosmology before it decays. Our main discussion will follow the simplest case of an X which dominates absolutely before it decays, comprising all of the energy density of the universe and effectively acting as a 'reheaton'. Afterwards, we will discuss the possibility of a 'thermal history' for X – a scenario where X is in thermal equilibrium with the two sectors, then chemically decouples at some high temperature and grows to dominate the cosmology before it decays. This scheme will result in additional stringent constraints on the viable parameter space.

Asymmetric Reheating

A $\mathbb{Z}_2$-even scalar X which is a total singlet under the SM and twin gauge groups admits the renormalisable interactions

$$V \supset \lambda_x X(X + x)\left(|H_A|^2 + |H_B|^2\right) + \frac{1}{2}m_X^2 X^2, \tag{5.25}$$

where m_X is the mass of X (neglecting corrections from mixing that will be shown below to be tiny), λ_x is a dimensionless coupling and x is a dimensionful parameter, which one may imagine identifying as a vacuum expectation value (vev) of X in an UV theory. Note that these interactions preserve the accidental $SU(4)$ symmetry of the Twin Higgs. The X field may additionally possess self-interactions, which we omit here as they do not play a significant role in what follows.

The interactions in (5.25) allow X to decay into light states in the Standard Model and twin sectors. If X reheats the universe through out-of-equilibrium decays, the reheating temperatures of the two sectors will be determined by its partial decay widths, assuming that the decay products do not equilibrate. In the instantaneous decay approximation,

179

X decays when the Hubble parameter falls to its decay rate $\Gamma_X \sim H$. As we will show in Section 5.2.3, in order to evade cosmological constraints we need the X to decay mostly into the SM, so we may estimate $\Gamma_X \sim \Gamma(X \to \text{SM})$. Then the energy that was contained in the X is transferred into radiation energy density, with the resulting temperature of the radiation given by (see [364])

$$T \sim 1.2 \sqrt{\frac{\Gamma_X M_{\text{pl}}}{\sqrt{g_\star}}} \tag{5.26}$$

where $g_\star$ is the effective number of relativistic degrees of freedom, as defined in Section 5.2.2, of the particles that are being reheated. Our numerical calculation of the reheating temperature, which will be presented in Section 5.2.3, indicates that the approximation $T \sim 0.1 \sqrt{\Gamma_X M_{\text{pl}}}$ reliably reproduces the reheating temperature over the range of interest.

As shown in Section 5.2.2, the two sectors thermally decouple when the temperature falls below $T_{\text{decoup}} \sim 1$ GeV, so reheating must take place to below this temperature. At even lower temperatures, big bang nucleosynthesis (BBN) places strong constraints on energy injected into the SM at temperatures below $\mathcal{O}(1-10)$ MeV [365]. Requiring that the SM reheating temperature is above ~ 10 MeV, these constraints on the SM reheating temperature become constraints on the decay rate of the X into the SM, which in the above approximation becomes

$$5 \times 10^{-21} \text{ GeV} \lesssim \Gamma_X \lesssim 3 \times 10^{-16} \text{ GeV}. \tag{5.27}$$

This then constrains the couplings λ_x and x of the X to the Higgs sector. Importantly, it means that X must couple very weakly, in order to be long-lived enough to reheat to a low temperature, as will be shown below.

The asymmetry in partial widths arises from different effects depending upon the mass of X. For masses below the SM Higgs threshold, it is predominantly differences in mass mixing with the two Higgs doublets that produces the asymmetry, where the size of the mixing angles determines the effective coupling of X to the SM and twin particles and therefore its branching fractions. For masses below the twin scale, the relative size of the mixing scales inversely with the vevs in each sector. Thus the hierarchy $v \ll f$ already present in the Higgs sector can automatically gives rise to a hierarchy in partial widths. Note that additional threshold effects can enhance the asymmetry further, in particular when X has mass above threshold for a significant decay channel in the SM, but below the corresponding mass threshold in the twin sector. Decays into on-shell Higgses complicate this picture further. In what follows, we first give an analytic calculation of the mass mixing effect, then present a more precise calculation of the decay widths into each sector.

To lowest order, X decays via its interactions with the SM and twin Higgs, and only to other fermions and gauge bosons through its mass mixing with the Higgs scalars. Expanding the X potential after the $SU(4)$ is spontaneously broken, the mixing term between X and h_A in the scalar mass matrix is $\sqrt{2}\lambda_x x v_A$, while that between X and h_B is $\sqrt{2}\lambda_x x v_B$. The h_A and h_B components of the X mass eigenstate, which we denote respectively as δ_{XA} and δ_{XB}, can then be determined. The expressions for the mixing angles are in general complicated, but they simplify in limits $m_X < f$ and $m_X \gg f$:

$$(\delta_{XA}, \delta_{XB}) \approx \begin{cases} \frac{4\lambda_x x v_A}{m_X^2 - m_h^2} \left(\frac{1}{\sqrt{2}}, \frac{v_A}{f} \right) & m_X < f \\ \frac{\lambda_x x f}{m_X^2} \left(\frac{\sqrt{2} v_A}{f}, 1 \right) & m_X \gg f \end{cases} \tag{5.28}$$

to lowest order in $(v/f)^2$ and κ/λ. The partial width for the decay of X into SM states

(excluding the Higgs) is

$$\Gamma(X \to \text{SM}) \approx |\delta_{XA}|^2 \, \Gamma_h(m_h = m_X), \qquad (5.29)$$

where $\Gamma_h(m_h = m_X)$ denotes the decay width of a SM Higgs if it were to have mass m_X. Note that the Higgs partial width must be computed using the vev $v_A \approx v/\sqrt{2}$ to determine the masses and couplings of the SM particles. The partial width of the X into twin states is computed the same way using δ_{XB} and the vev $v_B \approx f/\sqrt{2}$.

From the mixing angles (5.28), it is already apparent over what mass range asymmetric reheating from X decays will work. These give

$$\frac{\Gamma(X \to \text{SM})}{\Gamma(X \to \text{Twin})} \sim \begin{cases} f^2/v_A^2 \gg 1 & m_X < f \\[2mm] v_A^2/f^2 \ll 1 & m_X \gg f. \end{cases} \qquad (5.30)$$

Thus when the mass of X is less than the twin scale, the Standard Model will be reheated to a higher temperature than the twin sector, but in the large mass limit this mechanism works in the opposite direction and would appear to lead to preferential reheating of the twin sector.

More precise statements about the relative branching ratios and resulting temperatures require additional care. In addition to decaying through mass mixing, X can decay into the Higgs mass eigenstates themselves if above threshold. As the energy is ultimately transferred to the SM and twin sectors, we then need to consider how these states decay and account for the further mixing of the Higgs mass eigenstates into Higgs gauge eigenstates.

For $m_X > 2m_h$, decay can occur into the lighter (SM-like) Higgs mass eigenstate h

with partial width

$$\Gamma(X \to hh) \approx \frac{\lambda_x^2 x^2}{16\pi m_X} \sqrt{1 - \left(\frac{2m_h}{m_X}\right)^2}. \tag{5.31}$$

Similarly, for $m_X > 2m_H$, decays can proceed into HH with a similar partial width, but with the h mass replaced with that of the H. Above the intermediate threshold $m_X > m_h + m_H$, there is also the mixed decay

$$\Gamma(X \to hH) \approx \frac{\lambda_x^2}{2\pi m_X} \sqrt{1 - \left(\frac{m_H + m_h}{m_X}\right)^2} \left(f\delta_{AX} + 2v_A\delta_{BX}\right)^2. \tag{5.32}$$

Here, $\delta_{AX} \approx -\delta_{hA}\delta_{XA} - \delta_{hB}\delta_{XB}$ is the component of the h_A gauge eigenstate in the X mass eigenstate and $\delta_{BX} \approx \delta_{hB}\delta_{XA} - \delta_{hA}\delta_{XB}$ is the corresponding component of the h_B gauge eigenstate, where δ_{hA} and δ_{hB} are, respectively, the components of the SM Higgs in the h_A and h_B gauge eigenstates to zeroth order in λ_x. Combining all ingredients, this decay width is of order $\lambda_x^4 x^2$. Since it is only the total decay width that is constrained to be small by the demand that the SM reheating temperature lie in the required window, this fixes only a product of λ_x and x. If $x \sim v$, then the mixed decay to hH is effectively second order in the small coupling λ_x^2 and can be neglected relative to the other partial widths. Conversely if $x \ll v$, then λ_x is much larger and this decay cannot be neglected. In what follows we will work in the region of parameter space where mixed decays to hH are negligible.

The rate of heat flow into each sector may be well approximated by adding the decay rates of X into each channel and weighting these by the fraction of energy transferred into the particular sector. Of course, when X decays into Higgs particles, these in turn decay out of equilibrium into both the Standard Model and twin sectors. As the Higgs decays are almost instantaneous, the fraction of energy transferred into each sector is simply that carried by the Higgs decay products multiplied by their branching fractions

for each sector. The total rate at which X particles are transferred into the SM plasma is

$$
\begin{aligned}
W(X \to SM) \approx{} & \Gamma(X \to SM) + \Gamma(X \to hh)Br(h \to SM) \\
& + \Gamma(X \to HH)(Br(H \to SM) + Br(H \to hh)Br(h \to SM)). \quad (5.33)
\end{aligned}
$$

The corresponding rate for energy deposition into the twin sectors is simply given by the replacement of $SM \mapsto$ Twin. The first term is the rate at which X decays directly into the SM through mass mixing with the Higgs. The second is the fraction of X energy that is transferred into lighter Higgs states that subsequently decay into the SM. The third is the analogous term for decays into the heavy Higgs, where cascade decays of the H into the h and subsequently other SM particles must be included. Note that decays of the heavy Higgs into the light Higgs make up a majority of decay width, because of the large quartic coupling required for the twin Higgs potential.

Below the hh threshold, it is possible for X to decay via one on-shell and one off-shell Higgs boson. The partial width for off-shell Higgs production was calculated for $X \to hh^* \to hb\bar{b}$ and found to be negligible compared to two-body decays through mass mixing and so we omit three-body decay widths in what follows.

Ultimately, the complete partial widths for the decay of X into the Standard Model and twin sectors includes the sum of decays into Higgs bosons h and H and direct decays into the fermions and gauge bosons of the two sectors. We compute the latter to an intended level of accuracy of $\sim 10\%$ (including, e.g., NLO QCD corrections to decays into light-flavor quarks), mostly following [366]. The resulting partial widths into the Standard Model and twin sectors are shown as a function of m_X in Figure 5.5 with the ratio of branching fractions displayed in Figure 5.6.

Over much of the space below the Higgs mass, the branching ratio exhibits the ex-
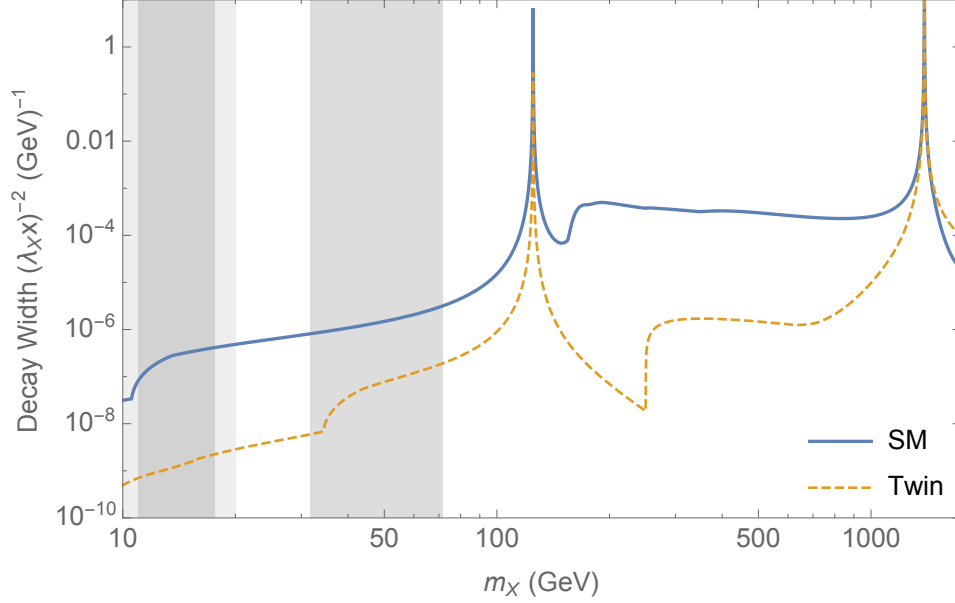
Figure 5.5: The partial widths of the X into the SM (solid blue line) and twin sector (dashed orange) for $f/v = 3$ in units of $(\lambda_x x)^2$. The light gray bands indicate regions of QCD-related uncertainty in the SM calculation, while the darker gray bands indicate the corresponding regions of uncertainty for the twin calculation.

pected $(f/v)^2$ scaling from the mass mixing. Below ~ 40 GeV, suppression of the twin partial width arises because the twin bottom quark pair production threshold is crossed. As m_X nears m_h, the SM branching fraction grows by ~ 4 orders of magnitude as the WW^*, ZZ^*, and then WW and ZZ decays go above threshold. Since the analogous thresholds are at much higher energies in the twin sector, the enhancement is not paralleled by decays into the twin sector until m_X is close to the twin scale. There is therefore a large range of masses $m_h \lesssim m_X \lesssim m_H$ over which the SM branching fraction dominates by several orders of magnitude.

Above the $X \to hh$ threshold, the ratio of decay widths is roughly constant in mass up to the HH threshold. The twin sector decay rate is dominated by decays of on-shell light Higgs into twin states, $\Gamma(X \to \text{Twin}) \approx \Gamma(X \to hh)\text{Br}(h \to \text{Twin}) \propto 1/m_X$ as in (5.31). If the SM were also predominantly reheated through this channel, then the
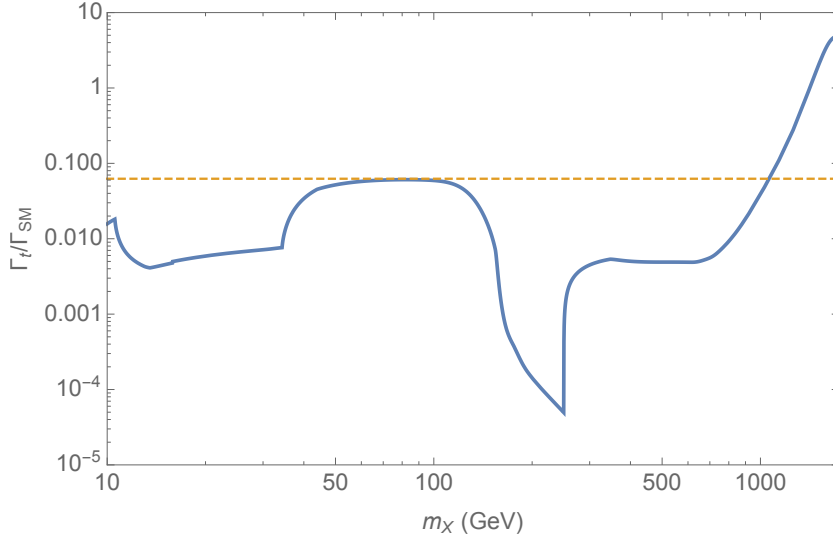
Figure 5.6: The ratio of branching fractions of the X into the SM and twin sectors at $f/v = 3$. The dashed line gives the expected $(v/f)^2$ scaling from the mass mixing; deviations are due to various mass threshold effects.

ratio of branching fractions would again be approximately $\delta_{hA}^2/\delta_{hB}^2 \approx (f/v)^2$. However, the SM decay width also receives a larger contribution from decays through mass mixing between the X and the Higgs gauge eigenstates.

For masses $m_X > 2m_h$, decays through mass mixing are dominated by the SM WW and ZZ channels. In this mass region, the decay rate of a Higgs into longitudinally polarized vector bosons scales as $\Gamma(h \to WW, ZZ) \sim m_X^3$, but the mixing angle scales as $\delta_{AX}^2 \sim 1/m_X^4$ (as in (5.28)), resulting in the same $\sim 1/m_X$ scaling and thus a roughly constant ratio in this range of masses. Near $m_X \sim 1$ TeV, decays into twin vector bosons through mass mixing begin to dominate, and there is no favourable asymmetry in the branching fractions, as discussed in this section. Even at higher masses, the effects of heavy Higgs decays into light Higgs do not compensate sufficiently, as this partial width scales with m_X in the same way as the partial width for longitudinally polarised weak bosons.

The constraint on the decay width from the required reheating temperature (5.27)

186

translates into a constraint on the size of the coupling $\lambda_x x$. For $m_X \gtrsim m_h$, this gives $10^{-8.5}\,\text{GeV} \lesssim \lambda_x x \lesssim 10^{-6}\,\text{GeV}$, while for lower masses, this range increases to $10^{-7}\,\text{GeV} \lesssim \lambda_x x \lesssim 10^{-5.5}\,\text{GeV}$ at $m_X \sim 20\,\text{GeV}$.

The gray bands in Figure 5.5 highlight regions where our analytic estimates of the partial widths encounter enhanced uncertainties arising from the bottom and charm thresholds in both sectors. Over most of these ranges, we estimate the size of these uncertainties to be either $\sim 10\%$ or confined to very small subregions. The thicknesses of these bands have been chosen conservatively, and ultimately the branching ratios should be accurate to within a factor of $\pm\Lambda_{QCD}$ of the bottom and charm mass thresholds. In particular, the prescription of [367] has been followed for approximating the bottom partial width close to the open flavour threshold. Resonant decay into gluons from bottomonia mixing has been neglected, although these resonant mass ranges are expected to be only $\sim$ MeV wide at the CP-even, spin-0 bottomonia masses $m_X = m_{\chi_{b_i}}$ (see [367] and [368]). It should be noted, however, that at temperatures above that of the QCD phase transition, the quark decay products behave differently compared to that expected in a low temperature environment. In particular, for hot enough temperatures, the b or c quarks may not hadronise and the partonic partial widths may more reliable. The applicability of the treatment of the flavour thresholds used here may therefore not be valid if the decay occurs in the hot early universe. However, it is only very close to the threshold itself (within several GeV) that this uncertainty becomes significant. Finally, quark masses have been neglected in the gluon partial width. For m_X close to the flavour thresholds, this approximation breaks down, but the gluon branching fraction is only $\sim 10\%$ and so the error does not contribute to the uncertainty of the total width by more than this order (it is this uncertainty that is responsible for most of the extension of the length of the gray bands about the flavour threshold).

Close to the charm threshold, the analogous uncertainties are even more poorly under-

stood. Below the charm threshold, hadronic decays of a light scalar are highly uncertain (see [369] for discussion). We avoid these regions altogether by restricting our consider-ations to m_X roughly above the twin charm threshold. Note that below the SM charm threshold, the smaller decay rate of a Higgs-like scalar necessitates larger couplings $\lambda_X x$ for X to have a lifetime within the required reheating window. The larger couplings then imply potentially stronger constraints from invisible mesonic decays. See [368, 369, 370] for further discussion and recent analysis of the pertinent experimental constraints.

Taken together, the results in Figures 5.5 and 5.6 bear out the expectation that a scalar X with symmetric couplings to the Standard Model and twin sectors may nonethe-less inherit a large asymmetry in partial widths from the hierarchy between the scales v and f. Across a wide range of masses m_X, the asymmetry is proportional to (or greater than) v^2/f^2, tying the reheating of the two sectors to the hierarchy of scales.

Before proceeding to our computation of cosmological observables, we comment on an alternative variation on the reheating mechanism presented here that involves having X odd under the twin parity. This permits two renormalisable interactions with the Higgses to give a Higgs potential of the form:

$$\mathcal{V} \supset m_0^2 \left(|H_A|^2 + |H_B|^2 \right) + \lambda_0 \left(|H_A|^4 + |H_B|^4 \right) + \epsilon X^2 \left(|H_A|^2 + |H_B|^2 \right) + \tilde{\epsilon} X \left(|H_A|^2 - |H_B|^2 \right) \tag{5.34}$$

If X then acquires a vev at some scale, it may be possible to arrange for the resulting spontaneous breaking of the $\mathbb{Z}_2$ to give that required in the Higgs potential. However, we find that, in order for X to be long-lived and reheat the universe, its couplings to the Higgs must be highly suppressed and therefore that the resulting vev of X required to explain the soft $\mathbb{Z}_2$-breaking in the Higgs potential must be many orders of magnitude above the twin scale. If this is to be identified with the characteristic mass scale of X, then a UV-completion of the twin Higgs is required for anything further to be said of the

prospects of this possibility. However, if such a UV completion has similar structure to the couplings in (5.34), then asymmetric reheating may require a cancellation between the odd and even couplings of X to the Higgs potential in order to suppress its twin-sector branching fraction (because the odd coupling appears with opposite signs in the coupling between X and the h_A and h_B states). We do not consider this possibility further.

Imprints on the CMB

For appropriate values of m_X, the out-of-equilibrium decay of X reheats the two sectors to different temperatures and effectively dilutes the energy density in the twin sector. We obtain an analytic estimate of the effects of the X decay on the number of light degrees of freedom observed from the CMB by approximating both the decay of X and the decoupling of species as instantaneous in Section 5.2.3. We then demonstrate that this estimate is reliable over most of the parameter space of interest with a numerical calculation in Section 5.2.3. In Section 5.2.3 we consider neutrino masses and their joint constraints with N_{eff}.

Analytic estimate of N_{eff}   If X dominates the energy density of the universe and then decays, depositing energy ρ_{SM} and ρ_{t} into the SM and twin sectors respectively, then the temperature ratio is determined by

$$\frac{\rho_{\text{t}}}{\rho_{\text{SM}}} = \frac{g_\star^{\text{t}}(T_{\text{reheat}}^{\text{t}})}{g_\star^{\text{SM}}(T_{\text{reheat}}^{\text{SM}})} \left(\frac{T_{\text{reheat}}^{\text{t}}}{T_{\text{reheat}}^{\text{SM}}}\right)^4 \approx \frac{\Gamma(X \to \text{Twin})}{\Gamma(X \to \text{SM})}, \qquad (5.35)$$

where $T_{\text{reheat}}^{\text{SM}}$ and $T_{\text{reheat}}^{\text{t}}$ are the reheating temperatures for each sector, while $g_\star^{\text{SM}}$ and $g_\star^{\text{t}}$ are the SM and twin effective number of relativistic degrees of freedom, respectively. We have assumed that the two sectors are cool enough that they have already decoupled. We point out that not only does the number of effective degrees of freedom in each sector

need to be evaluated at the temperature of that sector, but that $g_\star^{\mathrm{t}}$ and $g_\star^{\mathrm{SM}}$ differ as functions of temperature due to the differences in the spectra of the sectors, as seen in Figure 5.1. As is well-known [364], reheating is a protracted process that occurs over a time-scale given by the lifetime of the reheaton. During this time, the temperature of the plasma cools slowly because, while the energy is being replenished by the decay of the reheaton, it is simultaneously diluted and redshifted with the expansion of the universe. It is assumed in (5.35) that any primordial energy density in either sector is subdominant.

The temperatures of both sectors then redshift in the same way, so the only additional differences between their temperatures arise from changes to the effective number of degrees of freedom in each sector. By conservation of comoving entropy within each sector, each evolves as $T_{eq}^i/T_{\mathrm{reheat}}^i = \left(g_\star^i(T_{\mathrm{reheat}}^i)/g_\star^i(T_{eq}^i)\right)^{1/3} a(T_{\mathrm{reheat}})/a(T_{eq})$ where T_{eq}^i is the temperature of the sector at matter-radiation equality, which the CMB probes as explained in Section 5.2.2, and $a(T)$ is the scale factor as a function of temperature. In the mirror Twin Higgs model, the two sectors have the same number of light degrees of freedom at recombination (three neutrinos and a photon, assuming that the neutrinos are still relativistic), so

$$\left(\frac{T_{eq}^{\mathrm{t}}}{T_{eq}^{\mathrm{SM}}}\right)^4 = \left(\frac{T_{\mathrm{reheat}}^{\mathrm{t}}}{T_{\mathrm{reheat}}^{\mathrm{SM}}}\right)^4 \left(\frac{g_\star^{\mathrm{t}}(T_{\mathrm{reheat}}^{\mathrm{t}})}{g_\star^{\mathrm{SM}}(T_{\mathrm{reheat}})}\right)^{4/3} = \frac{\Gamma(X \to \mathrm{Twin})}{\Gamma(X \to \mathrm{SM})} \left(\frac{g_\star^{\mathrm{t}}(T_{\mathrm{reheat}}^{\mathrm{t}})}{g_\star^{\mathrm{SM}}(T_{\mathrm{reheat}})}\right)^{1/3} . \quad (5.36)$$

As our range of reheat temperatures encompasses the QCD phase transitions of both sectors, the factors of $g_\star$ can be important.

Given the temperatures of the two sectors after X decays, we can obtain a simple estimate of the contribution to N_{eff} that neglects the impact of masses of the twin neutrinos

discussed in Section (5.2.2),

$$(\Delta N_{\text{eff}})_{m_\nu=0} = \frac{4}{7}\left(\frac{11}{4}\right)^{4/3} g_\star^{\text{SM}}(T_{eq}^{\text{SM}}) \frac{\rho_{\text{t}}(T_{eq}^{\text{t}})}{\rho_{\text{SM}}(T_{eq}^{\text{SM}})} \tag{5.37}$$

$$\approx 7.4 \times \frac{\text{Br}(X \to \text{Twin})}{\text{Br}(X \to \text{SM})} \left(\frac{g_\star^{\text{t}}(T_{\text{reheat}}^{\text{t}})}{g_\star^{\text{SM}}(T_{\text{reheat}}^{\text{SM}})}\right)^{1/3}. \tag{5.38}$$

In this limit the most recent Planck data give a 2σ bound of $\Delta N_{\text{eff}} \lesssim 0.40$ assuming pure $\Lambda\text{CDM}+N_{\text{eff}}$ [318]. This translates into the requirement $\frac{\rho_{\text{t}}(T_{eq}^{\text{t}})}{\rho_{\text{SM}}(T_{eq}^{\text{SM}})} \approx \frac{\Gamma(X \to \text{Twin})}{\Gamma(X \to \text{SM})} \lesssim 0.05$, ignoring possible differences in $g_\star$.

Of course, as discussed in Section 5.2.2, the twin neutrino masses are relevant at the temperature of matter-radiation equality, so we can obtain a more meaningful estimate of ΔN_{eff} using the results of Section 5.2.2 evaluated at the twin temperature determined above:

$$\Delta N_{\text{eff}} = \left(\frac{11}{4}\right)^{4/3} \frac{120}{7\pi^2 \left(T_{eq}^{\text{SM}}\right)^4} \left(\rho_\gamma^{\text{t}}\left(T_{eq}^{\text{t}}\right) + \sum_\alpha 3 w_{\nu_\alpha}^{\text{t}}\left(T_{eq}^{\text{t}}\right)\rho_{\nu_\alpha}^{\text{t}}\left(T_{eq}^{\text{t}}\right)\right) \tag{5.39}$$

$$T_{eq}^{\text{t}} = T_{eq}^{\text{SM}} \left(\frac{\Gamma(X \to \text{Twin})}{\Gamma(X \to \text{SM})}\right)^{1/4} \left(\frac{g_\star^{\text{t}}(T_{\text{reheat}}^{\text{t}})}{g_\star^{\text{SM}}(T_{\text{reheat}}^{\text{SM}})}\right)^{1/12} \tag{5.40}$$

with $T_{eq}^{\text{SM}} \approx 0.77$ eV [318] the photon temperature. While the right-hand side of this equality has implicit dependence on T_{eq}^{t} through $g_\star^{\text{t}}$, this is only important if the reheating occurs between the SM and twin QCDPTs and the neglecting of the factors of $g_\star$ is otherwise reliable. With the further inclusion of Standard Model neutrino masses or an extra sterile neutrino, the bound described above weakens to $\Delta N_{\text{eff}} \lesssim 0.7$. As discussed in Section 5.2.2, we are not aware of any analyses specific to our model involving both pure dark radiation and three sterile neutrinos with masses of order the photon decoupling temperature of the CMB and possibly cooler temperatures. In the absence of such an analysis, we use the inequality $\Delta N_{\text{eff}} \lesssim 0.7$ to indicate where the present CMB

measurements are likely to constrain the light degrees of freedom of this model, leaving

a more detailed analysis of the CMB constraints as future work. In this case, the bound

on the decay width ratio is $\frac{\Gamma(X \to \text{Twin})}{\Gamma(X \to \text{SM})} \lesssim 0.09$. The next generation of CMB experiments

are projected to strengthen this constraint to $\Delta N_{\text{eff}} \lesssim 0.02$ at the 1σ level [371].

Numerical Calculation of N_{eff}   A more precise study of the effect of X decay on the

number of effective neutrino species at recombination may be performed by numerically

solving a system of differential equations for the entropy in X and the two sectors as a

function of time. Following the analysis of Chapter 5.3 of [364] we have

$$H = \frac{1}{a}\frac{da}{dt} = \sqrt{\frac{1}{3M_{Pl}^2}(\rho_X + \rho_{SM} + \rho_t)} \tag{5.41}$$

$$\frac{d\rho_X}{dt} + 3H\rho_X = -\Gamma_X \rho_X \tag{5.42}$$

$$\rho_i = \frac{3}{4}\left(\frac{45}{2\pi^2 g_\star^i}\right)^{1/3} S_i^{4/3} a^{-4} \tag{5.43}$$

$$S_i^{1/3}\frac{dS_i}{dt} = \left(\frac{2\pi^2 g_\star^i}{45}\right)^{1/3} a^4 \left(\rho_X \Gamma_{X \to i} + \frac{dq_{j \to i}}{dt}\right), \tag{5.44}$$

where S_i are comoving entropy densities and it has been assumed that X is cold by the

time it decays so that $\rho_X = m_X n_X$ with number density n_X (this is reliable as we only

consider $m_X > 10$ GeV, which is above the decoupling temperature of ~ 1 GeV). The

rate of heat flow from sector j to i per proper volume, $\frac{dq_{j \to j}}{dt}$, is defined in (5.6). To

account for the temperature-dependence of the effective number of relativistic degrees of

freedom in each sector, these equations are solved iteratively in the profiles of $g_\star^i(T^i)$.

The equations are solved in three stages: before, during and after the decoupling of

the SM and twin sectors. The ratio f/v is fixed to 4 for this analysis. Initial conditions

were chosen with $\rho = 10^{-12}\rho_X$, for combined SM and twin energy densities ρ. However,

it is only the requirement that the initial energy density of X dominates over that of

the SM and twin sectors that is important for simulating the cosmology over the times of interest here, as the entirety of the latter is then generated by the subsequent decay. The results close to the decoupling and reheating epochs are otherwise insensitive to the initial conditions and ultimately match onto the standard outcome [364] expected by equating the Hubble rate with the decay rate of X. The sectors are assumed to be in thermal equilibrium and sharing entropy until a temperature of 10 GeV, below which they are evolved separately with the heat flows $\frac{dq_{i \to j}}{dt}$ switched on. Elastic scatterings were neglected from the heat flow rate to accelerate the computation. It was verified for the results found below that their contribution to the heat flow was always $\lesssim 10\%$ while the heat flow was itself not dominated by the Hubble rate. Heat flow was switched off again once the twin temperature reaches 0.1 GeV, by which time thermal decoupling is long-since complete, and the sectors are subsequently evolved separately. Again, although the strengthening of the colour force and the QCDPT make the perturbative tree-level computation of the scattering rates unreliable at temperatures below ~ 1 GeV, as found in Section 5.2.2 and also in the results below, the sectors decouple above these temperatures. Notably, the impact of X on the expansion rate causes decoupling to occur at slightly hotter temperatures than expected from the analysis of Section 5.2.2 for the decoupling in the standard cosmology.

The ratio of energy densities in each sector determines N_{eff}, from (5.39). A plot of this ratio over time is shown in Figure 5.7, with the expectation under the approximations of the previous section shown as well. This approximation is reliable as long as the lifetime of X is much longer than the temperature at which decoupling concludes, here ~ 1 GeV. The larger asymptotic value of the ratio of the blue line arises because the lifetime lies close to the decoupling period, so that a significant fraction of the energy is transferred while the sectors are thermalised or partially thermalised and does not contribute toward asymmetric reheating. Equivalently, as will be discussed below, insufficient time elapses
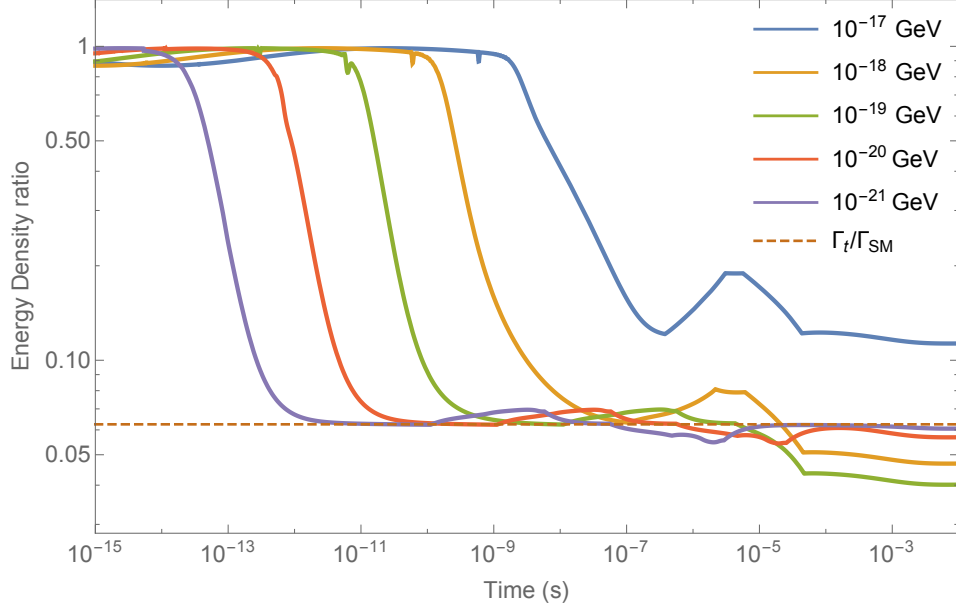
Figure 5.7: Ratio of twin to SM energy densities throughout decoupling and reheating, for different decay rates Γ_X. The dashed line corresponds to the prediction of from the ratio of decay widths, here selected to be $1/16$.

between decoupling and reheating for the twin energy density to dilute and be repopulated by the decays to the level predicted by (5.35). The subsequent bump represents the period between the reheating of the twin sector by its QCD phase transition followed by that of the SM. The green and orange lines correspond to reheating temperatures that lie between SM and twin QCD phase transitions. In these cases, the reheating of the SM from the subsequent SM QCD phase transition raises its energy density relative to the twin sector above that expected from the ratio of branching fractions. As this occurs after the lifetime of the reheaton, the estimate of the reheating temperatures presented in (5.36) is still good as subsequent changes in the ratio due to the evolution of $g_\star$ are accounted for in our analysis of the reheating scenarios.

The steep drop in the energy density ratio corresponds to the brief period during which the energy density of the twin sector present at decoupling dilutes and redshifts, which continues until it reaches a comparable size to the energy density that is being

replenished by reheating. If the twin-sector branching fraction is highly suppressed, as can occur in the "valley" region in Figure 5.6 with $m_h \lesssim m_X \lesssim 2m_h$, then a longer time is required for this to happen, especially close to the decay epoch where the diminishing of the X population also contributes to a reduced reheating rate. These effects can prolong the time required for the energy density ratio to converge to the asymptotic prediction of (5.35).

Contour plots of ΔN_{eff} as a function of m_X and f/v appear in Figure 5.8, along with current and predicted bounds using the analytic results of Section 5.2.3. The minimum neutrino mass configuration with Dirac masses has also been assumed, although the results are relatively insensitive to this provided that the twin neutrino masses are not well above the eV scale. A SM reheating temperature of 0.7 GeV has been assumed. At this temperature, we have verified using the numerical calculation of Section 5.2.3 that the twin sector reheating temperature is always roughly above the twin neutrino decoupling temperature over the parameter space of the figure, ensuring that the neutrinos thermalise once produced in the decays and hence that the predictions of Section 5.2.3 are valid. A treatment of the case in which the twin neutrinos are produced below their decoupling temperature is beyond the scope of this analysis, but would involve the computation of the phase space spectrum of the neutrino decay products of the X.

Also, as discussed in Section 5.2.2, a large temperature difference may partially relax back if reheating occurs close to sector decoupling. However, a reliable calculation of the heat flow at the temperatures of interest here must incorporate non-perturbative effects. We do not perform such a computation, but note that, at a slightly higher SM reheating temperature of 2 GeV where this computation is more reliable, ΔN_{eff} in Figure 5.8 can be raised by up to an order of magnitude in the region with $f/v \lesssim 4$ and $150\,\text{GeV} \lesssim m_X \lesssim 200\,\text{GeV}$, notably where the twin sector partial width is suppressed relative to the SM by several orders of magnitude. The resulting ΔN_{eff} prediction is, nevertheless, still

out of observable reach. At the lower SM reheating temperature assumed in Figure 5.8, it is expected that decoupling will be further advanced and the enhancement in ΔN_{eff} would be weaker.

We emphasize that, if the lifetime of X is sufficiently close to the time of decoupling, or equivalently, that the reheating temperature is sufficiently close to the decoupling temperature, then the residual twin energy density left-over may be comparable to or greater than that regenerated by reheating. Consequently, the suppression in ΔN_{eff} would be less than that predicted in (5.36). In this respect, the projection of Figure 5.8 should be regarded as a lower bound on ΔN_{eff}. In the regions of high suppression, such as the "valley" region, the full asymmetry may not be generated before the complete decay of X when the reheating temperature is of similar order as the decoupling temperature. In particular, for the reheating temperature chosen here of 0.7 GeV and branching fraction $\text{Br}(X \to \text{Twin}) \sim 10^{-5}$, the numerical calculation of the energy density ratio saturates at $\sim 4 \times 10^{-5}$. We do not include this effect in Figure 5.8 as its only impact is to mildly shift the unobservably small $\Delta N_{\text{eff}} = 10^{-4}$ contour. Lower reheating temperatures would agree with the prediction of (5.35) were it not for the caveat that the twin neutrinos may be produced out of equilibrium. However, this minimum value at which ΔN_{eff} is saturated can grow significantly with hotter reheating temperatures upon which it is highly dependent.

CMB-S4 observations will be able to probe a large portion of the most natural parameter space, save the region $m_h \lesssim m_X \lesssim 2m_h$ where decays into the Standard Model dominate well beyond the ratio f^2/v^2, as previously discussed. Significantly, precision Higgs coupling measurements at the LHC are unlikely to probe the mirror Twin Higgs model beyond $f \sim 4v$, so that the observation of additional dark radiation may be the *first* signature of a mirror Twin Higgs.
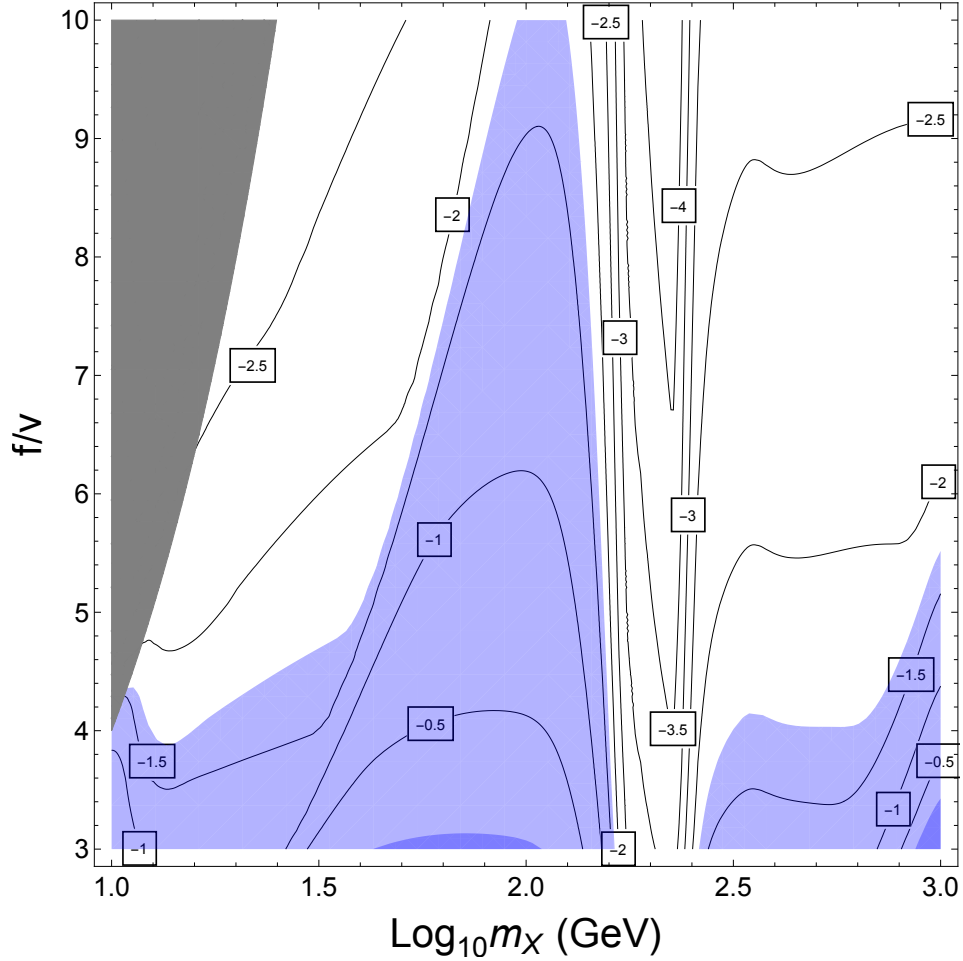
Figure 5.8: Contours of $\log_{10} \Delta N_{\mathrm{eff}}$ as a function of m_X and f/v, for $T_{\mathrm{reheat}}^{\mathrm{SM}} = 0.7$ GeV. The dark blue region is in tension with Planck, while the light blue region will be tested by CMB-S4. Gray regions are where the X mass is below the twin charm threshold and our calculation of the twin sector partial width is unreliable.

Neutrino Masses   In addition to the bounds on N_{eff}, we must also respect the bounds on neutrino masses. The analysis remains nearly the same as in Section 5.2.2, but now with the twin neutrinos at a lower temperature, as determined above. As mentioned above, for large enough f/v and SM reheating temperature sufficiently close to the lower bound, the reheating temperature of the twin sector may be below the twin neutrino decoupling temperature and the resulting energy density would be more difficult to compute. For simplicity, we choose $\lambda_x x$ large enough such that the twin reheating

temperature is always above the twin neutrino decoupling temperature.

As before, we compute m_ν^{eff} as

$$m_\nu^{\text{eff}} = \frac{n_\nu^{\text{t}}}{n_\nu^{\text{SM}}} \sum_\alpha m_{\nu_\alpha}^{\text{t}}. \tag{5.45}$$

In relating the scale factors at neutrino decoupling in each sector, we now have to use the above temperature ratio to find, analogously to Section 5.2.2, that

$$m_\nu^{\text{eff}} = \left(\frac{\Gamma_{\text{t}}}{\Gamma_{\text{SM}}}\right)^{3/4} \left(\frac{g_\star^{\text{t}}\left(T_{\text{reheat}}^{\text{t}}\right)}{g_\star^{\text{SM}}\left(T_{\text{reheat}}^{\text{SM}}\right)}\right)^{1/4} \left(\frac{f}{v}\right)^n \sum_\alpha m_{\nu_\alpha}^{\text{SM}}, \tag{5.46}$$

where, again, $n = 1$ for Dirac masses and $n = 2$ for Majorana masses. Interestingly, if the branching ratios scale as $\Gamma_{\text{t}}/\Gamma_{\text{SM}} = (v/f)^2$, then we have $m_\nu^{\text{eff}} \propto (f/v)^{-3/2+n}$, so the contribution grows with f/v for Majorana masses, but is suppressed for Dirac masses.

As before, we consider the minimal mass spectrum of $m_\nu = [0.0, 0.009, 0.06 \text{ eV}]$ and a degenerate spectrum of $m_\nu = [0.1 \text{ eV}, 0.1 \text{ eV}, 0.1 \text{ eV}]/3$. In Figure 5.9 we plot the predictions of the X reheating for ΔN_{eff} and m_ν^{eff} for both spectra and both Dirac and Majorana masses using the approximations of Section 5.2.2, for f/v from 3 to 10 and assuming the $\frac{\Gamma_{\text{t}}}{\Gamma_{\text{SM}}} \sim (v/f)^2$ scaling; there are regions in the space of m_X where the suppression of m_ν^{eff} would be much higher.

Dashed lines indicate the rough locations of present experimental limits from Planck 2015, and projected bounds from CMB-S4. As mentioned in Section 5.2.2, we are unaware of any study of bounds on both m_ν^{eff} and ΔN_{eff} treated jointly. In the absence of this, we show present and projected constraints on N_{eff} and $\sum m_\nu$ from [372] and [350], ignoring correlations, as described in Section 5.2.2.
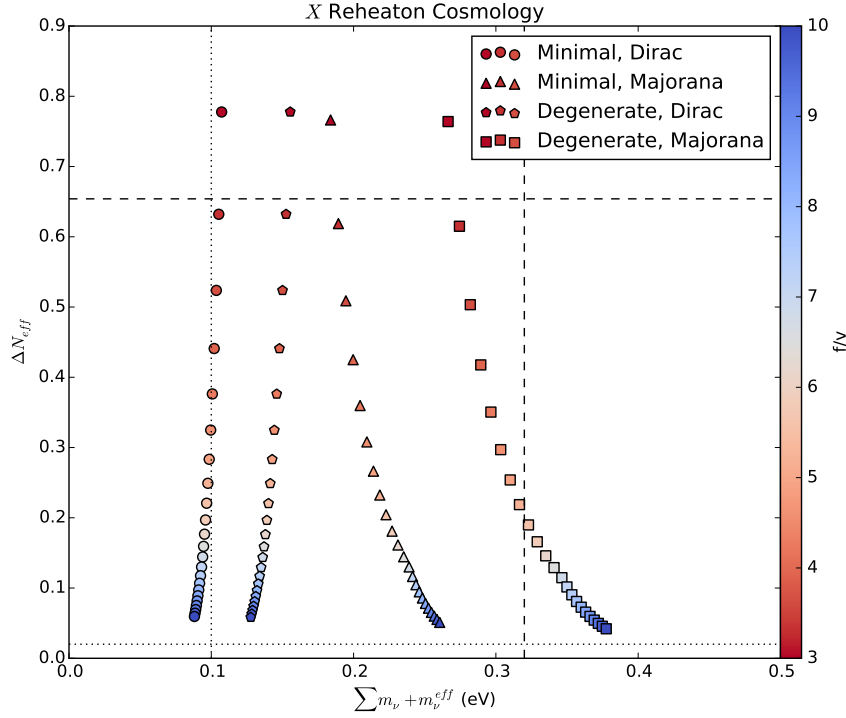
Figure 5.9: Predicted values of ΔN_{eff} and $\sum m_\nu + m_\nu^{\text{eff}}$ for minimal and degenerate neutrino mass spectra with both Dirac and Majorana masses for f/v from 3 to 10. The Planck 2015 [318] bounds on $\sum m_\nu$ and N_{eff}, as discussed in Section 5.2.2, are represented by the dashed lines, and the projected CMB-S4 constraints are given by the dotted lines. It has been assumed that $\frac{\Gamma_{\text{t}}}{\Gamma_{\text{SM}}} \sim (v/f)^2$. Note however, that, from Figure 5.8, this scaling of the partial widths holds only for the mass range $50\,\text{GeV} \lesssim m_X \lesssim 120\,\text{GeV}$, outside of which the twin partial width is more suppressed and the model is only testable through ΔN_{eff} over a smaller range in f/v.

Thermal Production

In our discussion up to this point, we have been agnostic about the origin of the cosmic abundance of X and have operated under the assumption that it absolutely dominates the cosmology before it decays. Here, we consider the possibility that X was thermally produced through freeze-out and subsequently dominates the universe as a relic before decaying. This thermal history is viable, but places strong constraints on the mass and couplings of the X.

The energy density of relativistic species redshifts as $\rho_r \propto a^{-4} \propto T^4$, while the energy density of non-relativistic, chemically decoupled matter scales as $\rho_m \propto a^{-3}$. The energy density contained in the X can therefore only grow relative to the energy density in the thermal bath once it becomes non-relativistic. We found in Section 5.2.3 that by recombination, $\rho_t/\rho_{\mathrm{SM}} \lesssim 0.09$ is needed to evade current bounds on ΔN_{eff}. Thus we need to have the energy density in the X dominate over the SM and twin plasmas by more than this factor when it decays. If X becomes non-relativistic instantaneously at the moment that its temperature reaches some fraction $c \sim \mathcal{O}(0.1)$ of its mass, then, as $T \propto 1/a$ and ρ_X is $\sim 1/g_\star$ of the total energy density, the mass is required to satisfy $m_X \gtrsim 10/c \times g_\star\,(T = m_X)\,T_{X\mathrm{reheat}}^{\mathrm{SM}}$. Since the SM reheating temperature is strongly constrained to be above BBN, this effectively puts a lower limit on the mass of the X. Importantly, X must freeze-out when relativistic or its energy density will be further Boltzmann suppressed. The lower limit on the mass of the X becomes an upper limit on the X's couplings - if it couples too strongly to the thermal bath, then it won't freeze out early enough to be hot.

In fact the situation is somewhat less favorable than the above analysis suggests, because it is relevant operators that must keep X in thermal equilibrium. For an X with the interactions introduced in Section 5.2.3, the annihilations have rates that scale with temperature as $\Gamma \sim n_X \langle \sigma v \rangle \sim T$ for $T \gtrsim m_X, m_h$ (where n_X is the number density of X and $\langle \sigma v \rangle$ is its thermally averaged annihilation cross section). However, in a radiation-dominated universe, $H \sim T^2$. Thus, at high enough temperatures, X is not in thermal equilibrium with the plasma and it is only once the universe cools enough that it may thermalise. Then, as the temperature drops, $XX \to q\bar{q}$ annihilations become suppressed by the Higgs mass and subsequent Boltzmann suppression causes X to freeze-out. Note that the rates of these annihilation processes are controlled by the coupling λ_x, independently of x, which is unconstrained by itself (other processes mediated by

$\lambda_x x$ are found to be subdominant in the ensuing analysis, for the range of λ_x over which thermal production is successful). If the coupling is too weak to begin with, then the X never thermalises and thermal production cannot happen. Thermal production therefore requires a careful balancing of parameters - small coupling λ_x is preferred for X to freeze-out hot and as early as possible, but the coupling is bounded from below by the requirement that X reach thermal equilibrium. This combination of constraints severely restricts the size of the parameter space over which thermal production is viable to cases in which the coupling is selected so that X enters and departs from thermal equilibrium at close to the same temperature.

To obtain numerical predictions for this scenario, the calculation of Section 5.2.3 was modified to account for the time after the freeze-out of X before it becomes non-relativistic. During this period we use (5.15) and (5.16) for the energy density of the X, approximating decays as being negligible, before switching over to (5.42) when the temperature drops below the mass of the X. The approximation that the X does not decay appreciably while it is relativistic must be good if there is to be sufficient time for it to grow to dominate between becoming non-relativistic and decaying. The decay width of X was fixed to 5×10^{-21} GeV, corresponding to a reheating temperature close to the ~ 10 MeV lower limit, in order to maximise the amount of time over which the energy density of X may grow relative to the SM plasma, thereby providing the greatest possible reheating.

The predictions for ΔN_{eff} from a thermally produced X are shown in Figure 5.10 for the small regions of parameter space where this is viable, with $f/v = 4$. We find that the dominant annihilation channels over this region are $XX \to t\bar{t}$ and $XX \to b\bar{b}$, mediated by the light Higgs, as well as their twin analogues, mediated by the heavy Higgs. As expected, the primordial energy density in the twin sector is too large compared to that generated by the X for the asymmetric reheating to be effective when m_X is too light

($\lesssim$ 100 GeV in this case). Similarly, when the coupling is too strong, the X is held in equilibrium for longer and freezes-out underabundant compared to the twin energy density. However, when the coupling is too weak (the gray region), X never thermalises to begin with (close to the boundary with this region, X freezes-out almost immediately after thermalising). The peak in the contours occurs because of the "H-funnel" in which the twin Higgs resonantly enhances annihilations into twin quarks. All of this region will be testable by CMB-S4.

5.2.4   Twinflation

As an alternative to the model presented above of late, out-of-equilibrium decays of a $\mathbb{Z}_2$-symmetric scalar, one may imagine that the field driving primordial inflation reheats only the Standard Model to below the decoupling temperature of the two sectors. Production of the twin particles then ceases at some time after the temperature drops below the decoupling temperature during reheating.

To make this consistent with a softly-broken $\mathbb{Z}_2$ symmetry, we extend the inflationary sector and introduce a 'twinflaton' that couples solely to the twin sector. The combined inflationary and twinflationary sectors respect the $\mathbb{Z}_2$ symmetry. However, if the two sectors are entirely symmetric then one generally expects both inflationary dynamics to happen coincidentally, which would result in identical reheating. We therefore rely on soft $\mathbb{Z}_2$-breaking to give an asymmetry between the two sectors that causes the twinflationary sector to dominate the universe first. With the right arrangement we can end up with two distinct periods of inflation - a first caused by the "twinflaton" and a second that then reheats the Standard Model to below the decoupling temperature, having diluted the sources of twin-sector reheating from the first period.

One simple mechanism for $\mathbb{Z}_2$-breaking which is well-suited for introducing asymme-
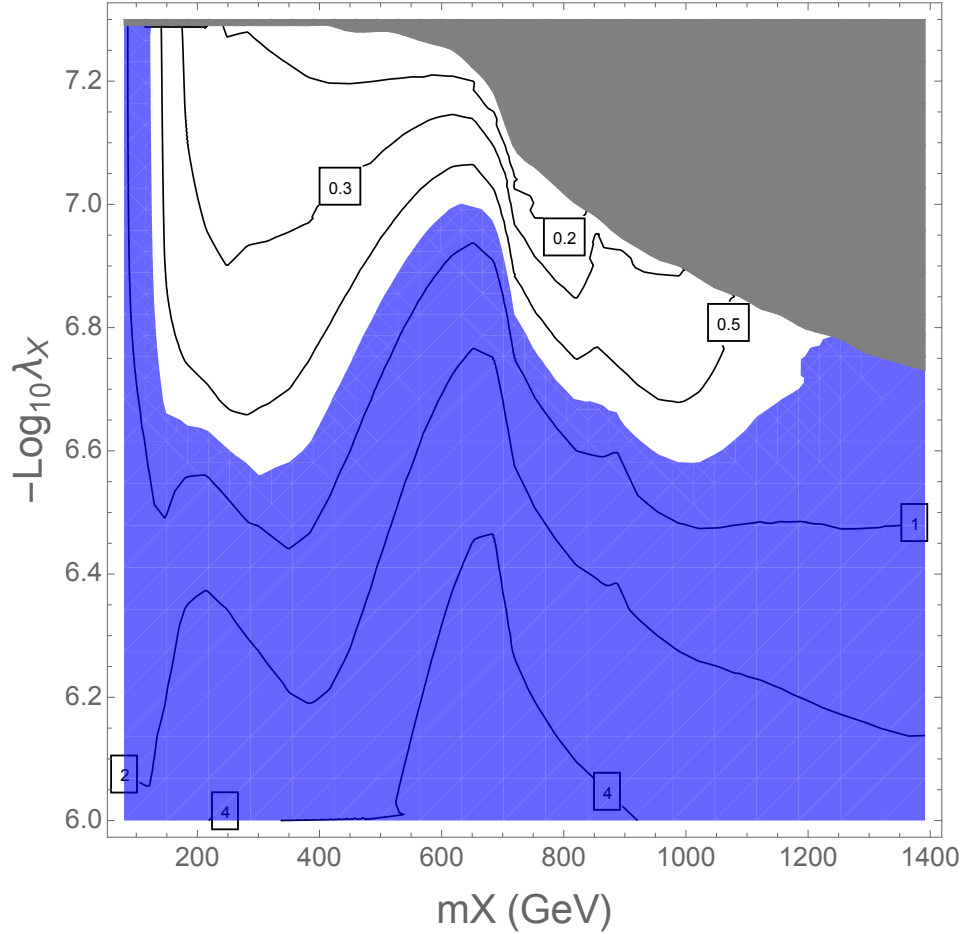
Figure 5.10: Parameter space where thermal production of X gives a large enough relic abundance to dilute the twin sector, for $f/v = 4$. In the gray region, the coupling is too weak for X to ever reach thermal equilibrium. The blue region is in tension with recent Planck measurements of ΔN_{eff}, whereas all of the white region will be tested by CMB-S4. Predictions presented here for ΔN_{eff} close to the gray boundary are more uncertain because of the high sensitivity of the freeze-out temperatures to the coupling.

try to inflationary sectors is to introduce an additional $\mathbb{Z}_2$-odd scalar field η (as was done in [373]). This admits linear and quadratic interactions to antisymmetric and symmetric combinations of the inflationary sector fields, respectively. When η acquires a vev, this introduces an asymmetry in the fields to which it was coupled, dependent on the combination of its vev and its couplings. If η is coupled to both the inflationary sectors and the Higgs sectors, it could be the sole source of $\mathbb{Z}_2$-breaking in a twinflationary theory. One

may generally imagine that, in some UV completion, the mechanism that softly breaks the symmetry in the Higgs potential could also be the origin of the soft breaking of the inflationary sector.

Cosmologically, this possibility may have similar observational signatures as the model discussed in Section 5.2.3, where the amount of twin-sector dark radiation is determined by the partial widths of the inflaton of the second inflationary epoch. If this dominantly couples to the SM, then ΔN_{eff} will be suppressed which, while successfully resolving the cosmological problems of the Mirror Twin Higgs, may also be observationally inaccessible. However, additional, distinctly inflationary signatures may make this potentially testable by other cosmological observations.

The mechanism of twinflation completes a catalog of models of asymmetric reheating by late decays, which may be indexed by representations of the twin parity: the case of a $\mathbb{Z}_2$-even particle, in which a kinematic asymmetry in the partial widths provides the reheating asymmetry, the case of a $\mathbb{Z}_2$-odd particle, which can also provide the spontaneous $\mathbb{Z}_2$-breaking required in the Higgs potential, and the case where two distinct, long-lived particles couple to each sector, which may also be related to inflation.

Toy Model

As a toy model we here consider 'twinning' the simple φ^2 chaotic inflation scenario. The inflationary dynamics in this case are easy to understand and we have the additional benefit that this inflationary model has been considered in the literature before as 'Double Inflation' (see [374], [375] and [376]). We furthermore specialize to 'double inflation with a break', where there are two distinct periods of inflation which produces a step in the power spectrum, and we consider the constraints that this places on our model. In this case, it is assumed that each inflaton field couples and therefore decays dominantly into the sector to which it belongs. We will comment briefly on the case without a break and

204

the additional signals one could look for in that case.

The potential of the inflationary sector for inflaton φ_A and twinflaton φ_B is

$$V = \frac{1}{2}m_A^2\varphi_A^2 + \frac{1}{2}m_B^2\varphi_B^2, \tag{5.47}$$

where $m_A \neq m_B$ may arise from soft $\mathbb{Z}_2$- breaking, perhaps related to the soft $\mathbb{Z}_2$-breaking in the Higgs potential. In order for the 'twinflation' to occur first, we require that the energy of the B field initially dominates the energy density of the universe. We take the initial positions of the fields to be the same and $m_B^2 \gg m_A^2$.[22] Call $\varphi_A(0) = \varphi_B(0) = n\sqrt{2}M_{\mathrm{pl}} = n\varphi_c$, where φ_c is the critical value at which inflation stops and $m_B = rm_A = rm$ with $n, r > 1$. The inflationary dynamics are then those of slowly-rolling scalar fields. At some point in the early universe we imagine that the slow-roll approximation holds for both fields and the inflationary sector dominates the universe. The dominating field then slow-rolls down its potential for $\frac{n^2-1}{2}$ e-folds, while the lighter field's velocity is suppressed by approximately $\frac{\varphi_A}{r^2\varphi_B}$. Solving the system numerically reveals that the motion of φ_A during this period can be neglected entirely.

After φ_B reaches the critical value $\sqrt{2}M_{\mathrm{pl}}$, it stops slow-rolling and begins oscillating around the minimum of its potential. For there to be two distinct periods of inflation, there must be a period where these oscillations dominate the universe, which requires that the energy densities of each inflaton ρ_A and ρ_B satisfy $\rho_B(\varphi_c) = r^2m^2M_{\mathrm{pl}}^2 > \rho_A(\varphi(0)) = n^2m^2M_{\mathrm{pl}}^2$ and therefore $r > n$. For a φ_B^2 potential, the energy in these oscillations redshifts as $\rho_B \sim a^{-3}$. Eventually, the energy density in φ_B drops below that of φ_A and a new epoch of inflation, driven by φ_A, begins. This provides a further $\frac{n^2-1}{2}$ e-folds of inflation to give $n^2 - 1$ in total, while the B-sector energy density is diluted away.

---

[22]Note that merely giving the twin field a much larger initial condition does not instigate twinflation. The dynamics of the subdominant field in this case are such that it will track the dominant field and both will reach the critical value at the same time. This is easily confirmed numerically.

Note that in order for our toy model to reheat below the decoupling temperature of the two sectors, reheating must occur well after the end of inflation. If, during the coherent oscillation of an inflaton, it becomes the case that the inflaton decay width $\Gamma \sim H$, then reheating will occur and result in temperature $T_{\text{reheat}} \sim 0.1\sqrt{\Gamma M_{\text{pl}}}$. However, if $\Gamma \gg H$ when inflation ends, then all of the energy in the inflaton is immediately transferred and we instead have reheating temperature $T_{\text{reheat}} \sim 0.1\sqrt{m_\alpha M_{\text{pl}}}$ for an inflaton of mass m_α. But in order for $T_{\text{reheat}} \lesssim 1$ GeV, it is required that $m_\alpha \lesssim 10^{-7}$ eV, so this possibility that the inflaton is short lived is not viable. The procedure of twinning inflationary potentials may be generalised to other, more realistic models, provided that this constraint upon the reheating temperature can be satisfied.

Observability

One could always make a twinflationary scenario consistent with observational constraints by letting the second inflationary period of inflation last long enough. In our toy model, this would correspond to setting n high enough that the momentum modes which left the horizon during the first inflation have not yet re-entered the horizon - such a scenario would look exactly like single-field chaotic inflation.

Alternatively, we may also allow for n small enough that all the momentum modes that left the horizon during the second inflation are currently sub-horizon. In this case, fluctuations at large enough wavenumbers (equivalently, small enough length scales) are 'processed' (cross the horizon) at a different inflationary energy scale than those that were processed earlier, giving a step in the power spectrum. While Planck has measured the primordial power spectrum for modes with 10^{-4} Mpc$^{-1} \lesssim k \lesssim 0.3$ Mpc^{-1} (where the lower bound is set by the fact that smaller modes have not yet re-entered the horizon), proposed CMB-S4 experiments will increase this range [350] somewhat, as will be discussed further below. We wish to show that the power spectrum of our toy model is not

ruled out and, furthermore, may be observed in the coming decades.

The height of the step in the primordial power spectrum is determined by the energy scale of each period of inflation, so modes crossing the horizon in the second inflationary period should be suppressed by a factor of $r^2 > n^2 \gtrsim 25$ compared to those exiting in the first period. This degree of suppression is ruled out by Planck for the range of modes over which it has reconstructed the power spectrum [372]. A computation of the primordial power spectrum for double inflation was given in [375]. It was found that significant damping does not occur for modes which cross outside the horizon during the first inflationary period, re-enter during the inter-inflationary period and again cross the horizon during the second inflationary period. It is only those scales which first cross the horizon during the second inflationary period that are significantly damped (although other features in the shape, such as oscillations, may be present for modes that are subhorizon during the intermediate period).

The relation of this characteristic scale to present-day observables is easily done using the framework given in [377]. Let the subscripts a, b, c, d, e respectively correspond to the beginning of the first inflationary period, the end of that period, the beginning of the second inflationary period, the end of that period, and the beginning of radiation domination. During the coherent oscillation periods, the inflaton acts as matter and the energy density falls as $\rho \propto a^{-3}$. Let k_i be the momentum whose mode is horizon-size at the i epoch; $k_i = a_i H_i$. The scales k_i can be related using the number of e-folds in each period, which are themselves determined from the first Friedmann equation. Denoting $N_{ij} = \ln \frac{a_j}{a_i}$, we have $k_a = e^{-N_{ab}} n k_b$, $k_b = e^{\frac{1}{2} N_{bc}} k_c$ and similarly for the other characteristic modes, where, in particular, slow-roll inflation predicts that $N_{ab} = N_{cd} = \frac{n^2 - 1}{2}$. The evolution of the characteristic momentum scales is shown schematically in Figure 5.11.
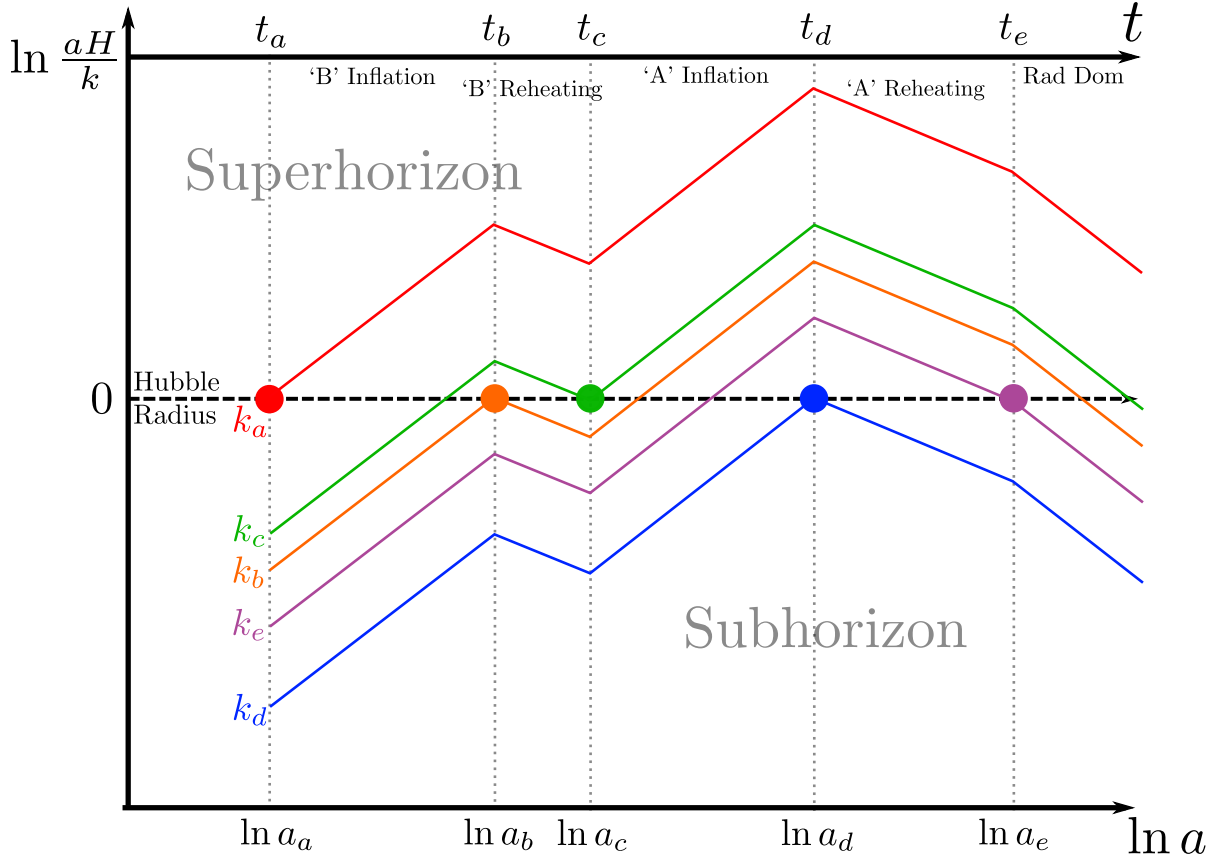
Figure 5.11: Schematic evolution of the characteristic scales in Twinflation, as seen by comparing wavenumbers to the Hubble radius over time. Note that the time axis is not a linear scale.

Finally, k_e can be determined using the conservation of comoving entropy:

$$k_e \;\; = \;\; \frac{\pi g_\star^{1/3}(T_0) g_\star^{1/6}(T_{\mathrm{reheat}}) T_0 T_{\mathrm{reheat}}}{3\sqrt{10}\, M_{\mathrm{pl}}}, \tag{5.48}$$

where T_0 and a_0 are the temperature and scale factor today and T_{reheat} is the reheating temperature (which is sufficiently low that only SM particles are produced). We work explicitly with the convention $a_0 = 1$. The characteristic modes associated with the break can then be determined.

As mentioned above, [375] shows that damping occurs for modes that exit the horizon

only during the second inflationary period, so we should take the characteristic damping scale to be the smallest such scale, which here corresponds roughly to k_b This can be determined as

$$
\begin{aligned}
k_b &= n e^{\frac{1}{2}N_{bc} - N_{cd} + \frac{1}{2}N_{de}} k_e \\
&= n \left(\frac{r}{n}\right)^{1/3} \exp\left(-\frac{n^2 - 1}{2}\right) \left[\frac{\frac{1}{2}m^2 M_{\text{pl}}^2}{\frac{\pi^2}{30} g_\star(T_{\text{reheat}}) T_{\text{reheat}}^4}\right]^{1/6} \frac{\pi g_\star^{1/3}(T_0) g_\star^{1/6}(T_{\text{reheat}}) T_0 T_{\text{reheat}}}{3\sqrt{10}\, M_{\text{pl}}}
\end{aligned}
\tag{5.49}
$$

where k_c only differs by the factor of $(r/n)^{1/3}$ (which is roughly close to unity). Once again, between k_b and k_c are oscillatory features, so k_b should merely be taken as the rough characteristic scale of the damping.
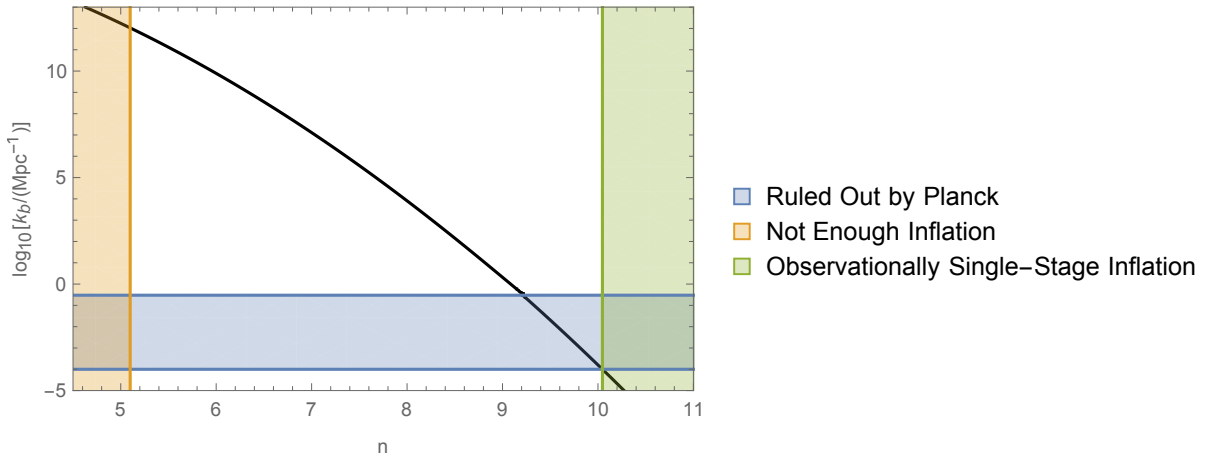
Figure 5.12: The prediction for the characteristic suppression scale as a function of the initial values of the fields. The mapped regions should be interpreted not as having hard boundaries, but rather fuzzy endpoints where they break down. Here we have used $T_{\text{reheat}} = 10$ MeV and $r = 2n$.

Now the characteristic damping scale is determined by m, n, r, and T_{reheat}. Our observational bound on k_b is that Planck has not seen this suppression on momentum scales at which it has been able to reconstruct the primordial power spectrum from the angular temperature anisotropy power spectrum, which is roughly $k \lesssim 0.3$ Mpc^{-1}. We have constraints on the reheating temperature from rethermalization of the twin sector

or interrupted big bang nucleosynthesis $10\,\mathrm{MeV} \lesssim T_{\mathrm{reheat}} \lesssim 1$ GeV, on having a period of intermediate matter domination between the two inflations $r > n$ and on the total number of e-folds $n^2 - 1 \gtrsim 25$ to solve cosmological problems. Note that we require fewer e-folds of inflation than is typically assumed in the standard cosmology. Since the low reheating temperature gives fewer e-folds from reheating up to today, less inflation is needed to explain the large causal horizon and flatness.

The normalization of the spectrum provides a further constraint, the most recent measurement of which come from Planck [372]. The scalar power spectrum at $k_\star = 0.05$ Mpc^{-1} is measured to be $\mathcal{P}_\mathcal{R}(k_\star) = e^{3.094 \pm 0.034} \times 10^{-10}$. Then for $k_\star < k_c$ (i.e. $k_\star$ having left the horizon during the first period of inflation and not re-entered before the second, so no deviation from single-field inflation would be seen at this scale), the spectrum of [375] yields the constraint

$$2.03 \times 10^{-6} = \frac{r^2 m^2}{M_{\mathrm{pl}}^2} \ln\left(\frac{k_b}{k_\star}\right) \left(\ln\frac{k_b}{k_\star} + \frac{n^2}{2}\right). \qquad (5.50)$$

The characteristic scale (5.49) depends much more strongly on n than it does on any of the other parameters. In Figure 5.12, we give a rough idea of the scale as a function of n, having set $T_{\mathrm{reheat}} = 10$ MeV and $r = 2n$, while m is chosen to satisfy the normalization condition. We also show the constraint on k_b set by Planck. Note again that the region described as "observationally single-stage inflation" *does* still provide a solution to the problem of reconciling cosmology with the mirror Twin Higgs.

CMB-S4 will improve the constraint on k_b through its improved measurement of polarization anisotropies [350]. With only precision measurements of temperature anisotropies, the un-lensed power spectrum cannot be so easily reconstructed from the lensed spectrum. The effects of gravitational lensing of CMB place an upper limit on the size of primordial temperature anisotropies that can be measured [378], which Planck has

saturated. However, the polarization anisotropy power spectrum allows the removal of lensing noise from the temperature spectrum so that higher primordial modes can be detected. The polarization power spectrum itself also gives us another window into the high-ℓ modes of the primordial power spectrum, as the signal does not become dominated by polarized foreground sources until higher scales near $\ell \sim 5000$. CMB-S4 is projected to make cosmic variance limited measurements of both the temperature and polarization anisotropy power spectra up to the modes where they become foreground-contaminated and so provide additional information on the shape of the primordial power spectrum [350]. The map from measurements of angular modes ℓ to contraints on spatial modes k depends on the evolution of the power spectrum between inflation and the CMB, so forecasting constraints requires careful study. However, these improvements will not test most of the parameter space presented in Figure 5.12, where the step is predicted on extremely small distance scales.

We have discussed a twinflationary model of double inflation with a break for simplicity, but there is a parametric regime where double inflation without a break gives the required amount of asymmetric reheating into the Standard Model. With two periods of inflation, the second period dilutes the energy density of the heavier field sufficiently that there is no observable signal of it produced in reheating. However, even with only one period, inflation can continue for long enough after the inflaton turns the corner in field space such that, at late times, the fraction of the inflaton in the B state relative to the A state is small enough that the expected energy densities that are transferred into each sector satisfy $\rho_B/\rho_A < 0.1$. This occurs as long as $r \gtrsim 1.2$, assuming that the mixing angle of the slow-rolling field with the φ_A and φ_B fields entirely determines the fraction of its energy that reheats each sector. There is thus a much larger range of r where this toy model of inflation passes N_{eff} bounds than our above analysis shows. The resulting imprint on the CMB could resemble that of the long-lived decay model of

Section 5.2.3, with ΔN_{eff} again being related to the ratio of branching fractions, although this is dependent upon the UV completion of the Twin Higgs.

When there is only one period of inflation, the step is smoothed out and less pronounced and it is necessary to locate the feature numerically. Furthermore, having multiple degrees of freedom available allows for non-trivial evolution of momentum modes after they become super-horizon, which does not occur in single-field inflation but may be calculated from the full solution to the field equations [376]. While a twinned potential leading to two periods of inflation generally predicts a step in the power spectrum, when there is no break the predictions, and thus constraints, this prediction become more model-dependent. Therefore we leave detailed predictions in that case for future study using realistic models and merely state that the range of $r = 1$ to n interpolates between the single field spectrum and that with a step, as one would expect.

There are also at least two other detectable effects one might expect in double inflation without a break and in general realistic twinflationary models. Interactions between inflaton fields may produce primordial non-Gaussianities, while the presence of additional oscillating degrees of freedom may produce isocurvature perturbations. These do not appear in our toy model because the heavy field is exponentially damped during the second inflation. CMB-S4 is projected to improve Planck's bounds on non-Gaussianities by a factor of ~ 2 and on isocurvature perturbations by perhaps an order of magnitude (though model-independent projections have not been made), so may be able to detect or place useful constraints on realistic twinflationary models [350].

We have introduced twinflation as a mirror Twin Higgs model which suppresses the cosmological effects of twin light degrees of freedom. It extends the mirror symmetry to the inflationary sector. The soft Z_2 symmetry-breaking of the Higgs sector may be used in the inflationary sector to cause distinct periods of inflation. There exists a parametric region where this is cosmologically indistinct from single-stage inflation, but also another

in which it may be observable. As the direct product of inflation and the Mirror Twin Higgs, this is in some sense a minimal solution.

5.2.5   Conclusion

In this section we have considered scenarios in which cosmology provides meaningful insight on solutions to the electroweak hierarchy problem. In particular, we have demonstrated several simple mechanisms in which the cosmological history of a mirror Twin Higgs model is reconciled with current CMB constraints and provides signatures accessible in future CMB experiments. In the case of out-of-equilibrium decays, we have found that decays of $\mathbb{Z}_2$-even scalars sufficiently dilute the energy density in the twin sector without the addition of any new sources of $\mathbb{Z}_2$-breaking. In much of the parameter space, the residual contribution to ΔN_{eff} is directly proportional to the ratio of vacuum expectation values v^2/f^2 parameterizing the mixing between Standard Model and twin sectors (as well as the tuning of the electroweak scale), and may be within reach of CMB-S4 experiments. In the case of twinflation, we have found that a (broken) $\mathbb{Z}_2$-symmetric inflationary sector may successfully dilute the energy density in the twin sector, as well as potentially leave signatures in the form of a step in the primordial power spectrum or in departures of primordial perturbations from adiabaticity and Gaussianity. In both cases, these models raise the tantalizing possibility that signatures of electroweak naturalness may first emerge in the CMB, rather than the LHC.

There are a variety of possible directions for future work. Here we have focused on the cosmological consequences of late-decaying scalars and twinned inflationary sectors without specifying their origin in a microscopic model. It would be interesting to construct complete models (where, e.g., supersymmetry or compositeness protect the scale f from UV contributions) in which the existence and couplings of late-decaying scalars

arise as intrinsic ingredients of the UV completion. Likewise, we have considered only a toy model of twin chaotic inflation; it would be interesting to see if twinflation may be realized in complete inflationary models that match the observed spectral index and constraints on the tensor-to-scalar ratio.

While we have taken care to ensure that our scenarios respect the well-measured cosmological history beneath $T \sim 1$ MeV, we have not addressed the origin of the observed baryon asymmetry. In the case of out-of equilibrium decays, there are a number of possibilities. It is plausible that a somewhat larger baryon asymmetry is generated through various conventional mechanisms and diluted by late decays. Alternatively, the decay mechanism itself may possibly be expanded to generate a baryon asymmetry or some other late decay may generate the baryon asymmetry below ~ 1 GeV. In the case of twinflation, inflationary dilution of pre-existing baryon asymmetry requires that baryogenesis occur in association with reheating or via another mechanism at temperatures below ~ 1 GeV. It would be worthwhile to study models for the baryon asymmetry consistent with these scenarios. Steps in this direction have been taken in [338], which attempted to relate this to asymmetric dark matter in the twin sector.

Likewise, any investigation of dark matter, be it related directly to the twin mechanism or otherwise, must also address implications of the dilution. Previous work attempting to construct dark matter candidates in the twin sector [323, 333, 334, 335, 336, 337, 338, 339]) has relied upon explicit $\mathbb{Z}_2$-breaking that is not present in the mirror model. Dark matter may alternatively be unrelated to the Twin Higgs mechanism, such as a a WIMP in some minimal extension of the electroweak sector that freezes-out as an overabundant thermal relic and is then diluted to the observed density during reheating. Alternatively, it may be that the dark matter abundance is produced directly during reheating. It would be interesting to study extensions of our scenarios that incorporate dark matter candidates directly related to the mechanism of dilution.

Finally, we have only approximately parameterized Planck constraints and the reach of CMB-S4 on twin neutrinos and twin photons. Ultimately, more precise constraints and forecasts may be obtained via numerical CMB codes. This strongly motivates the future study of CMB constraints on scenarios with three sterile neutrinos and additional dark radiation whose temperatures differ from the Standard Model thermal bath.

5.3 Freeze-*Tw*in Dark Matter

5.3.1 Introduction

In this work, we build on a MTH framework where 'hard' breaking of the $\mathbb{Z}_2$ is absent. In [13, 379], it was realized that late-time asymmetric reheating of the two sectors could arise naturally in these models if the spectrum were extended by a single new state. This asymmetric reheating would dilute the twin energy density and so attune the MTH with the cosmological constraints. This dilution of twin energy density to negligible levels would seem to hamper the prospect that twin states might constitute the dark matter, and generating dark matter was left as an open question. This presents a major challenge toward making such cosmologies realistic. However, we show that asymmetric reheating perfectly sets the stage for a MTH realization of the 'freeze-in' mechanism for dark matter production [380, 381, 382, 383, 384, 385, 386, 387].

Freeze-in scenarios are characterized by two assumptions: 1) DM has a negligible density at some early time and 2) DM interacts with the SM so feebly that it never achieves thermal equilibrium with the SM bath.[23] This second assumption is motivated in part by the continued non-observation of non-gravitational DM-SM interactions. Both assumptions stand in stark contrast to freeze-out scenarios.

---

[23]We note that the feeble connection between the two sectors may originate as a small dimensionless coupling or as a small ratio of mass scales, either of which deserves some explanation.

Freeze-twin dark matter is a particularly interesting freeze-in scenario because both assumptions are fulfilled for reasons orthogonal to dark matter considerations: 1) the negligible initial dark matter abundance is predicted by the asymmetric reheating already necessary to resolve the MTH cosmology, and 2) the kinetic mixing necessary to achieve the correct relic abundance is of the order expected from infrared contributions in the MTH. To allow the frozen-in twin electrons and positrons to be DM, we need only break the $\mathbb{Z}_2$ by a relevant operator to give a Stueckelberg mass to twin hypercharge. Additionally, the twin photon masses we consider can lead to dark matter self-interactions at the level relevant for small-scale structure problems [388].

The next sections are organized as follows: In Section 5.3.2, we review the MTH and its cosmology in models with asymmetric reheating, and in Section 5.3.3 we introduce our extension. In Section 5.3.4, we calculate the freeze-in yield for twin electrons and discuss the parameter space to generate dark matter and constraints thereon. We discuss future directions and conclude in Section 5.3.5. For the interested reader, we include some discussion of the irreducible IR contributions to kinetic mixing in the MTH in Appendix A.

5.3.2 The Mirror Twin Higgs & Cosmology

The mirror twin Higgs framework [267] introduces a twin sector B, which is a 'mirror' copy of the Standard Model sector A, related by a $\mathbb{Z}_2$ symmetry. Upgrading the $SU(2)_A \times SU(2)_B$ gauge symmetry of the scalar potential to an $SU(4)$ global symmetry adds a Higgs-portal interaction between the A and B sectors:

$$V = \lambda \left(|\mathcal{H}|^2 - f^2/2 \right)^2 , \tag{5.51}$$

where $\mathcal{H} = \begin{pmatrix} H_A \\ H_B \end{pmatrix}$ is a complex $SU(4)$ fundamental consisting of the A and B sector

Higgses in the gauge basis. The SM Higgs is to be identified as a pseudo-Goldstone mode

arising from the breaking of $SU(4) \to SU(3)$ when $\mathcal{H}$ acquires a vacuum expectation

value (vev) $\langle \mathcal{H} \rangle = f/\sqrt{2}$. Despite the fact that the global $SU(4)$ is explicitly broken

by the gauging of $SU(2)_A \times SU(2)_B$ subgroups, the $\mathbb{Z}_2$ is enough to ensure that the

quadratically divergent part of the one-loop effective action respects the full $SU(4)$. The

lightness of the SM Higgs is then understood as being protected by the approximate

accidental global symmetry up to the UV cutoff scale $\Lambda \lesssim 4\pi f$, at which point new

physics must come in to stabilize the scale f itself.

We refer to twin particles by their SM counterparts primed with a superscript ', and

we refer the reader to [267, 268] for further discussion of the twin Higgs mechanism.

The thermal bath history in the conventional MTH is fully dictated by the Higgs

portal in Eq. (5.51) which keeps the SM and twin sectors in thermal equilibrium down to

temperatures $\mathcal{O}(\text{GeV})$. A detailed calculation of the decoupling process was performed

in [13] by tracking the bulk heat flow between the two sectors as a function of SM

temperature. It was found that for the benchmark of $f/v = 4$, decoupling begins at a

SM temperature of $T \sim 4$ GeV and by ~ 1 GeV, the ratio of twin-to-SM temperatures

may reach $\lesssim 0.1$ without rebounding. While heat flow rates become less precise below

~ 1 GeV due to uncertainties in hadronic scattering rates, especially close to color-

confinement, decoupling between the two sectors is complete by then for $f/v \gtrsim 4$. For

larger f/v, the decoupling begins and ends at higher temperatures.

As mentioned above, one class of solutions to this N_{eff} problem uses hard breaking of

the $\mathbb{Z}_2$ at the level of the spectra [268, 323, 324, 325, 326] while keeping a standard cos-

mology. An alternative proposal is to modify the cosmology with asymmetric reheating to

dilute the energy density of twin states. For example, [379] uses late, out-of-equilibrium

decays of right-handed neutrinos, while [13] uses those of a scalar singlet. These new particles respect the $\mathbb{Z}_2$, but dominantly decay to SM states due to the already-present soft $\mathbb{Z}_2$-breaking in the scalar sector. In [379], this is solely due to extra suppression by f/v-heavier mediators, while in [13], the scalar also preferentially mass-mixes with the heavier twin Higgs. [13] also presented a toy model of 'Twinflation', where a softly-broken $\mathbb{Z}_2$-symmetric scalar sector may lead to inflationary reheating which asymmetrically re-heats the two sectors to different temperatures. In any of these scenarios, the twin sector may be diluted to the level where it evades Planck bounds [389] on extra radiation, yet is potentially observable with CMB Stage IV [390].

We will stay agnostic about the particular mechanism at play, and merely assume that by $T \sim 1$ GeV, the Higgs portal interactions have become inefficient and some mechanism of asymmetric reheating has occurred such that the energy density in the twin sector has been largely depleted, $\rho_{\text{twin}} \approx 0$.[24] This is consistent with the results of the decoupling calculation in [13] given the uncertainties in the rates at low temperatures, and will certainly be true once one gets down to few $\times 10^2$ MeV.

One may be concerned that there will be vestigial model-dependence from irrelevant operators induced by the asymmetric reheating mechanism which connect the two sectors. However, these operators will generally be suppressed by scales above the reheating scale, as in the example studied in [13]. Prior to asymmetric reheating, the two sectors are in thermal equilibrium anyway, so these have little effect. After the energy density in twin states has been diluted relative to that in the SM states, the temperature is far below the heavy masses suppressing such irrelevant operators, and thus their effects are negligible. So we may indeed stay largely agnostic of the cosmological evolution before asymmetric reheating as well as the details of how this reheating takes place. We take the absence

---

[24]If asymmetric reheating leaves some small $\rho_{\text{twin}} > 0$, then mirror baryon asymmetry can lead to twin baryons as a small subcomponent of dark matter [391].

of twin energy density as an initial condition, but emphasize that there are external, well-motivated reasons for this to hold in twin Higgs models, as well as concrete models that predict this occurrence naturally.

5.3.3   Kinetic Mixing & A Massive Twin Photon

In order to arrange for freeze-in, we add to the MTH kinetic mixing between the SM and twin hypercharges and a Stueckelberg mass for twin hypercharge. At low energies, these reduce to such terms for the photons instead, parametrized as

$$\mathcal{L} \mathrel{+}= -\frac{\epsilon}{2} F_{\mu\nu} F'^{\mu\nu} - \frac{1}{2} m_{\gamma'}^2 A'_\mu A'^\mu. \tag{5.52}$$

This gives each SM particle of electric charge Q an effective twin electric charge ϵQ.[25] The twin photon thus gives rise to a 'kinetic mixing portal' through which the SM bath may freeze-in light twin fermions in the early universe.

The Stueckelberg mass constitutes soft $\mathbb{Z}_2$-breaking,[26] but has no implications for the fine-tuning of the Higgs mass since hypercharge corrections are already consistent with naturalness [268]. We will require $m_{\gamma'} > m_{e'}$, to prevent frozen-in twin electron/-positron annihilations, and $m_{\gamma'} > 2m_{e'}$, to ensure that resonant production through the twin photon is kinematically accessible. Resonant production will allow a much smaller kinetic mixing to generate the correct relic abundance, thus avoiding indirect bounds

---

[25]Note that twin charged states do not couple to the SM photon. Their coupling to the SM Z boson has no impact on freeze-in at the temperatures under consideration. Furthermore, the miniscule kinetic mixing necessary for freeze-in has negligible effects at collider experiments. See Ref. [392] for details.

[26]While we are breaking the $\mathbb{Z}_2$ symmetry by a relevant operator, the extent to which a Stueckelberg mass is truly *soft* breaking is not clear. Taking solely Eq. (5.52), we would have more degrees of freedom in the twin sector than in the SM, and in a given UV completion it may be difficult to isolate this $\mathbb{Z}_2$-asymmetry from the Higgs potential. One possible fix may be to add an extremely tiny, experimentally allowed Stueckelberg mass for the SM photon as well [393], though we note this may be in violation of quantum gravity [394, 395] or simply be difficult to realize in UV completions without extreme fine-tuning. We will remain agnostic about this UV issue and continue to refer to this as 'soft breaking', following [392].

from supernova cooling. We note that while taking $m_{\gamma'} \ll f$ does bear explanation, the parameter is technically natural.

On the other hand, mixing of the twin and SM $U(1)$s preserves the symmetries of the MTH EFT, so quite generally one might expect it to be larger than that needed for freeze-in. However, it is known that in the MTH a nonzero ϵ is not generated through three loops [267]. While such a suppressed mixing is phenomenologically irrelevant for most purposes, here it plays a central role. In Appendix A, we discuss at some length the vanishing of infrared contributions to kinetic mixing through few loop order. If nonzero contributions appear at the first loop order where they are not known to vanish, kinetic mixing of the order $\epsilon \sim 10^{-13} - 10^{-10}$ is expected.

The diagrams which generate kinetic mixing will likely also generate higher-dimensional operators. These will be suppressed by (twin) electroweak scales and so, as discussed above for the irrelevant operators generated by the model-dependent reheating mechanism, freeze-in contributions from these operators are negligible.

5.3.4 Freezing-*Tw*in Dark Matter

As we are in the regime where freeze-in proceeds while the temperature sweeps over the mass scales in the problem, it is not precisely correct to categorize this into either "UV freeze-in" or "IR freeze-in". Above the mass of the twin photon, freeze-in proceeds through the marginal kinetic mixing operator, and so a naive classification would say this is IR dominated. However, below the mass of the twin photon, the clearest approach is to integrate out the twin photon, to find that freeze-in then proceeds through an irrelevant, dimension-six, four-Fermi operator which is suppressed by the twin photon mass. Thus, at temperatures $T_{\text{SM}} \lesssim m_{\gamma'}$, this freeze-in looks UV dominated. This leads to the conclusion that the freeze-in rate is largest at temperatures around the mass of the

twin photon. Indeed, this is generally true of freeze-in — production occurs mainly at temperatures near the largest relevant scale in the process, whether that be the largest mass out of the bath particles, mediator, and dark matter, or the starting thermal bath temperature itself in the case that one of the preceding masses is even higher.

As just argued, freeze-in production of dark matter occurs predominantly at and somewhat before $T \sim m_{\gamma'}$. We require $m_{\gamma'} \ll 1$ GeV so that most of the freeze-in yield comes from when $T < 1$ GeV, which allows us to retain 'UV-independence' in that we need not care about how asymmetric reheating has occurred past providing negligible density of twin states at $T = 1$ GeV. Specifically, we limit ourselves to $m_{\gamma'} < 2m_{\pi^0}$, both for this reason and to avoid uncertainties in the behavior of thermal pions during the epoch of the QCD phase transition. However, we emphasize that freeze-in will remain a viable option for producing a twin DM abundance for even heavier dark photons. But the fact that the freeze-in abundance will be generated simultaneously with asymmetric reheating demands that each sort of asymmetric reheating scenario must then be treated separately. Despite the additional difficulty involved in predicting the abundance for larger twin photon masses, it would be interesting to explore this part of parameter space. In particular, it would be interesting to consider concrete scenarios with twin photons in the range of tens of GeV [396].

To calculate the relic abundance of twin electrons and positrons, we use the Boltzmann equation for the number density of e':

$$\dot{n}_{e'} + 3Hn_{e'} = \sum_{k,l} - \langle \sigma v \rangle_{e'\bar{e}' \to kl} \left(n_{e'}n_{\bar{e}'} - n_{e'}^{\mathrm{eq}}n_{\bar{e}'}^{\mathrm{eq}} \right), \tag{5.53}$$

where $\langle \sigma v \rangle_{e'\bar{e}' \to kl}$ is the thermally averaged cross section for the process $e'\bar{e}' \to kl$, the sum runs over all processes with SM particles in the final states and $e'\bar{e}'$ in the initial state, and $n_{e'}^{\mathrm{eq}}$ is the equilibrium number density evaluated at temperature T. As we are

in the parametric regime in which resonant production of twin electrons through the twin photon is allowed, $2 \to 2$ annihilation processes $\bar{f}f \to \gamma' \to \bar{e}'e'$, with f a charged SM fermion, entirely dominate the yield.

In accordance with the freeze-in mechanism, $n_{e'}$ remains negligibly smaller than its equilibrium number density throughout the freeze-in process, and so that term is ignored. It is useful to reparametrize the abundance of e' in terms of its yield, $Y_{e'} = n_{e'}/s$ where $s = \frac{2\pi^2}{45}g_{*s}T^3$ is the entropy density in the SM bath. Integrating the Boltzmann equation using standard methods, we then find the yield of e' today to be

$$
\begin{aligned}
Y_{e'} = \int_0^{T_i} dT \frac{(45)^{3/2}}{\sqrt{2}\,\pi^3\sqrt{g_*}\,g_{*s}} \frac{M_{Pl}}{T^5}\left(\frac{1}{T} + \frac{\partial_T g_{*s}}{3g_{*s}}\right) \\
\times \sum_{\bar{f}f\to\bar{e}'e'} \langle\sigma v\rangle_{\bar{f}f\to\bar{e}'e'}\, n_{\bar{e}'}^{\text{eq}} n_{e'}^{\text{eq}},
\end{aligned}
\tag{5.54}
$$

where $T_i = 1$ GeV is the initial temperature of the SM bath at which freeze-in begins in our setup, $g_*(T)$ is the number of degrees of freedom in the bath, and M_{Pl} is the reduced Planck mass. We will calculate this to an intended accuracy of 50%. To this level of accuracy, we may assume Maxwell-Boltzmann statistics to vastly simplify the calculation [397]. As a further simplification, we observe that the $\partial_T g_{*s}$ term is negligible compared to $1/T$ except possibly during the QCDPT - where uncertainties on its temperature dependence remain [345] - and so we ignore that term. The general expression for the thermally averaged cross section of the process $12 \to 34$ is then

$$
\begin{aligned}
\langle\sigma v\rangle\, n_1^{\text{eq}} n_2^{\text{eq}} = \frac{T^4}{2^9\pi^5 s_{34}}\int_{\text{Max}\left(\frac{m_1+m_2}{T}, \frac{m_3+m_4}{T}\right)}^{\infty} dx\, x^2 \\
\times \sqrt{[1,2]}\sqrt{[3,4]}K_1(x)\int d(\cos\theta)\, |\mathcal{M}|^2_{12\to34},
\end{aligned}
\tag{5.55}
$$

where s_{34} is 1 if the final states are distinct and 2 if not, $x = \sqrt{s}/T$, $\sqrt{[i,j]} =$

$\sqrt{1 - \left(\frac{m_i + m_j}{xT}\right)^2} \sqrt{1 - \left(\frac{m_i - m_j}{xT}\right)^2}$, and $|\mathcal{M}|^2_{12 \to 34}$ is the matrix element squared summed (not averaged) over all degrees of freedom. To very good approximation, the yield results entirely from resonant production, and so we may analytically simplify the matrix element squared for $\bar{f}f \to \bar{e}'e'$ using the narrow-width approximation

$$\int d\left(\cos\theta\right) |\mathcal{M}|^2_{\bar{f}f \to \bar{e}'e'} \approx \frac{256\pi^3 \alpha^2 \epsilon^2}{3} \left(2m_f^2 + m_{\gamma'}^2\right) \tag{5.56}$$
$$\times \frac{\left(2m_{e'}^2 + m_{\gamma'}^2\right)}{\Gamma_{\gamma'} m_{\gamma'}^2 T} \delta\left(x - m_{\gamma'}/T\right).$$

$\Gamma_{\gamma'}$ is the total decay rate of the twin photon.

For the range of $m_{\gamma'}$ we consider, the twin photon can only decay to twin electron and positron pairs. Thus, its total decay rate is

$$\Gamma_{\gamma'} = \frac{\alpha \left(m_{\gamma'}^2 + 2m_{e'}^2\right)}{3m_{\gamma'}} \sqrt{1 - \frac{4m_{e'}^2}{m_{\gamma'}^2}}. \tag{5.57}$$

Its partial widths to SM fermion pairs are suppressed by ϵ^2, and so contribute negligibly to its total width.

The final yield of twin electrons is then

$$Y_{e'} \approx \frac{3m_{\gamma'}^2}{2\pi^2} \frac{(45)^{3/2} M_{Pl}}{\sqrt{2}\,\pi^3} \sum_f \int_{T_f}^{T_i} dT \Gamma_{\gamma' \to \bar{f}f} \frac{K_1(\frac{m_{\gamma'}}{T})}{\sqrt{g_*}\, g_{*s} T^5}, \tag{5.58}$$

where $T_f = \Lambda_{\mathrm{QCD}}$ for quarks, $T_f = 0$ for leptons, $\Gamma_{\gamma' \to \bar{f}f}$ is the partial decay width of the twin photon to $f\bar{f}$, and the sum is over all SM fermion-antifermion pairs for which $m_{\gamma'} > 2m_f$.

Since we have approximated the yield as being due entirely to on-shell production and decay of twin photons, the analytical expression for the yield in Eq. (5.58) exactly agrees with the yield from freezing-in γ' via 'inverse decays' $\bar{f}f \to \gamma'$, as derived in [384].

We have validated our numerical implementation of the freeze-in calculation by success-fully reproducing the yield in similar cases found in [397, 398]. We have furthermore checked that reprocessing of the frozen-in dark matter [385, 399] through $e'\bar{e}' \to e'\bar{e}'e'\bar{e}'$ is negligible here,[27] as is the depletion from $e'\bar{e}' \to \nu'\bar{\nu}'$.

An equal number of twin positrons are produced as twin electrons from the freeze-in processes. Requiring that ϵ reproduce the observed DM abundance today, we find

$$\epsilon = \sqrt{\frac{\Omega_\chi h^2 \rho_{\mathrm{crit}}/h^2}{2 m_{e'} \tilde{Y}_{e'} s_0}}, \tag{5.59}$$

where $\Omega_\chi h^2 \approx 0.12$, $\rho_{\mathrm{crit}}/h^2 \approx 1.1 \times 10^{-5} \mathrm{GeV/cm^3}$, and $s_0 \approx 2900/\mathrm{cm^3}$ [45]. $\tilde{Y}_{e'}$ is the total yield with the overall factor of ϵ^2 removed. This requisite kinetic mixing appears in Fig. 5.13 as a function of the twin photon mass $m_{\gamma'}$ for the two benchmark f/v values 4 and 10. In grey, we plot constraints from anomalous supernova cooling. To be conservative, we include both, slightly different bounds from [402, 403]. The dashed regions of the lines show approximately where self-interactions through Bhabha scattering are relevant in the late universe, $\sigma_{\mathrm{elastic}}/m_{e'} \gtrsim 1 \mathrm{\ cm^2/g}$. Self-interactions much larger than this are constrained by the Bullet Cluster [404, 405, 406] among other observations. Interestingly, self-interactions of this order have been suggested to fix small-scale issues, and some claimed detections have been made as well. We refer the reader to [388] for a recent review of these issues.

As mentioned above and discussed further in Appendix A, the level of kinetic mixing required for freeze-in is roughly of the same order as is expected from infrared contri-butions in the MTH. It would be interesting to develop the technology to calculate the high-loop-order diagrams at which it may be generated. In the context of a complete

---

[27]To be conservative, we calculate the rate assuming all interactions take place at the maximum $\sqrt{s} \simeq m_{\gamma'}$ and find that it is still far below Hubble. We perform the calculation of the cross section using MadGraph [400] with a model implemented in Feynrules [401].
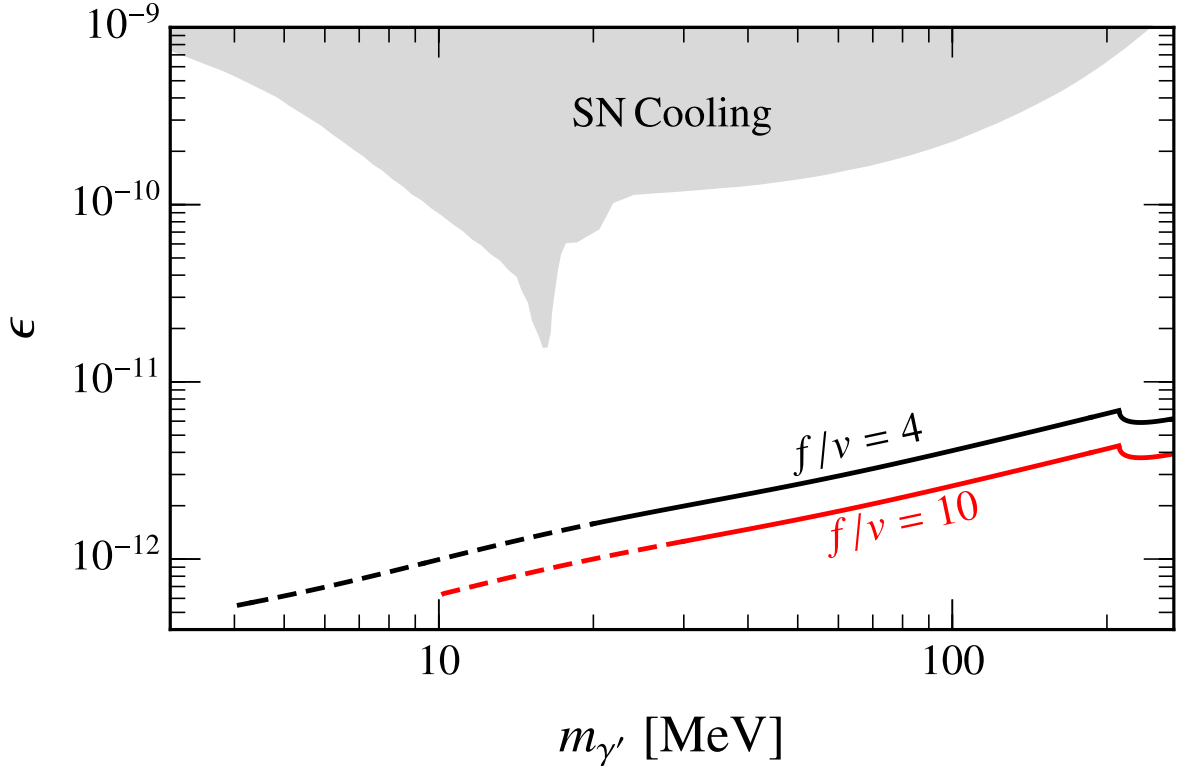
Figure 5.13: Contours in the plane of twin photon mass $m_{\gamma'}$ and kinetic mixing ϵ which freeze-in the observed DM abundance for two values of f/v. The dip at high masses corresponds to additional production via muon annihilations. In the dashed segments, self-interactions occur with $\sigma_{\mathrm{elastic}}/m_{e'} \gtrsim 1\mathrm{cm}^2/\mathrm{g}$. Also included are the combined supernova cooling bounds from [402, 403].

model of the MTH where kinetic mixing is absent in the UV, ϵ is fully calculable and depends solely on the scale at which kinetic mixing is first allowed by the symmetries. Calculating ϵ would then predict a minimal model at some $m_{\gamma'}$ to achieve the right dark matter relic abundance, making this effectively a zero-parameter extension of MTH models with asymmetric reheating. Importantly, even if ϵ is above those shown in Fig. 5.13, that would simply point to a larger value of $m_{\gamma'}$ which would suggest that the parameter point depends in more detail on the mechanism of asymmetric reheating. We note that in the case that the infrared contributions to ϵ are below those needed here, the required kinetic mixing may instead be provided by UV contributions and the scenario

is unaffected.

5.3.5   Conclusion

The mirror twin Higgs is perhaps the simplest avatar of the Neutral Naturalness program, which aims to address the increasingly severe little hierarchy problem. Understanding a consistent cosmological history for this model is therefore crucial, and an important step was taken in [13, 379]. As opposed to prior work, the cosmology of the MTH was remedied without hard breaking of the $\mathbb{Z}_2$ symmetry by utilizing asymmetric reheating to dilute the twin energy density. Keeping the $\mathbb{Z}_2$ as a good symmetry should simplify the task of writing high energy completions of these theories, but low-scale reheating may slightly complicate cosmology at early times. These works left as open questions how to set up cosmological epochs such as dark matter generation and baryogenesis in such models. We have here found that at least one of these questions has a natural answer.

In this work, we have shown that twin electrons and positrons may be frozen-in as dark matter following asymmetric reheating in twin Higgs models. This requires extending the mirror twin Higgs minimally with a single free parameter: the twin photon mass. Freezing-in the observed DM abundance pins the required kinetic mixing to a level expected from infrared contributions in MTH models. In fact, the prospect of calculating the kinetic mixing generated in the MTH could make this an effectively parameter-free extension of the MTH. Compared to generic freeze-in scenarios, it is interesting in this case that the "just so" stories of feeble coupling and negligible initial density were already present for reasons entirely orthogonal to dark matter.

This minimalism in freeze-twin dark matter correlates disparate signals which would allow this model to be triangulated with relatively few indirect hints of new physics. If

226

deviations in Higgs couplings are observed at the HL-LHC or a future lepton collider, this would determine f/v [407, 408, 12, 10], which would set the dark matter mass. An observation of anomalous cooling of a future supernova through the measurement of the neutrino 'light curve' might allow us to directly probe the $m_{\gamma'}, \epsilon$ curve [402, 403], though this would rely on an improved understanding of the SM prediction for neutrino production.[28] Further astrophysical evidence of dark matter self-interactions would point to a combination of f/v and $m_{\gamma'}$. All of this complementarity underscores the value of a robust experimental particle physics program whereby new physics is pursued via every imaginable channel.

---

[28]We thank Jae Hyeok Chang for a discussion on this point.

Chapter 6

Neutral Naturalness in the Ground

> There is nothing like looking, if you want to find something. You certainly usually find something, if you look, but it is not always quite the something you were after.
>
> J.R.R. Tolkien
>
> *The Hobbit*, 1937 [409]

Long-Lived Particles

When we introduced models of Neutral Naturalness in Section 4.1.2, we were motivated by the lack of signals of new, SM-charged particles at colliders. However, experimentalists are very clever, and it is in fact possible to observe the effects of such particles if you know where to look.

In particular, Neutral Naturalness models usually exhibit a 'hidden valley' type phenomenology, wherein a dark sector is connected to the SM only through some heavy states. At a collider, a high energy collision can transfer energy from our sector to these

heavy dark sector states, which may then decay just as heavy SM particles do. If the dark sector contains absolutely stable particles, then the energy may cascade into those states, which simply leave the detector. But without a symmetry dictating that stability, dark sector states may be destabilized by the effects of that same high-energy link to the SM. That connection may induce higher-dimensional operators which allow the decay of the dark sector states back into SM states. Since this decay channel comes from interactions of heavy states, the width may be highly suppressed, leaving to a macroscopically long lifetime, even when the scale set by the particle's mass is microscopically small. This leads to the appearance of SM particles out of nowhere inside a detector a macroscopic distance away from the interaction point in a collider.

In fact there are many reasons an unstable particle may be long-lived, and we see a multitude examples in the SM itself. In analogy with the hidden valley phenomenology, charged pions are long-lived compared to the QCD scale because their decay must take place through a heavy W boson. Neutrons are similarly very long-lived as a result of their small mass splitting with protons. Protons themselves are *very* long-lived as a result of an approximate global symmetry. The SM Higgs is long-lived compared to the electroweak scale because its leading decay mode proceeds through the small bottom quark Yukawas. Given the genericity of long-lived particles in our sector, it's entirely reasonable to imagine that some dark sector will have similar phenomenology. So even aside from our Neutral Naturalness motivation, such searches are generically useful and interesting things to look for.

In this section we forecast how well an electron-positron collider will be able to probe Higgs decays to long-lived particles, using the parameters for some proposed machines and search strategies we designed. Such forecasting is crucially important at the present time, as the community is still discussing what the next collider is that we will build. It's clearly necessary to know what sort of physics program we expect we can carry out

before we build a machine to do it. And these studies are used to motivate different types of detectors, their detailed design features and how trigger bandwidth is allocated.

6.1   Introduction

Following the discovery of the Higgs boson in 2012 [410, 411], the precision study of its properties has rapidly become one of the centerpieces of the physics program at the LHC. The expansion of this program beyond the LHC has become one of the key motivators for proposed future accelerators, including lepton colliders such as CEPC [412, 413], FCC-ee [414], the ILC [415, 416], and CLIC [417, 418] that would operate in part as Higgs factories.

The potential gains of a precision Higgs program pursued at both the LHC and future colliders are innumerable. Confirmation of Standard Model predictions for Higgs properties would mark a triumphant validation of the theory and illuminate phenomena never before seen in nature. The observation of deviations from Standard Model predictions, on the other hand, would point the way directly to additional physics beyond the Standard Model. Such deviations could take the form of changes in Higgs couplings to itself or other Standard Model states, or they could appear as exotic decay modes not predicted by the Standard Model. The latter possibility has been extensively explored for *prompt* exotic decay modes in the context of both the LHC (see e.g. [419, 420]) and future Higgs factories [421].

However, an equally compelling possibility is for new physics to manifest itself in exotic decays of the Higgs boson to long-lived particles (LLPs). Such signals were first considered in the context of Hidden Valleys [422, 423, 424] and subsequently found to arise in a variety of motivated scenarios for physics beyond the Standard Model, including solutions to the electroweak hierarchy problem [268] and models of baryogenesis [425;

for an excellent recent overview, see [426]. The search for exotic Higgs decays into LLPs necessarily involves strategies outside the scope of typical analyses. The non-standard nature of these signatures raises the compelling possibility of discovering new physics that has been heretofore concealed primarily by the novelty of its appearance.

There is a rich and rapidly growing program of LLP searches at the LHC. A variety of existing searches by the ATLAS, CMS, and LHCb collaborations (e.g. [427, 428, 429]; for a recent review see [430]) constrain Higgs decays into LLPs at roughly the percent level across a range of LLP lifetimes. Significant improvements in sensitivity are possible in future LHC runs with potential advances in timing [431], triggers [432, 433, 434], and analysis strategies [329, 328]. Most notable among these is the possible implementation of a track trigger [432, 433], which would significantly lower the trigger threshold for Higgs decays into LLPs and potentially allow sensitivity to branching ratios on the order of 10^{-6} in zero-background scenarios.

While studies of prompt exotic Higgs decays at future colliders [421] have demonstrated the potential for significantly improved reach over the LHC, comparatively little has been said about the prospects for constraining exotic Higgs decays to long-lived particles at the same facilities.[29] In this work we take the first steps towards filling this gap by studying the sensitivity of e^+e^- Higgs factories to hadronically-decaying new particles produced in exotic Higgs decays with decay lengths ranging from microns to meters. For the sake of definiteness we restrict our attention to circular Higgs factories operating at or near the peak rate for the Higgsstrahlung process $e^+e^- \rightarrow hZ$, namely CEPC and FCC-ee, while also sketching the corresponding sensitivity for the $\sqrt{s} = 250$ GeV stage of the ILC. While essentially all elements of general-purpose detectors may be brought to bear in the search for long-lived particles, the distribution of decay lengths for a given

---

[29] A notable exception is CLIC, for which a study of tracker-based searches for Higgs decays to LLPs has been recently performed [435]. For preliminary studies of other non-Higgs LLP signatures at future lepton colliders, see e.g. [436]. For studies of LLP signatures at future electron-proton colliders, see [437].

average lifetime make it advantageous to exploit detector elements close to the primary interaction point. We thus focus on signatures that can be identified in the tracker. In order to provide a faithful forecast accounting for realistic acceptance and background discrimination, we employ a realistic (at least at the level of theory forecasting) approach to the reconstruction and isolation of secondary vertices.

A key question is the extent to which future Higgs factories can improve on the LHC sensitivity to Higgs decays to LLPs, insofar as the number of Higgs bosons produced at the LHC will outstrip that of proposed Higgs factories by more than two orders of magnitude. Higgs decays to LLPs are sufficiently exotic that appropriate trigger and analysis strategies at the LHC should compensate for the higher background rate and messier detector environment. As we will see, there are two natural avenues for improved sensitivity at future lepton colliders: improved vertex resolution potentially increases sensitivity to LLPs with relatively short lifetimes, while lower backgrounds and a cleaner detector environment improves sensitivity to Higgs decays into lighter LLPs whose decay products are collimated.

This chapter is organized as follows: In Section 2 we present a simplified signal model for Higgs decays into pairs of long-lived particles, which in turn travel a macroscopic distance before decaying to quark pairs. We further detail the components of our simulation pipeline and lay out an analysis strategy aimed at eliminating the majority of Standard Model backgrounds. In Section 3 we translate this analysis strategy into the sensitivity of future lepton colliders to long-lived particles produced in Higgs decays as a function of the exotic Higgs branching ratio and the mass and decay length of the LLP. While these forecasts are generally applicable to any model giving rise to the signal topology, we additionally interpret the forecasts in terms of the parameter space of several motivated models in Section 4. We summarize our conclusions and highlight avenues for future development in Section 5.

6.2   Signal and analysis strategy

Exotic decays of the Higgs to long-lived particles encompass a wide variety of inter-
mediate and final states. The decay of the Higgs itself into LLPs can proceed through a
variety of different topologies. Perhaps the most commonly-studied scenario is the decay
of the Higgs to a pair of LLPs, $h \to XX$, though decays involving additional visible or
invisible particles (such as $h \to X +$ invisible or $h \to XX +$ invisible) are also possible.
The long-lived particles in turn may have a variety of decay modes back to the Standard
Model, including $X \to \gamma\gamma, jj, \ell\bar{\ell}$, or $jj\ell$, including various flavor combinations. These
decay modes may also occur in the company of additional invisible states. Moreover, a
given long-lived particle may possess a range of competing decay modes, as is the case for
LLPs whose decays back to the Standard Model are induced by mixing with the Higgs.

Our aim here is to be representative, rather than comprehensive, as each production
and decay mode for a long-lived particle is likely to require a dedicated search strategy.
For the purposes of this study, we adopt a simplified signal model in which the Higgs
decays to a pair of long-lived scalar particles X of mass m_X, which each decay in turn
to pairs of quarks at an average "proper decay length" $c\tau$.[30] Both the mass m_X and
proper decay length $c\tau$ are treated as free parameters, though they may be related in
models that give rise to this topology. For the sake of definiteness, for $m_X > 10$ GeV
we take a branching ratio of 0.8 to $b\bar{b}$ and equal branching ratios of 0.05 to each of
$u\bar{u}, d\bar{d}, s\bar{s}, c\bar{c}$, though the precise flavor composition is not instrumental to our analysis.
For $m_X \leq 10$ GeV we take equal branching ratios into each of the lighter quarks. We
further restrict our attention to Higgs factories operating near the peak of the $e^+e^- \to hZ$
cross section, for which the dominant production process will be $e^+e^- \to hZ$ followed
by $h \to XX$. The associated Z boson provides an additional invaluable handle for

---

[30]Of course, "proper decay length" is a bit of a misnomer, but we use it as a proxy for c times the
mean proper lifetime τ.

background discrimination. Here we develop the conservative approach of focusing on leptonic decays of the Z, though added sensitivity may be obtained by incorporating hadronic decays.

Given the signal, there are a variety of possible analysis strategies sensitive to Higgs decays to long-lived particles, exploiting various parts of a general-purpose detector. Tracker-based searches are optimal for decay lengths below one meter, with sensitivity to shorter LLP decay lengths all the way down to the tracker resolution. Timing information using timing layers between the tracker and electromagnetic calorimeter offers optimal coverage for slightly longer decay lengths, while searches for isolated energy deposition in the electromagnetic calorimeter, hadronic calorimeter, and muon chambers provides sensitivity to decay lengths on the order of meters to tens of meters. In principle, instrumenting the exterior of a general-purpose detector with large volumes of scintillator may lend additional sensitivity to even longer lifetimes. In this work we will focus on tracker-based searches at future lepton colliders, as these may be simulated relatively faithfully and ultimately are among the searches likely to achieve zero background while retaining high signal efficiency.

We define our signal model in `FeynRules` [401] and generate the signal $e^+e^- \rightarrow hZ \rightarrow XX + \ell\bar{\ell}$ at $\sqrt{s} = 240$ GeV using `MadGraph 5` [400]. Where appropriate, we will also discuss prospects for Higgs factories operating at $\sqrt{s} = 250$ GeV (potentially with polarized beams) such as the ILC by rescaling rates with the appropriate leading-order cross section ratios. In order to correctly simulate displaced secondary vertices, the decay of the LLP X and all unstable Standard Model particles is then performed in `Pythia 8` [438].

In addition to the signal, we consider some of the leading backgrounds to our signal process and develop selection cuts aimed at achieving a zero-background signal region. The most significant irreducible backgrounds from Standard Model processes include

$e^+e^- \to hZ$ with $Z \to \ell\bar{\ell}$ and $h \to b\bar{b}$ as well as $e^+e^- \to ZZ \to \ell\bar{\ell} + b\bar{b}$. Unsurprisingly, there are a variety of other Standard Model backgrounds, but they are typically well-controlled by imposing basic Higgsstrahlung cuts, and we do not simulate them with high statistics. In addition to irreducible backgrounds from hard collisions, there are possible backgrounds from particles originating away from the interaction point, including cosmic rays, beam halo, and cavern radiation; algorithmic backgrounds originating from effects such as vertex merging or track crossing; and detector noise. Such backgrounds are well beyond the scope of the current study, and will require dedicated investigation with full simulation of the proposed detectors.

Correctly emulating the detector response to LLPs using publicly-available fast simulation tools is notoriously challenging. In particular, we have found that the default clustering algorithms in the detector simulator `Delphes` [439] tends to cluster calorimeter hits from different secondary vertices into the same jets, significantly complicating the realistic reconstruction of secondary vertices. As such, we develop an analysis strategy using only ingredients from the `Pythia` output, although we do further run events through `Delphes` and utilize `ROOT` [440] for analysis.

We implement two distinct tracker-based analyses with complementary signal parameter space coverage, which we denote as the 'large mass' and 'long lifetime' pipelines. We shall eventually see that the former will be effective for $m_X \gtrsim 10$ GeV down to proper decay lengths $c\tau \gtrsim 1\mu$m, while the latter is able to push down in m_X by a factor of a few though is only fully effective for $c\tau \gtrsim 1$ cm. Full cut tables for both irreducible backgrounds and a variety of representative signal parameter points appear in Tables 6.1 and 6.2, respectively.

Table 6.1: Cut flow of the 'large mass' analysis for the CEPC with entries of acceptance × efficiency. The top set of rows gives the cut flow on 500k $Z(b\bar{b})Z(\ell\bar{\ell})$ events and 100k $h(b\bar{b})Z(\ell\bar{\ell})$ background events, which are used to confirm our analysis is in the no-background regime. The next sets of rows give cut flows on 5k signal events at representative parameter points, where the different columns are labeled by $m_X/\text{GeV}, c\tau/\text{m}$. The full row labels are given in the top set of rows and the labels below are abbreviations for the same cuts or selections.

Cut/Selection	ZZ Background	hZ Background
Dilepton Invariant Mass	0.97	0.98
Recoil Mass	0.006	0.94
Displaced Cluster ($\geq$ resolution)	0.004	0.94
Invariant Charged Mass (6 GeV)	0	0.00005
Invariant 'Dijet' Mass	0	0.00005
Pointer Track	0	0.00001

$m_X, c\tau$	7.5, 10^{-4}	7.5, 10^{-2}	7.5, 10^{0}	10, 10^{-4}	10, 10^{-2}	10, 10^{0}
$M_{\ell\ell}$	0.97	0.97	0.97	0.97	0.98	0.97
M_{recoil}	0.93	0.93	0.93	0.93	0.94	0.93
$|\vec{d}_{\text{cluster}}|$	0.93	0.93	0.41	0.93	0.94	0.50
M_{charged}	0.27	0.28	0.08	0.55	0.55	0.21
M_{cluster}	0.27	0.28	0.08	0.55	0.55	0.21
Pointer	0.25	0.28	0.08	0.50	0.55	0.21

$m_X, c\tau$	25, 10^{-4}	25, 10^{-2}	25, 10^0	50, 10^{-4}	50, 10^{-2}	50, 10^0
$M_{\ell\ell}$	0.97	0.97	0.98	0.97	0.98	0.97
M_{recoil}	0.92	0.92	0.93	0.92	0.92	0.93
$|\vec{d}_{\mathrm{cluster}}|$	0.92	0.92	0.80	0.92	0.92	0.93
M_{charged}	0.76	0.77	0.57	0.82	0.85	0.81
M_{cluster}	0.76	0.77	0.57	0.76	0.80	0.76
Pointer	0.73	0.76	0.57	0.75	0.77	0.76

Table 6.2: Cut flow of the 'long lifetime' analysis for the CEPC with entries of acceptance × efficiency. The top set of rows gives the cut flow on 500k $Z(b\bar{b})Z(\ell\bar{\ell})$ events and 100k $h(b\bar{b})Z(\ell\bar{\ell})$ background events, which are used to confirm our analysis is in the no-background regime. The next sets of rows give cut flows on 5k signal events at representative parameter points, where the different columns are labeled by $m_X/\mathrm{GeV}, c\tau/\mathrm{m}$. The full row labels are given in the top set of rows and the labels below are abbreviations for the same cuts or selections.

Cut/Selection	ZZ Background	hZ Background
Dilepton Invariant Mass	0.97	0.98
Recoil Mass	0.006	0.94
Displaced Cluster (≥ 3 cm)	0.004	0.62
Charged Invariant Mass (2 GeV)	0	0.002
'Dijet' Invariant Mass	0	0.002
Pointer Track	0	0.001
Isolation	0	0.00005

$m_X, c\tau$	2.5, 10^{-4}	2.5, 10^{-2}	2.5, 10^0	7.5, 10^{-4}	7.5, 10^{-2}	7.5, 10^0
$M_{\ell\ell}$	0.97	0.97	0.98	0.97	0.97	0.97
M_{recoil}	0.93	0.93	0.93	0.93	0.93	0.93
$|\vec{d}_{\text{cluster}}|$	0.21	0.89	0.15	0.41	0.89	0.41
M_{charged}	0	0.40	0.05	0	0.74	0.34
M_{cluster}	0	0.40	0.05	0	0.74	0.34
Pointer	0	0.40	0.05	0	0.74	0.34
Isolation	0	0.33	0.045	0	0.51	0.33
$m_X, c\tau$	15, 10^{-4}	15, 10^{-2}	15, 10^0	50, 10^{-4}	50, 10^{-2}	50, 10^0
$M_{\ell\ell}$	0.97	0.98	0.97	0.97	0.98	0.97
M_{recoil}	0.93	0.93	0.93	0.92	0.92	0.93
$|\vec{d}_{\text{cluster}}|$	0.59	0.87	0.65	0.64	0.69	0.92
M_{charged}	0.001	0.71	0.63	0	0.10	0.91
M_{cluster}	0.001	0.71	0.63	0	0.09	0.90
Pointer	0.001	0.65	0.60	0	0.08	0.84
Isolation	0.0002	0.42	0.58	0	0.05	0.77

As a first step in either analysis, we select Higgsstrahlung events by requiring that our events have an opposite sign electron (muon) pair in the invariant mass range $70 \leq M_{ee} \leq 110$ GeV ($81 \leq M_{\mu\mu} \leq 101$ GeV) and with recoil mass $M_{\text{recoil}}^2 \equiv \left((\sqrt{s}, \vec{0})^\mu - p_{\ell\ell}^\mu\right)^2$ in the range $120 \leq M_{\text{recoil}} \leq 150$ GeV, with $p_{\ell\ell}^\mu$ the momentum of the lepton pair. This allows us to limit our background considerations to the irreducible backgrounds mentioned above and cuts down severely on the $e^+ e^- \to ZZ$ background, as seen in Tables 6.1 and 6.2.

We next identify candidate secondary vertices using a depth-first 'clustering' algorithm, which roughly emulates that performed in the CMS search [441]. We perform this clustering using *all* particles in the event because at later points in the analysis we

need this truth-level assignment of neutral particles to clusters, but we expect that this (admittedly unrealistic) inclusion does not significantly modify the performance of this algorithm. Beginning with a single particle as the 'seed' particle for our algorithm, we look through all other particles in the event and create a 'cluster' of particles consisting of the seed particle and any others whose origins are within $\ell_{\text{cluster}} = 7\ \mu\text{m}$ (the projected tracker resolution of CEPC [413]) of the seed particle. We then add to that cluster any particles whose origins are within ℓ_{cluster} of any origins of particles in the cluster, and do this step iteratively until no further particles are added to the cluster. We then choose a new seed particle which has not yet been assigned to a cluster and begin this clustering process again. We repeat this process until all particles in the event have been assigned to clusters. We assign to each cluster a location $\vec{d}_{\text{cluster}}$ which is the average of the origins of all charged particles in the cluster. To ensure that our events contain displaced vertices, we impose a minimum bound on the displacement from the interaction point $|\vec{d}_{\text{cluster}}| > d_{\text{min}}$, and clusters satisfying this requirement constitute candidate secondary vertices. For our 'large mass' analysis we set d_{min} to be the impact parameter resolution ($\simeq 5\ \mu\text{m}$ for both CEPC and FCC-ee [413]), and so retain sensitivity to very short X lifetimes. For our 'long lifetime' analysis we set $d_{\text{min}} = 3$ cm, which removes the vast majority of clusters coming from B hadron decays in background events, as seen in Table 6.2. An upper bound $|\vec{d}_{\text{cluster}}| < r_{\text{tracker}}$ is imposed by the outer radius of the tracker, where $r_{\text{tracker}} = 1.81$m for CEPC and $r_{\text{tracker}} = 2.14$m for FCC-ee are proposed.

At this point an experimental analysis might sensibly examine dijets containing candidate secondary vertices and impose an upper bound on the dijet invariant mass to remove backgrounds coming from Standard Model H or Z decays. As discussed above we are limited to `Pythia` objects, but to mock up the (small) penalty to signal of such a selection we implement a selection on the total invariant mass of the clusters $M_{\text{cluster}}^2 \equiv (\sum_{i \in \text{ cluster}} p_i^\mu)^2$. Since this is truth-level information, to turn it into an analog

for the dijet invariant mass we apply a Gaussian smearing with a standard deviation of 10 GeV to account for the dijet resolution. We then select only candidate secondary vertices with $M_{\text{cluster}} < m_h/2$. This has no effect on background in our simulation pipeline as the background candidate secondary vertices are the result of hadronic decays, so the invariant masses of these clusters are not analogs for dijet invariant masses. As emphasized above, the imposition of this cut is strictly to account for possible selections that might appear in a more realistic experimental analysis.

While the total invariant mass of the clusters is not an experimental observable, the invariant mass of charged particles in the clusters $M_{\text{charged}} = (\sum_{i \text{ charged}} p_i^\mu)^2$ is experimentally accessible. For our 'large mass' analysis we select candidate secondary vertices with $M_{\text{charged}} > 6$ GeV, which gets rid of nearly all clusters from hadronic decays, as seen in Table 6.1. For the 'long lifetime' analysis, while the increased displacement requirement removes b hadrons it still allows c, s hadrons, and so we select $M_{\text{charged}} > 2$ GeV to address this, which Table 6.2 shows is again very effective.

Next we select the cluster closest to the beamline which passes the above selection requirements as our secondary vertex for the event. Choosing the closest one preferentially selects X decay clusters over hadronic decay clusters in the jets to which the X decays, though this can be fooled by a non-zero fraction of 'back-flowing' quarks in X decays (quarks with momenta pointing toward the beamline).

To remove displaced vertices coming from the decays of charged b hadrons we implement a 'pointer track' cut in both analyses as follows. For the cluster selected as the secondary vertex, we consider a sphere of radius $r = 0.5$ mm around the position $\vec{d}_{\text{cluster}}$. We look for any charged particles whose origins are outside this sphere and whose momenta (at the point at which they were created) point into it, and veto the event if there are any such particles. The main effect of this cut is to remove clusters which were produced from the decay of a charged hadron. The sphere size has been chosen to maxi-

mize this effect, though this allows a small effect on signal due to geometric coincidence. Since this cut is only on charged particles, roughly $\sim 30\%$ of background clusters are unaffected. For this cut we ignore the effect of the magnetic field in the tracker, which should not highly impact the trajectories on short scales.

For the 'long lifetime' analysis we further implement an 'isolation' cut to remove neutral hadronic background decays. Given the cluster selected as the secondary vertex, we consider the plane perpendicular to the sum of momenta of charged particles in the cluster which passes through $\vec{d}_{\text{cluster}}$. We project the paths of prompt (vertex within $3~\mu$m of the primary vertex, the planned CEPC vertex resolution [413]) charged particles onto this plane (again ignoring the magnetic field) and veto the event if any come within $R = 10$ cm of the position of the secondary vertex. This radius was chosen to maximally reduce background, and does have a deleterious effect on short decay lengths $\lesssim 10$ mm, as can be seen in Table 6.2. This cut is not perfectly effective at rejecting background due to the non-negligible presence of jets whose prompt components have neutral fraction 1.

6.3   Results and discussion

To confirm that our analysis pipelines put us in the zero-background regime we run both the 'long lifetime' and the 'large mass' analyses on 500k $e^+e^- \to Z(b\bar{b})Z(\ell\bar{\ell})$ events and 100k $e^+e^- \to h(b\bar{b})Z(\ell\bar{\ell})$ background events. For both pipelines we find that zero $e^+e^- \to Z(b\bar{b})Z(\ell\bar{\ell})$ events remain, while for $e^+e^- \to h(b\bar{b})Z(\ell\bar{\ell})$ we find efficiencies of 5×10^{-5} and 1×10^{-5} respectively. We then run each analysis on 5k signal events to get acceptance $\times$ efficiencies for each $(m_X, c\tau)$ point, for a selection of points with $m_X = 2.5$ from GeV to 50 GeV and $c\tau$ from 1 μm to 50 m. In Table 6.2 we give a cut table for both backgrounds and some representative signal parameter points for the 'long lifetime' analysis, and in Table 6.1 we do the same for the 'large mass' analysis.

In the zero-background regime, Poisson statistics rules out model points which predict 3 or more signal events to 95% confidence (or better) if no signal is detected. We may then find a projected 95% upper limit on branching ratio as

$$\mathrm{Br}(h \to XX)^{95} = \frac{N_{sig}}{\mathcal{L} \times \sigma(e^+e^- \to hZ) \times \mathrm{Br}(Z \to \ell\ell) \times A \times \varepsilon}, \qquad (6.1)$$

with $N_{sig} = 3$ and $A \times \varepsilon$ the result of our simulations. For both the CEPC and FCC-ee, the most recent integrated luminosity projections [413, 442] give $\mathcal{L} \times \sigma(e^+e^- \to hZ) = 1.1 \times 10^6$ Higgses produced.

In Figure 6.1 we show projected 95% upper limits on $\mathrm{Br}(h \to XX)$ as a function of X mass and proper decay length. While we plot separate lines for both CEPC and FCC-ee, we only use one set of signal events generated at $\sqrt{s} = 240$ GeV and only account for the difference in tracker radii, so these overlap entirely at smaller lifetimes. Approximate limits for the ILC can be obtained by multiplying the above branching ratio limits by a factor of ~ 1.8 (i.e. weakening the limit) to account for the leading order differences in center-of-mass energy, polarization, and integrated luminosity at the $\sqrt{s} = 250$ GeV ILC run, assuming comparable acceptance and efficiency. The ILC limits weaken slightly further for large decay lengths, as its proposed tracker radius is 1.25 m. Of course, adding the higher-energy ILC runs should significantly improve sensitivity given analyses suitable for the WW fusion production relevant at those energies.

For small masses we are only able to use the 'long lifetime' analysis, which requires large displacement from the beamline to cut out the SM b hadron background. As a result we only retain good sensitivity to X decay lengths comparable with the tracker size, though the fact that we only require one displaced vertex (out of two Xs per signal event) significantly broadens our sensitivity range. This fact also helps us retain efficiency at low masses, as we are able to get down to a projected branching ratio limit of 1×10^{-4}

for $m_X = 2.5$ GeV despite our 2 GeV cut on charged invariant mass of the decay cluster. For larger masses this cut has less effect, which allows it to push down to even lower branching ratios $\sim 5 \times 10^{-5}$.

The 'large mass' analysis begins working well for masses not far above the 6 GeV charged invariant mass cut and provides sensitivity to far shorter decay lengths, reaching all the way down to the impact parameter resolution and below. For $m_X = 10$ GeV, where we are aided by the boost factor, we project a limit of 1×10^{-4} for a proper decay length of 1 micron. The sensitivity to extremely small decay lengths drops for larger masses, but at $m_X = 50$ GeV we cross below the 10^{-4} threshold by 7.5 μm. For X masses high enough that the charged invariant mass cut does not remove a large amount of signal events, this analysis projects a branching ratio limit of $\sim 5 \times 10^{-5}$ across roughly the entire range of decay lengths corresponding to the geometric volume of the detector. There is a slight dip in sensitivity for $c\tau \sim 1$ mm, where the pair of dijets from the two X decays are most likely to overlap and trigger the cut on 'pointer' tracks.

The notable region of this parameter space to which our analyses do not provide good sensitivity is the low mass ($m_X \lesssim 6$ GeV) and short proper decay length ($c\tau \lesssim 1$ cm) regime. The difficulty is that, from the perspective of the tracker, the X here looks more and more like a neutral SM hadron. An analysis making use of the impact parameter distribution of particles in clusters may help here [441], but we leave this to future work. Taking advantage of calorimeter data to distinguish between clusters in single jets versus dijets is also likely to provide good sensitivity, but we again leave this to future exploration.

Broadly speaking, our results suggest a peak sensitivity of $\mathrm{Br}(h \to XX) \sim 5 \times 10^{-5}$, weakening to $\sim 10^{-4}$ for lower-mass LLPs. Significant additional improvement could be expected with the inclusion of hadronic Z decays, but this requires further study to ensure the control of corresponding Standard Model backgrounds. These limits are
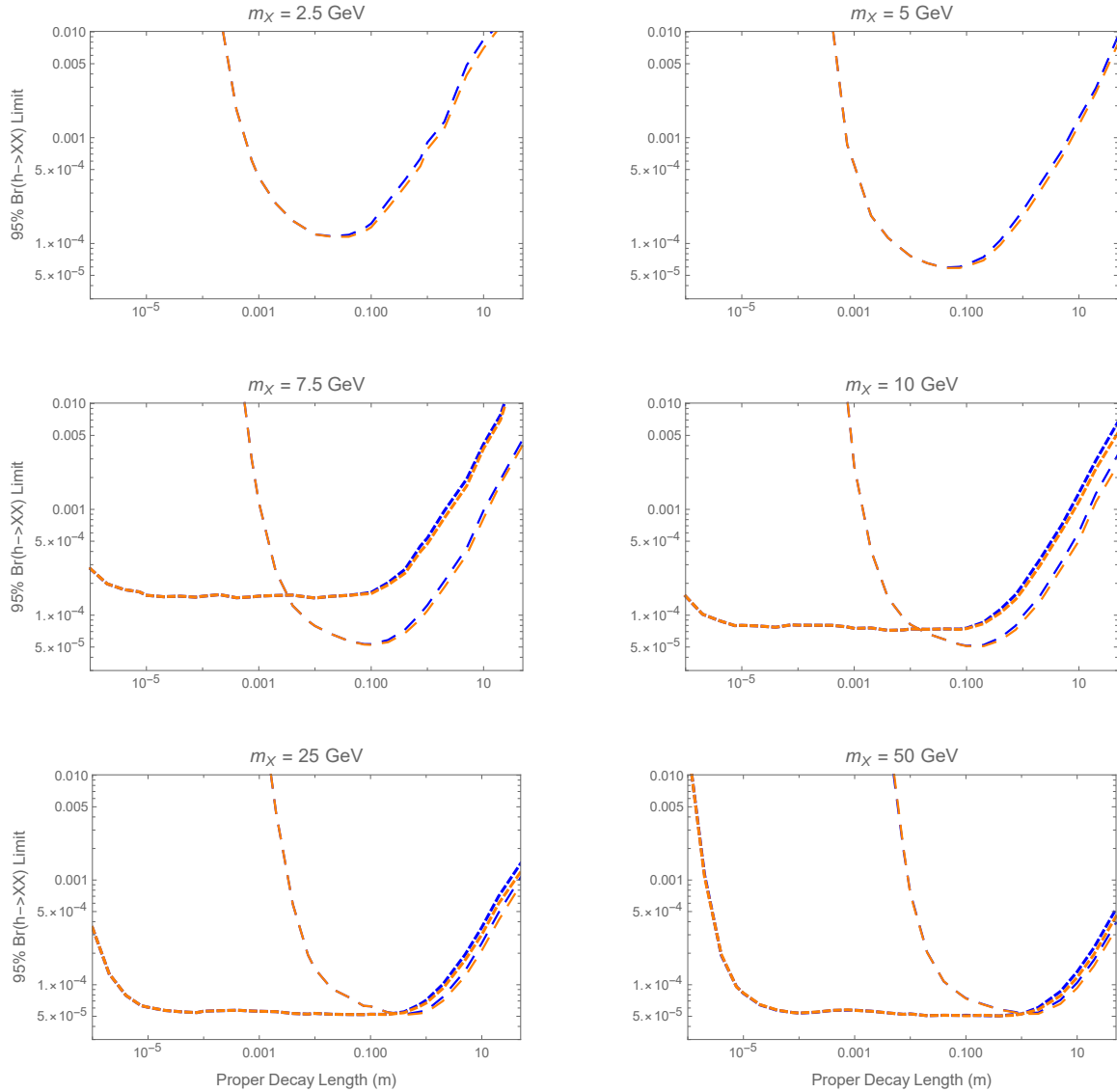
Figure 6.1: Projected 95% $h \to XX$ branching ratio limits as a function of proper decay length for a variety of X masses. Blue lines are for CEPC and orange lines are for FCC-ee, and where only one is visible they overlap. The larger dashes are the 'long lifetime' analysis and the smaller dashes are the 'large mass' analysis.

competitive with LHC forecasts based on conventional Higgs triggers [329, 328], noting that these latter forecasts assume zero background. However, the lepton collider limits are potentially superseded by an efficient CMS track trigger [432, 433] for higher-mass LLPs, again assuming zero background is achievable with high signal efficiency across a range of lifetimes. In this respect, the primary strengths of the Higgs factories in searching for exotic Higgs decays to LLPs are the potential to push down to shorter decay lengths and lighter LLPs. In particular, the relatively clean and low-background environment of lepton colliders should enable efficient LLP searches even when the LLP decay products become collimated, which remains a weakness of the corresponding LHC searches.

6.4   Signal interpretations

While the bounds presented in the previous section apply to any scenario in which the Higgs decays into pairs of long-lived particles which in turn decay (at least in part) into pairs of quarks, it is also useful to interpret these bounds in the context of specific models that relate the Higgs branching ratio to LLPs (and the LLP lifetime) to underlying parameters. This illustrates the potential for LLP searches at future lepton colliders to constrain motivated scenarios for physics beyond the Standard Model and allows us to explore the potential complementarity between LLP searches and precision Higgs coupling measurements. To this end, we consider the implications of the LLP limits presented here in the context of both the original Higgs portal Hidden Valley model and a variety of models of neutral naturalness.

6.4.1   Higgs portal

As a general proxy model for Higgs decays into LLPs, we first consider the archetypal Higgs portal Hidden Valley [423]. This entails the extension of the Standard Model by an additional real singlet scalar ϕ, which couples to the Standard Model through the Higgs portal [443, 444, 445] via

$$
\begin{aligned}
\mathcal{L} \supset &-\frac{1}{2}(\partial_\mu \phi)^2 - \frac{1}{2}M^2\phi^2 - A|H|^2\phi - \frac{1}{2}\kappa|H|^2\phi^2 \\
&-\frac{1}{3!}\mu\phi^3 - \frac{1}{4!}\lambda_\phi\phi^4 - \frac{1}{2}\lambda_H|H|^4.
\end{aligned}
\tag{6.2}
$$

If ϕ respects a $\mathbb{Z}_2$ symmetry under which $\phi \to -\phi$, this additionally sets $\mu = A = 0$, such that the singlet scalar only couples to the Standard Model via the quartic interaction $|H|^2\phi^2$. After electroweak symmetry breaking, in unitary gauge $H = \left(0, \frac{1}{\sqrt{2}}(h+v)\right)$, but the CP-even scalars h and ϕ do not mix. Nonetheless, the quartic interaction nonetheless provides a significant portal for the production of ϕ, as ϕ may be pair produced via the decay $h \to \phi\phi$ for $m_\phi < m_h/2$. Of course, ϕ is stable if the $\mathbb{Z}_2$ symmetry is exact, rendering it a potential (albeit highly constrained) dark matter candidate [446, 447, 448].

This model gives rise to long-lived particle signatures [423] if the $\mathbb{Z}_2$ is broken by a small amount, such that $A \neq 0$ but e.g. $A^2/M^2 \ll \kappa$. The relative smallness of A is technically natural, as the $\mathbb{Z}_2$ symmetry is restored when $A \to 0$. This then leads to mass mixing between the CP even scalars. As long as A is small compared to M and v, the mass eigenstates consist of an SM-like Higgs h_{SM} and a mostly-singlet scalar s, related to the gauge eigenstates by

$$
\begin{aligned}
h_{\mathrm{SM}} &= h\cos\theta + \phi\sin\theta \tag{6.3}\\
s &= -h\sin\theta + \phi\cos\theta, \tag{6.4}
\end{aligned}
$$

where $\theta \ll 1$ is the mixing angle. There are now two parametrically distinct processes: pair production of the scalar s via Higgs decays, governed by the size of the $\mathbb{Z}_2$-preserving coupling κ, and decay of the s scalar back to the Standard Model, governed by the size of the $\mathbb{Z}_2$-breaking coupling A. In the limit of small mixing, the former process is of order

$$\Gamma(h \to ss) \approx \frac{\kappa^2 v^2}{32\pi m_h} \sqrt{1 - 4\frac{m_s^2}{m_h^2}}, \tag{6.5}$$

where we are neglecting subleading corrections proportional to $\lambda_H \sin^2 \theta$. The latter process proceeds into whatever Standard Model states Y are kinematically available, with partial widths

$$\Gamma(s \to YY) = \sin^2 \theta \times \Gamma(h_{\mathrm{SM}}[m_s] \to YY), \tag{6.6}$$

where $h_{\mathrm{SM}}[m_s]$ denotes a Standard Model-like Higgs of mass m_s. This naturally leads to a scenario in which the s scalars may be copiously produced via Higgs decays but travel macroscopic distances before decaying back to Standard Model particles.

This scenario may be constrained not only by direct searches for Higgs decays to LLPs (with the scalar s playing the role of the LLP), but also by precision Higgs coupling measurements. Higgs coupling deviations in this scenario arise from two parametrically distinct effects: tree-level deviations proportional to θ^2 due to Higgs-singlet mixing, and one-loop deviations proportional to κ due to s loops. Both effects result in a universal modification of Higgs couplings, which is best constrained at lepton colliders via the precision measurement of the $e^+ e^- \to hZ$ cross section [449, 450]. The net deviation in the $e^+ e^- \to hZ$ cross section due to these effects in the limit of small mixing is

$$\frac{\delta\sigma_{hZ}}{\sigma_{hZ}^{\mathrm{SM}}} \approx -\theta^2 - \mathrm{Re} \left. \frac{d\mathcal{M}_{hh}}{dp^2} \right|_{p^2 = m_h^2}, \tag{6.7}$$
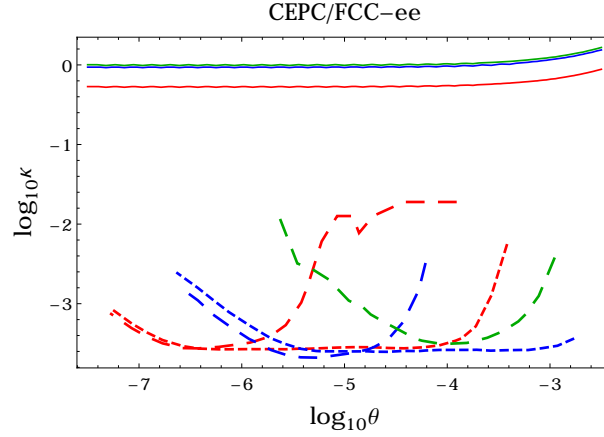
Figure 6.2: Projected 95% limits on the Higgs portal Hidden Valley model in the κ, θ plane for three choices of m_s; green lines correspond to $m_s = 2.5$ GeV, blue to $m_s = 10$ GeV, and red to $m_s = 50$ GeV. The solid lines are the projected lower limits from precision Higgs measurements, taking the CEPC projections [413] for definiteness. The dashed lines are projected limits from this work, which are essentially identical for CEPC and FCC-ee. Long dashes are from the 'long lifetime' analysis and short dashes from the 'large mass' analysis.

where the radiative correction [450]

$$\frac{d\mathcal{M}_{hh}}{dp^2}\bigg|_{p^2=m_h^2} = -\frac{1}{16\pi^2}\frac{\kappa^2 v^2}{2m_h^2} \tag{6.8}$$

$$\times \left(1 + \frac{4m_s^2}{m_h^2}\sqrt{\frac{m_h^2}{m_h^2 - 4m_s^2}}\tanh^{-1}\left[\sqrt{\frac{m_h^2}{m_h^2 - 4m_s^2}}\right]\right)$$

is approximated at $\theta = 0$. Either effect can dominate depending on the relative size of A/M and κ.

Constraints from a direct search for Higgs decays to LLPs and precision Higgs measurements as a function of the underlying parameters θ and κ are shown in Figure 6.2 for the illustrative benchmarks $m_s = 2.5, 10$, and 50 GeV. Unsurprisingly, in the regime where s is long-lived, the bounds from precision Higgs coupling measurements are modest and direct searches provide the leading sensitivity.

248

6.4.2   Neutral naturalness

Higgs decays to LLPs are also motivated by naturalness considerations, arising frequently in models of neutral naturalness that address the hierarchy problem with SM-neutral degrees of freedom [267, 276]. In these models, partially or entirely SM-neutral partner particles that couple to the Higgs boson are charged under an additional QCD-like sector. Confinement in the additional QCD-like sector leads to a variety of bound states that couple to the Higgs and may be pair-produced in exotic Higgs decays with predictive branching ratios. The bound states with the same quantum numbers as the physical Higgs scalar typically decay back to the Standard Model by mixing with the Higgs. These decays occur on length scales ranging from microns to kilometers, making them a motivated target for LLP searches at colliders [268, 328].

For simplicity, here we will restrict our focus to scenarios with the sharpest predictions for the Higgs branching ratio to LLPs. In these cases, the LLPs in question are typically glueballs of the additional QCD-like sector, of which the $J^{PC} = 0^{++}$ is typically the lightest. The coupling of the SM-like Higgs to these LLPs is predominantly due to top partner loops, for which the scales and couplings are directly related to the naturalness of the parameter space. In the Fraternal Twin Higgs [268], the entirely SM-neutral fermionic partners of the top quark induce Higgs couplings to twin gluons, which then form glueballs; the 0^{++} states are the lightest in the twin QCD spectrum only if the other twin quarks are sufficiently heavy. In addition, there are tree-level deviations in Higgs couplings due to the pseudo-goldstone nature of the SM-like Higgs. In Folded SUSY [273], the scalar top partners carry electroweak quantum numbers, leading to radiative corrections to standard Higgs decays as well as the existence of exotic decay modes. Loops of the scalar top partners again induce Higgs couplings to twin gluons, and without light folded quarks the 0^{++} glueball is generically the lightest state in the folded

QCD spectrum. While there are no tree-level Higgs coupling deviations in this case, the electroweak quantum numbers of the scalar top partners induce significant corrections to the branching ratio $h \to \gamma\gamma$. Finally, in the Hyperbolic Higgs [281] (see also [282]), the scalar top partners are entirely SM-neutral, and induce couplings to 0^{++} glueballs that are generically the lightest states in the hyperbolic QCD spectrum. As with the Fraternal Twin Higgs, however, there are also tree-level Higgs coupling deviations due to mass mixing among CP-even neutral scalars.

In each of these scenarios, the branching ratio of the SM-like Higgs can be parameterized as follows:

$$
\mathrm{Br}(h \to 0^{++}0^{++}) \approx \left(2v^2 \frac{\alpha'_s(m_h)}{\alpha_s(m_h)} \left[\frac{y^2}{M^2} \right] \right)^2
$$
$$
\times \mathrm{Br}(h \to gg)_{\mathrm{SM}} \times \sqrt{1 - \frac{4m_0^2}{m_h^2}} \tag{6.9}
$$

Here α'_s denotes the coupling of the additional QCD-like sector (whether twin, folded, or hyperbolic), which is necessarily of the same order as the SM QCD coupling α_s, and m_0 is the mass of the glueball, which is determined in terms of the QCD-like confinement scale. Adopting the schematic notation of [328], the parameter $\left[\frac{y^2}{M^2} \right]$ encodes the model-dependence of the Higgs coupling to pairs of gluons in the QCD-like sector, with

$$
\left[\frac{y^2}{M^2} \right] \approx \begin{cases} -\frac{1}{2v^2} \frac{v^2}{f^2} & \text{Fraternal Twin Higgs} \\ \frac{1}{4v^2} \frac{m_t^2}{m_{\tilde{t}}^2} & \text{Folded SUSY} \\ \frac{1}{4v^2} \frac{v}{v_{\mathcal{H}}} \sin\theta & \text{Hyperbolic Higgs} \end{cases} \tag{6.10}
$$

For the Fraternal Twin Higgs, f denotes the overall twin symmetry-breaking scale $f^2 = v^2 + v'^2$ in terms of the SM weak scale v and the fraternal weak scale v'. For Folded SUSY, $m_{\tilde{t}}$ denotes the mass of the scalar top partners, neglecting possible mixing effects.

For the Hyperbolic Higgs, $v_{\mathcal{H}}$ is the hyperbolic scale and $\tan\theta \approx \frac{v}{v_{\mathcal{H}}}$ encodes tree-level mixing effects. In each case, the scales appearing in the effective coupling are related to the fine-tuning of the model, drawing a direct connection between the Higgs exotic branching ratio and the naturalness of the weak scale.

In each case, the 0^{++} glueballs of the additional QCD-like sector decay back to the Standard Model by mixing with the SM-like Higgs, with a partial width to pairs of SM particles Y given by

$$\Gamma(0^{++} \to YY) = \left(\frac{1}{12\pi^2} \left[\frac{y^2}{M^2} \right] \frac{v}{m_h^2 - m_0^2} \right)^2 \left(4\pi\alpha_s^B F_{0^{++}}^S \right)^2$$
$$\times \Gamma(h_{SM}[m_0] \to YY), \qquad (6.11)$$

where $4\pi\alpha_s^B F_{0^{++}}^S \approx 2.3m_0^3$ and, as before, $h_{\mathrm{SM}}[m_0]$ denotes a Standard Model-like Higgs of mass m_0.

Constraints on each model from a direct search for Higgs decays to LLPs and precision Higgs measurements are shown in Figure 6.3 as a function of the LLP mass m_0 and the relevant scale (f, $v_{\mathcal{H}}$, and $m_{\tilde{t}}$, respectively). For precision Higgs measurements we use the CEPC projections from [413]. In the Fraternal Twin Higgs and Hyperbolic Higgs, the dominant indirect constraint is from σ_{Zh}, while for Folded SUSY it is from $\mathrm{Br}(h \to \gamma\gamma)$. For both the Fraternal Twin Higgs and the Hyperbolic Higgs, tree-level Higgs coupling deviations make precision Higgs measurements the strongest test of the model. However, the sensitivity of LLPs searches provides valuable complementarity in the event that Higgs coupling measurements yield a discrepancy from Standard Model predictions. In particular, the size of an observed Higgs coupling deviation would single out the relevant overall mass scale (f or $v_{\mathcal{H}}$), providing a firm target for LLP searches that would then validate or falsify these models as an explanation of the deviation. Note also that in the Fraternal Twin Higgs there may be additional contributions to the Higgs branching
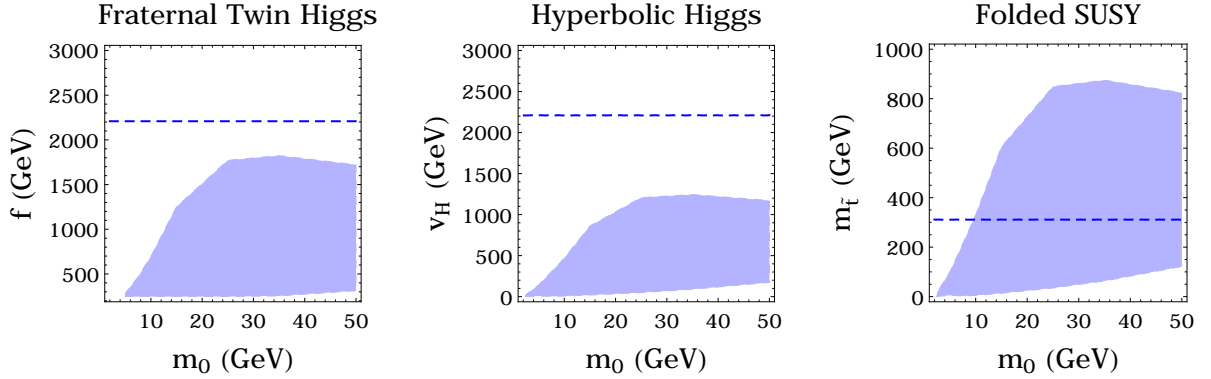
Figure 6.3: Projected 95% limits on the underlying scale as a function of the LLP mass m_0 in three models of neutral naturalness: the Fraternal Twin Higgs (f), the Hyperbolic Higgs ($v_{\mathcal{H}}$), and Folded SUSY ($m_{\tilde{t}}$). The blue dashed line denotes the limit coming from precision Higgs coupling measurements, taking for definiteness the CEPC projections from [413]. For the Fraternal Twin Higgs and Hyperbolic Higgs, the dominant indirect constraint is from σ_{Zh}, while for Folded SUSY it is from $\mathrm{Br}(h \to \gamma\gamma)$. The shaded region denotes the projected limits from direct LLP searches obtained in this work.

ratio into LLPs coming from the production of twin bottom quarks, which could lead to sensitivity in the LLP search comparable to that of Higgs couplings. In the case of Folded SUSY, the absence of tree-level Higgs coupling deviations and the relatively weaker constraints on $\mathrm{Br}(h \to \gamma\gamma)$ make the LLP search the leading test of this model at Higgs factories.

6.5 Conclusion

The exploration of exotic Higgs decays is an integral part of the physics motivation for future lepton colliders. New states produced in these exotic Higgs decays may themselves decay on a variety of length scales, necessitating a range of search strategies. While considerable attention has been devoted to the reach of future lepton colliders for promptly-decaying states produced in exotic Higgs decays, the reach for long-lived particles is relatively unexplored.

252

In this paper we have made a first attempt to study the reach of proposed circular Higgs factories such as CEPC and FCC-ee (as well as approximate statements for the $\sqrt{s} = 250$ GeV run of the ILC) for long-lived particles produced in exotic Higgs decays, focusing on the pair production of LLPs and their subsequent decay to pairs of quarks. We have developed a realistic tracker-based search strategy motivated by existing LHC searches that entails the reconstruction of displaced secondary vertices. Rather than relying on existing public fast simulation tools, which do not necessarily give a sensible parameterization of signal and background efficiencies for long-lived particle searches, we have implemented a realistic approach to clustering and isolation. This allows us to characterize some of the leading irreducible Standard Model backgrounds to our search and determine reasonable analysis cuts necessary for a zero-background analysis. We obtain forecasts for the potential reach of CEPC and FCC-ee on the Higgs branching ratio to long-lived particles with a range of lifetimes. The projected reach is competitive with LHC forecasts and potentially superior for lower LLP masses and shorter lifetimes. In addition to our branching ratio limits, which may be freely interpreted in a variety of model frameworks, we interpret our results in the parameter space of a Higgs portal Hidden Valley and various incarnations of neutral naturalness, demonstrating the complementarity between direct searches for LLPs and precision Higgs coupling measurements.

There are a variety of directions for future work. While we have attempted to investigate some of the leading irreducible backgrounds and impose realistic cuts, we have not attempted to estimate possible backgrounds coming from cosmic rays; algorithmic, detector, or beam effects; or other contributions. Our tracker-based analysis has focused on Higgs decays to pairs of hadronically-decaying LLPs, but a comprehensive picture of exotic Higgs decays would also suggest the investigation of Higgs decays to various LLP combinations as well as the consideration of additional LLP decay modes. Moreover, tracker-based searches for displaced vertices are but one of many possible avenues

to discover long-lived particles. Analogous searches based on timing or on isolated energy deposition in outer layers of the detector (including either the electromagnetic or hadronic calorimeter, the muon chambers, or potentially instrumented volumes outside of the main detector) would be valuable for building a complete picture of LLP sensitivity across a range of lifetimes.

More broadly, it is an ideal time to study the potential sensitivity of future Higgs factories to long-lived particles, as the results are likely to inform the design of detectors for these proposed colliders. This is a necessary step in motivating the physics case of future Higgs factories and ensuring that they enjoy optimal coverage of possible physics beyond the Standard Model.

Chapter 7

New Trail for Naturalness

7.1   Introduction

At its heart, the electroweak hierarchy problem is a question of how an infrared (IR) scale can emerge from an ultraviolet (UV) scale without fine-tuning of UV parameters. Given the sensitivity of the Standard Model Higgs mass to UV scales, the expectation of effective field theory (EFT) is that the two should coincide. Conventional solutions to the hierarchy problem introduce both symmetries that control UV contributions to the Higgs potential and dynamics that generate IR contributions, leading to considerable structure at the weak scale and correspondingly sharp experimental tests. Ongoing exploration of the weak scale has given no evidence for these solutions, despite their theoretical soundness.

In the face of increasingly powerful LHC data in excellent agreement with the Stan-

dard Model, it's worth taking seriously the possibility that Nature may be leading us to the conclusion that *there is no new physics at the weak scale*. While this is often taken to suggest the existence of considerable fine-tuning in the Higgs potential, here we pursue an alternative idea. Perhaps the apparent violation of EFT expectations at the weak scale is a sign of the breakdown of EFT itself. We'll use the broad term 'UV/IR mixing' to denote any effects that the UV has on low-energy physics which goes past that expected in EFT.

In this work we pursue the idea that such UV/IR mixing may have more direct effects on the SM by considering noncommutative field theory (NCFT) as such a toy model. These theories model physics on spaces where translations do not commute [311, 312], and have many features amenable to a quantum gravitational interpretation — indeed, noncommutative geometries have been found arising in various limits of string theory [313, 314, 315, 316].[31]

This noncommutativity bears out the general expectation that the general-relativistic notion of spacetime should break down in a theory of quantum gravity [463]. Its realization here leads directly both to UV/IR mixing in the form of a violation of decoupling and to nonlocal effects in interactions. This gives rise to many interesting effects, but particularly fascinating for our purposes is that UV divergences present in the S-matrix elements of QFTs on commutative spaces can be transmogrified into *new infrared poles* in the corresponding field theory on noncommutative space [464]. An effective field theorist living in a noncommutative space would have no way to understand the appearance of this infrared scale; its existence is intrinsically linked to the geometry of spacetime and to the far UV of the theory. Such an effective field theorist would see a surprising lack

---

[31]Noncommutative branes arising in gauge theory matrix models have also been found to contain emergent gravitational effects, and so have been suggested as novel quantum theories of gravity [452, 453, 454, 455, 456, 457, 458, 459, 460, 461]. We do not pursue this perspective here, but refer the reader to [462] for a review of this approach.

of new physics accompanying this pole to explain its presence.

It is clear from the outset that the direct application of NCFT to understand the hierarchy problem is immediately hindered by the Lorentz invariance violation which is inherent to these theories. Precisely how fatal this might be is not entirely clear; results regarding the extent to which 'generic' Lorentz violation is empirically ruled out [465] are partly circumvented here by the fact that the Lorentz violation is *not* generic, but comes as part of some larger structure. In this case the novel effects of UV/IR mixing in fact only appear in nonplanar loop diagrams [466] and care is required when interpreting EFT constraints on Lorentz violation — a point we will emphasize in Section 7.2. Even so, it is difficult to imagine that observed properties of the weak scale and the wide range of constraints on Lorentz violation leave room for NCFT to be directly relevant to puzzles of the Standard Model.

Thus we make no claim about having solved the hierarchy problem. The value of this work is in the exploration of this toy model of UV/IR mixing, which possesses the intriguing feature that ultraviolet dynamics generate a scale whose lightness would be baffling to an effective field theorist. As this is the only model (of which we are aware) with this feature — and this feature, at the level of words, increasingly matches the experimental situation with the Higgs — it's worth understanding its appearance in as much detail as possible.

To make this work self-contained for the contemporary particle theorist, we begin with an extensive introduction. In Section 7.2, we review quantum field theory on non-commutative spaces with an emphasis on the violation of EFT expectations. In Section 7.3 we use this technology to go over the classic result of [464] which first identified this emergent infrared pole in a Euclidean ϕ^4 theory. We compute also the effect in dimensional regularization to evince the regularization-independence of the UV/IR mixing effects.

In Section 7.4 we ask how general the effect of UV/IR mixing is within NCFT, which leads us to study noncommutative Yukawa theory in detail. We find that the scalar propagator again develops a new infrared pole at one loop, in contrast with previous work. Intriguingly, the pole in this case is accessible in s-channel scattering in the Lorentzian theory, making Yukawa theory a promising setting for probing phenomenological consequences of UV/IR mixing.

In Section 7.5 we upgrade our model to the softly-broken Wess-Zumino model to study the interplay between UV-finiteness and UV/IR mixing effects. When the fermion is kept in the spectrum of the theory below the cutoff, the lack of UV sensitivity of the field theory removes the light pole. As the fermion is taken above the cutoff, an effective theorist again sees effects past those observed in Wilsonian EFT. These results are expected, but this model affords us a concrete demonstration that UV/IR mixing can only have interesting low-energy effects if the field theory is UV sensitive, and puts this naturalness strategy in stark contrast to conventional approaches. Of course, this also makes addressing the hierarchy problem with UV/IR mixing a potentially Pyrrhic victory: to generate an IR scale, the field theory alone cannot be fully predictive.

Finally, in Section 7.6 we examine the appearance of the emergent light pole in NCFT from more general arguments, so as to ascertain the relative importance of nonlocality and Lorentz-violation for these effects. The conclusion is inevitably that in this case the two are inexorably linked, and no strong conclusion about the possibility of finding a light pole in a theory with only one or the other is available. However, we provide some direction toward future explorations into both of these possibilities. We wrap up in Section 7.7.

7.2   Noncommutative Field Theory

In this section we review the salient features of the formulation of noncommutative field theories and the standard formalism for studying their perturbative physics. Useful general references for this background include [467, 468]. Readers familiar with NCFT may wish to skip to Section 7.3, but we emphasize that our interest is necessarily non-perturbative in the parameter controlling the noncommutativity, unlike much of the earlier phenomenological literature.

Physics on noncommutative spaces involves the introduction of a nonzero commutator between position operators

$$[\hat{x}_\mu, \hat{x}_\nu] = i\theta_{\mu\nu}, \tag{7.1}$$

where we will refer to $\theta_{\mu\nu} = -\theta_{\nu\mu}$ as the noncommutativity tensor, and we emphasize that it is covariant under Lorentz transformations. So while it does break Lorentz invariance, it only does so in the way that turning on a magnetic field in your lab chooses a preferred frame, and it can indeed be thought of as simply a background field. This basic definition is reminiscent of the introduction of a nonzero commutator in passing from classical mechanics to quantum mechanics. Indeed much of the structure is precisely analogous, including importantly the construction of noncommutative versions of familiar commutative theories via a quantization map. At an even more basic level, the above nonzero commutator induces an uncertainty relation

$$\Delta\hat{x}_\mu \Delta\hat{x}_\nu \geq \frac{|\theta_{\mu\nu}|}{2}, \tag{7.2}$$

which immediately makes apparent the presence of UV/IR mixing in this theory. If you attempt to create a wavepacket which is very small in one direction it will necessarily be elongated in another, and so we see already the non-trivial mixing of UV and IR modes.

This clearly violates the separation of scales which is baked in to EFT. Thus purely from the defining relation of noncommutative geometry, we see already an indication that noncommutative theories should violate EFT expectations.

Field theories on this space may be conveniently formulated in terms of fields that are functions of *commuting* coordinates imbued with a new field product, known as a Groenewold-Moyal product (or star-product), with position-space representation

$$f(x) \star g(x) = \exp\left(\frac{i}{2}\theta_{\mu\nu}\partial_y^\mu\partial_z^\nu\right) f(y)g(z)\Big|_{y=z=x} = f(x)\exp\left(\frac{i}{2}\overleftarrow{\partial}^\mu\theta_{\mu\nu}\overrightarrow{\partial}^\nu\right) g(x). \quad (7.3)$$

We derive this procedure in Appendix B. It is important to observe that this is a nonlocal product, since it contains an infinite series of derivative operators. So we see again that one of the tenets of EFT has been violated by our basic definition of field theory on noncommutative spaces.

With this in hand we may now write down noncommutative versions of familiar theories *in terms of commuting coordinates*, which will then allow us to use normal QFT methods to analyze them. First note that this noncommutative quantization will not affect the quadratic part of the tree-level action due to momentum conservation and the antisymmetry of the noncommutativity tensor. For the interacting part of the action the effects of noncommutative quantization are not so trivial, but are easy to analyze classically. As an example, for a simple ϕ^n theory we find

$$\mathcal{L}_{\text{int}}^{(NC)} = \frac{\lambda}{n!} \overbrace{\phi(x) \star \phi(x) \star \cdots \star \phi(x)}^{\text{n copies}}. \quad (7.4)$$

Note, importantly, that the star-product has endowed our vertices with a notion of ordering, as it is only cyclically invariant. If we now Fourier transform the action to momentum space, we find that we can account for the effects of quantization on the

tree-level action with a simple modification of the momentum-space vertex factor:

$$\tilde{V}(k_1, \ldots, k_n) = \delta(k_1 + \cdots + k_n) \exp\left(\frac{i}{2} \sum_{i<j}^{n} k_i^\mu k_j^\nu \theta_{\mu\nu}\right). \tag{7.5}$$

A word of caution is in order. We can now express the action in momentum space as

$$\mathcal{S}_{\text{int}}^{(NC)} = \frac{\lambda}{n!} \int \left(\prod_i^n \mathrm{d}^4 k_i\right) \delta(k_1 + \cdots + k_n) \phi(k_1)\phi(k_2)\ldots\phi(k_n) \exp\left(\frac{i}{2} \sum_{i<j}^{n} k_i^\mu k_j^\nu \theta_{\mu\nu}\right), \tag{7.6}$$

and so — as good effective field theorists — we may be tempted to expand the exponential for small momenta $\sim |k^2| |\theta| \ll 1$. Indeed, doing so would give us a series of irrelevant operators which would correct the leading interaction. However, once the theory is truncated at some finite order in θ, we are left with a perfectly local EFT. In other scenarios where an infinite series of operators appears, this is a valid approximation procedure and allows one to calculate the leading corrections a theory predicts. But here our definition of NCFT introduces UV/IR mixing which we expect to violate EFT expectations. Truncating the series removes these effects entirely, and a theory so defined no longer has anything to do with NCFT — at least not in the effects we will be interested in, which are nonperturbative in θ as we shall see explicitly in the following sections. There has been much work expended on these 'noncommutative-inspired' theories, but they do not contain UV/IR mixing, and do not capture the most striking and most interesting features of physics on a noncommutative space, from our perspective.[32]

With that in mind, we may now proceed to do perturbative quantum field theory calculations, but we must worry about keeping track of all the phases from each of the vertices. In fact there is another simplification that occurs, as found by Filk [466], which

---

[32]We are not the first to issue a warning of this sort — see e.g. [469, 470] in the context of connecting noncommutativity to the real world, and [471] which discusses the general case of nonlocal interactions.

allows us to simplify the process of finding the phase factor for a diagram to a graph-topological statement. Filk proved two simple rules for the phase factors:

1. An internal line which ends on two different vertices can be contracted while keeping the ordering of the other lines fixed.

$$\tilde{V}\left(k_1, \ldots, k_{n_1}, p\right) \tilde{V}\left(-p, k_{n_1+1}, \ldots, k_{n_2}\right) = \tilde{V}\left(k_1, \ldots, k_{n_2}\right) \delta(k_1 + \cdots + k_{n_1} + p) \quad (7.7)$$

2. A loop which doesn't cross any lines can be eliminated. Note that the fixed ordering of the lines at a vertex means that we can now meaningfully speak of lines which do or don't cross each other.

$$\tilde{V}\left(k_1, \ldots, k_{n_1}, p, k_{n_1+1}, \ldots, k_{n_2}, -p\right) = \tilde{V}\left(k_1, \ldots, k_{n_1}, k_{n_1+1}, \ldots, k_{n_2}\right) \quad \text{if} \sum_{i=n_1+1}^{n_2} k_i = 0$$
$$(7.8)$$

The proof of these facts relies only on the antisymmetry of $\theta^{\mu\nu}$ and the fact that each vertex contains a momentum-conserving delta function. We may make use of this to simply find the phase factor of any Feynman diagram. Using the first rule, we can reduce any diagram to a single vertex, which is a rosette of the external lines and closed loops. The second rule allows us to eliminate loops which don't cross other lines.

If the graph was planar (including, importantly, any tree-level graph), then by definition all loops can be eliminated. So all contributions to phase factors from internal lines cancel, and we're only left with an overall phase corresponding to the ordering of the external lines, which has remained fixed throughout the reduction process.

For a nonplanar graph, in this representation it is easy to see that we only pick up phase factors from lines which cross. The loop gives vertex legs with $\pm p^\mu$, and for an external line which doesn't cross this loop, both loop legs will be on the same side of it

in the cyclic ordering, and so the two terms will cancel in the sum. Only for an external line which crosses it are the $\pm p$ on different sides, and so the antisymmetry of θ will make the two negative signs cancel to give a coherent phase for this vertex. Thus we define I_{ij}, the intersection matrix of an oriented graph:

$$I_{ij} = \begin{cases} 1 & \text{line } j \text{ crosses } i \text{ from right} \\ -1 & \text{line } j \text{ crosses } i \text{ from left} \\ 0 & \text{line } j \text{ does not cross } i \end{cases} \tag{7.9}$$

Then for any graph $\mathcal{G}$, the contribution $\Gamma(\mathcal{G})$ of the phase factors is just

$$\Gamma(\mathcal{G}) = \tilde{V}\left(\{\text{external momenta}\}\right) \times \exp\left(\frac{i}{2}\sum_{ij} I_{ij} k_i \wedge k_j\right), \tag{7.10}$$

where we've defined $k_i \wedge k_j \equiv k_i^\mu \theta_{\mu\nu} k_j^\nu$.

In what follows we will omit the overall external phase when evaluating diagrams, as it will not be important for our purposes. We have now simplified perturbative field theory on noncommutative spaces down to the simple task of marking line-crossings, at least at the level of writing down integrands of amplitudes. The triviality of this task for tree-level graphs leads to the interesting feature that tree-level amplitudes on noncommutative spaces are the same as on commutative manifolds, and it is only at loop-level that we find deviations. We will see in the next section that the loop integration will bring surprising features.

An important issue for the interpretation of NCFTs is that of their unitarity. There is no problem in Euclidean space, but for Lorentzian spacetimes with noncommutativity in the time directions ('timelike' or 'space-time' noncommutativity when $-k^\mu \theta_{\mu\rho} \theta^{\rho\nu} k_\nu \equiv k \circ k < 0$ is allowed), one may find a breakdown of unitarity by taking cuts of one-loop

diagrams [472, 473].[33] This may be interpreted physically as being due to the production of tachyonic states, which if added to the Fock space of the theory result in a formal restoration of the cutting relations whilst making the nonunitarity explicit [476].

This failure of unitarity is well-understood from the stringy perspective. Spatial noncommutativity appears from a background magnetic field and the field theory limit to a spacelike NCFT is smooth [315]. In the case of timelike noncommutativity, however, approaching the field theory limit forces an electric field to supercritical values whence pair-production of charged strings destabilizes the vacuum [477]. Study of string theories with timelike noncommutativity (e.g. 'noncommutative open string theory' [477, 478]) is outside our scope, but there are at least some hints of similar UV/IR mixing effects as those in the NCFT [479]. We note in passing that there are further interesting connections between NCFTs and string theories — not only do particles on noncommutative spaces act in many ways like rods of size $L \sim p\theta$ (see e.g. [480, 481, 482, 483, 484]), mimicking the behavior of extended objects, but there have been many hints in the spacelike theories that the curious IR effects in the NCFT are reproducing effects from closed strings, despite the fact that these have been decoupled (e.g. [464, 485, 486, 487, 488, 489, 490, 479, 491, 492]).

Within the realm of field theory, there have long been suggestions that this difficulty is pointing to the need for a modified definition of quantum field theories on timelike noncommutative spaces (for some early references, see [493, 494, 495, 496, 497, 498, 499, 500]). From this perspective, the issue is that such field theories are non-local in time, which renders nonsensical the normal time-ordering involved in the perturbative Dyson series (at the least). That is, our effective definition of these theories above via the diagrammatic expansion may be too naïve. An interesting line of work is to formulate a

---

[33]Though it is interesting to note that the special case of 'lightlike' noncommutativity is also unitary [474, 475].

modification of the standard quantum field theory machinery to non-local-in-time theories which avoids the unitarity issue by construction. We note that the same UV/IR mixing effects of interest in the two-point function have been seen to persist in at least some of these approaches (e.g. [495]). For some recent work on the formulation and properties of nonlocal field theories, see e.g. [501, 502, 503, 471, 504, 505].

Below we will begin in Euclidean space, where $k \circ k \geq 0$ is guaranteed for any $\theta_{\mu\nu}$, but will then venture into Lorentzian signature. All of our calculations and the general features we find, including finding new infrared poles, will hold robustly in spacelike non-commutative theories. However we will comment also on how these features are modified when timelike noncommutativity is turned on, taking license from the aforementioned hints that unitary completions/reformulations of timelike NCFT may retain the UV/IR mixing exhibited in the naïve approach.

7.3   Real Scalar ϕ^4 Theory

In this section we review the perturbative physics of the noncommutative real scalar ϕ^4 theory at one loop, which was first studied in detail by Minwalla, Van Raamsdonk, and Seiberg in [464].[34]

In four Euclidean dimensions the action on noncommutative space becomes

$$S = \int \mathrm{d}^4 x \left(\frac{1}{2} \partial_\mu \phi \partial^\mu \phi + \frac{1}{2} m^2 \phi^2 + \frac{g^2}{4!} \phi \star \phi \star \phi \star \phi \right), \tag{7.11}$$

where we have already used the fact that the quadratic part of the noncommutative action is the same as the commutative theory to eliminate the star product there. Our object of interest will be the one-loop correction to the two-point function. In the com-

---

[34]Some early results in this model may also be found in [506, 507].

mutative theory this is given by a single Feynman diagram, but the noncommutative theory contains both a planar diagram and a nonplanar diagram.

The expressions for these two diagrams now differ — not only in symmetry factor but also due to the phase in the integrand. We find

$$
\begin{aligned}
\Gamma_{1,\text{planar}}^{(2)} &= \frac{g^2}{3\,(2\pi)^4} \int \frac{\mathrm{d}^4 k}{k^2 + m^2} \\
\Gamma_{1,\text{nonplanar}}^{(2)} &= \frac{g^2}{6\,(2\pi)^4} \int \frac{\mathrm{d}^4 k}{k^2 + m^2} e^{ik^\mu \theta_{\mu\nu} p^\nu}.
\end{aligned}
\tag{7.12}
$$

We may already see that something interesting should happen, as in the nonplanar diagram the phase mixes the internal and external momenta. One may intuit that the rapidly oscillating phase in the UV of the loop integration will dampen the would-be divergence, and indeed we will see that nonplanar diagrams are finite. However, unlike in the case where the vertex factor vanishes rapidly for large Euclidean momenta and so ensures UV-finiteness [505], here the damping is in some sense 'marginal'. This fact will be responsible for the interesting feature we will find presently.

The simplest method to evaluate noncommutative diagrams is to use Schwinger parameters, recalling the identity $\frac{1}{k^2 + m^2} = \int_0^\infty \mathrm{d}\alpha \; e^{-\alpha\left(k^2 + m^2\right)}$. The presence of the phase in the nonplanar diagram means we must complete the square before going to spherical coordinates to get a Gaussian integral. This means that after the momentum integrals

we end up with

$$
\Gamma^{(2)}_{1,\text{planar}} = \frac{g^2}{48\pi^2} \int \frac{\mathrm{d}\alpha}{\alpha^2} e^{-\alpha m^2}
$$
$$
\Gamma^{(2)}_{1,\text{nonplanar}} = \frac{g^2}{96\pi^2} \int \frac{\mathrm{d}\alpha}{\alpha^2} e^{-\alpha m^2 - \frac{p\circ p}{4\alpha}}
$$

(7.13)

where again $p \circ q = -p^\mu \theta^2_{\mu\nu} q^\nu$. Moving to Schwinger space trades large-k divergences for small-α divergences, which we now smoothly regulate by multiplying the integrands by $\exp\left(-1/(\Lambda^2 \alpha)\right)$ so that the small α region will be driven to zero. Note that a term of this form already exists in the expression for the nonplanar diagram. After introducing the regulator, we can evaluate the integrals to find

$$
\Gamma^{(2)}_{1,\text{planar}} = \frac{g^2}{48\pi^2} \left(\Lambda^2 - m^2 \log\left(\frac{\Lambda^2}{m^2}\right) + \mathcal{O}(1) \right)
$$
$$
\Gamma^{(2)}_{1,\text{nonplanar}} = \frac{g^2}{96\pi^2} \left(\Lambda^2_{\text{eff}} - m^2 \log\left(\frac{\Lambda^2_{\text{eff}}}{m^2}\right) + \mathcal{O}(1) \right),
$$

(7.14)

where we've defined

$$
\Lambda^2_{\text{eff}} \equiv \frac{1}{1/\Lambda^2 + p \circ p/4},
$$

(7.15)

which is the effective cutoff of the nonplanar diagram.

The first thing to note is that it seems the UV divergence of the nonplanar diagram has disappeared — the graph is finite in the limit $\Lambda \to \infty$, and so appears to have been regulated by the noncommutativity of spacetime. In fact the effect is more subtle, as alluded to earlier, and now the UV and IR limits of this amplitude do not commute. If we first take an infrared limit $p \circ p \to 0$ we find that $\Lambda_{\text{eff}} \to \Lambda$ and the ultraviolet divergence of the commutative theory reappears. If we take the UV limit $\Lambda \to \infty$ first we find an IR divergence $\frac{1}{p \circ p}$, so the noncommutativity has transmogrified the UV divergence into an IR one.[35]

---

[35]We note here that the failure of a 'correspondence principle' between commutative and noncom-

Turning to the question of renormalizability, one may naïvely ask if we can absorb all UV divergences into a finite number of counterterms. Under this criterion, it is clear that this procedure works in the noncommutative theory at least when the commutative version is renormalizable. In the current case, we may absorb the UV divergences of this correction to the two-point function into a redefinition of the physical mass, $M^2 = m^2 + \frac{g^2 \Lambda^2}{48\pi^2} - \frac{g^2 m^2}{48\pi^2} \log \frac{\Lambda^2}{m^2}$, and so write down a one-particle irreducible quadratic effective action which has a finite $\Lambda \to \infty$ limit:

$$S_{1\text{PI}}^{(2)} = \int \frac{\mathrm{d}^4 p}{(2\pi)^4} \frac{1}{2} \Bigg(p^2 + M^2 + \frac{g^2}{96\pi^2 \left(\frac{p \circ p}{4} + \frac{1}{\Lambda^2} \right)} \tag{7.16}$$
$$- \frac{g^2 M^2}{96\pi^2} \log \frac{1}{M^2 \left(\frac{p \circ p}{4} + \frac{1}{\Lambda^2} \right)} + \cdots + \mathcal{O}(g^4) \Bigg) \phi(p)\phi(-p).$$

However, in the $\Lambda \to \infty$ limit one finds that at one loop the propagator now has *two* poles. The first is a standard radiative correction to the free pole, but the second has appeared ex nihilo at one loop:

$$
\begin{aligned}
p^2 &= -m^2 + \mathcal{O}(g^2) \\
p \circ p &= -\frac{g^2}{24\pi^2 m^2} + \mathcal{O}(g^4),
\end{aligned}
\tag{7.17}
$$

where we have assumed that $\theta^{\mu\nu}$ is full rank. The former is to be interpreted as the on-shell propagation of the particles associated to our fundamental field ϕ. If $\theta^{\mu\nu}$ has only one eigenvalue $1/\Lambda_\theta^2$ — with Λ_θ thought of as the scale associated with the breakdown of classical geometry — we have $p \circ p = \frac{p^2}{\Lambda_\theta^4}$. We see that the new pole appears at $p^2 \propto g^2 \frac{\Lambda_\theta^4}{m^2}$, and so if our field ϕ lives in the deep UV of the theory, our new pole appears

mutative theories as $\theta^{\mu\nu} \to 0$ is clearly intrinsically linked to the appearance of UV/IR mixing. This failure doesn't violate Kontsevich's proof of the existence of deformation quantization for any symplectic manifold [508], as that is confined solely to 'formal' deformation quantization — that is, the production of a formal power series expansion of the algebra of observables in terms of the deformation parameter. As was noted in Section 7.2 and is now on prime display, the physics of the theory with nonperturbative θ-dependence is starkly different from that of any truncation.

at parametrically low energy scales. To the extent that poles are particles, we appear to have generated a new light particle from ultraviolet dynamics.

The interpretation of the new pole can be sharpened by considering more carefully the criteria for renormalizability in Wilsonian EFT. In a Wilsonian picture, we upgrade our Lagrangian parameters to running parameters, and define our theory at the scale Λ as

$$S_{Wilson}(\Lambda) = \int \mathrm{d}^4 x \left(\frac{1}{2} Z(\Lambda) \partial_\mu \phi \partial^\mu \phi + \frac{1}{2} Z(\Lambda) m^2(\Lambda) \phi^2 + \frac{Z^2(\Lambda) g^2(\Lambda)}{4!} \phi \star \phi \star \phi \star \phi \right).$$

$$(7.18)$$

It is immediately apparent from the above calculation that we cannot write the action at a lower scale $\Lambda_0 < \Lambda$ in this same form by choosing appropriate definitions for $Z(\Lambda), m(\Lambda), g(\Lambda)$ — there's nowhere to put the $\frac{1}{p \circ p}$ term![36]

Stated more precisely, for Wilsonian renormalizability we require that we can define the running couplings such that correlation functions computed from this action converge uniformly to their $\Lambda \to \infty$ limits. However, this requirement is flatly violated by the noncommutation of the UV and IR limits of the diagrams. For any finite value of Λ, the effective action of Equation 7.16 differs significantly from its limiting value for small momenta $p \circ p \ll \frac{1}{\Lambda^2}$. This is the precise sense in which the violation of Wilsonian EFT appears in this one-loop correction.

This brings up the question of how an effective field theorist would describe the

---

[36]There has been much work on understanding renormalizability of NCFTs, especially with an eye toward finding a mathematically well-defined four-dimensional quantum field theory with a non-trivial continuum limit. Renormalizability has been proven for modifications of NCFTs where the free action is supplemented by an additional term which adjusts its long-distance behavior. Such an action is manufactured either by requiring it manifest 'Langman-Szabo' duality [509] $p_\mu \leftrightarrow 2(\theta^{-1})_{\mu\nu} x^\nu$ [510, 511] or by adding a $1/p \circ p$ term to the free Lagrangian [512], the latter of which directly has the interpretation of adding 'somewhere to put the $1/p \circ p$ counterterm'. For recent reviews of these and related efforts we refer the reader to [513, 514]. It would be interesting to understand fully the extent to which the physics of these schemes agrees with the interpretation of the IR effects as coming from auxiliary fields [464, 488].

universe if they unknowingly lived on a noncommutative space. A consistent Wilsonian interpretation can be regained by including a degree of freedom which can absorb the new infrared dynamics of the quadratic effective action. Since we need this to involve the ϕ momentum, this new particle must mix linearly with the ϕ field. We manufacture its tree-level Lagrangian such that the problematic inverse $p \circ p$ term in the quadratic effective action of ϕ is replaced with its $\Lambda \to \infty$ value for all values of Λ, to satisfy our precise condition for Wilsonian renormalizability. To see how this works, we add to our tree level Wilsonian action

$$\Delta S(\Lambda) = \int \mathrm{d}^4 x \left(\frac{1}{2} \partial \chi \circ \partial \chi + \frac{1}{2} \frac{\Lambda^2}{4} (\partial \circ \partial \chi)^2 + i \frac{1}{\sqrt{24\pi^2}} g \chi \phi \right). \qquad (7.19)$$

Since χ appears quadratically, we may integrate it out exactly at tree level to find a contribution to the effective action

$$\Delta S_{\mathrm{1PI}}(\Lambda) = \int \frac{\mathrm{d}^4 p}{(2\pi)^4} \frac{1}{2} \left(-\frac{g^2}{96\pi^2 \left(\frac{p \circ p}{4} + \frac{1}{\Lambda^2} \right)} + \frac{g^2}{24\pi^2 p \circ p} \right) \phi(p)\phi(-p) \qquad (7.20)$$

This precisely subtracts off the problematic term in the original 1PI quadratic effective action and adds back its $\Lambda \to \infty$ limit, as we had wanted. Ignoring the logarithmic term,[37] we are left with an effective action which is manifestly independent of the cutoff Λ, and so satisfies our criterion for Wilsonian renormalizability.[38] We discuss the generalization

---

[37]Discussion of the interpretation of logarithmic singularities as being due to auxiliary fields propagating in extra dimensions may be found in [488].

[38]In Equation 7.19, the four-derivative quadratic action of the auxiliary field can be rewritten as two fields with two-derivative actions, one of which is of negative norm and may be thought of as the 'Lee-Wick partner' of the positive norm state [515], viz.

$$\mathcal{L} = \frac{1}{2} \partial \chi' \circ \partial \chi' - \frac{1}{2} \partial \tilde{\chi} \circ \partial \tilde{\chi} - \frac{1}{2} \frac{4}{\Lambda^2} \tilde{\chi}^2 + i \frac{1}{\sqrt{24\pi^2}} g \left(\chi' - \tilde{\chi} \right) \phi, \qquad \chi' \equiv \chi + \tilde{\chi} \qquad (7.21)$$

One may then wonder if the lightness of the new IR pole may be understood through the regularization performed by the Lee-Wick field, as is done for the Higgs in the 'Lee-Wick standard model' [516]. However, in that theory the Higgs is kept light because every particle comes with a Lee-Wick partner, and so all diagrams contributing to corrections to the Higgs mass are made finite. The presence of the

of this procedure in Appendix C.

Now while we have written down an action which identifies the new observed IR pole with a field and in doing so gives our effective action a Wilsonian interpretation, the extent to which χ can be taken seriously as a fundamental degree of freedom is unclear.[39] The new pole is inaccessible in Euclidean space — so one does not immediately conclude there is a tachyonic instability — and relatedly, when we naïvely analytically continue this result to Lorentzian spacetime this new pole is inaccessible in the s-channel.[40] However, its presence is still enough to break unitarity for this theory [472], and in fact may still be interpreted as being due to the presence of tachyons [476]. As discussed in Section 7.2, it is possible this may be resolved if analytical continuation is adjusted for nonlocal-in-time theories, or it may be that a UV theory cures this apparent violation.

Separately, it is not obvious much has been gained by attributing the new pole to a new, independent field, past acting as a formal tool to regain a notion of renormalizability. Since the only interaction of χ above is linear mixing, its action is not renormalized — any divergences are instead absorbed into the running of ϕ parameters — and so no

---

Higgs' Lee-Wick partner alone is not enough to keep it light. Here, the lightness of χ can be understood diagrammatically as being simply due to the fact that its only interaction is linear mixing with ϕ, and so any correction to its two-point function is absorbed into that of the two-point function of ϕ. A further issue with the Lee-Wick rewriting is that the seeming perturbative unitary of the theory is normally guaranteed by the Lee-Wick partner being heavy and unstable. But as we take the $\Lambda \to \infty$ limit in our Wilsonian action, we see that the Lee-Wick partner becomes massless as well, in accordance with the result that this theory is non-unitary [472].

[39]We note that in matrix models containing dynamical noncommutative geometries it has been argued that emergent infrared singularities should be associated with the dynamics of the geometry (see e.g. [517, 462]). As our field theories are formulated on fixed noncommutative backgrounds, this interpretation is unavailable to us.

[40]Note that this peculiar connection regarding (in)accessibility is due to the Lorentz violation. While the normal pole which is inaccessible in Euclidean signature becomes accessible for timelike momenta in Lorentzian signature, the Wick rotation affects the noncommutative momentum contraction differently. When taking $x_4 \to -ix_0$, one also rotates $\theta_{4\nu} \to -i\theta_{0\nu}$ such that Equation 7.1 continues to hold for the same numerical $\theta_{\mu\nu}$. For the simplest configuration of full-rank noncommutativity with $\theta_{\mu\nu}$ block-off-diagonal and only one eigenvalue $1/\Lambda_\theta^2$, the Euclidean $p \circ p = p^2/\Lambda_\theta^4$ becomes a Lorentzian $p \circ p = (p_0^2 - p_1^2 + p_2^2 + p_3^2)/\Lambda_\theta^4$. So a noncommutative pole which is inaccessible in the Euclidean theory becomes accessible in the Lorentzian theory for spacelike momenta, while a noncommutative pole which can be accessed in the Euclidean theory becomes accessible in the s-channel in Lorentzian signature.

interactions are generated. Furthermore one is obstructed from integrating out the heavy field ϕ to come up with an effective action of χ at low energies by the fact that the kinetic terms of χ are non-standard, which prevents diagonalization of the quadratic terms in the Lagrangian. Thus it seems it is intrinsically linked with the heavy scalar which begat it.

There are further obstructions to asking that this specific mechanism be responsible for the lightness of an observed particle such as the Higgs. Prime among these is the modified dispersion relation of the new field, $p \circ p = \mathcal{O}(g^2)$, which means that the free propagation of this field would be Lorentz violating.[41] We will explore these issues further in the next sections, as in the Yukawa theory of Section 7.4 the new pole will appear with the opposite sign and so will offer the prospect of appearing as an s-channel pole.

We emphasize that a new infrared scale whose lightness is unexplained in the context of Wilsonian effective field theory is an exciting feature that makes further exploration of UV/IR mixing an interesting pursuit. The fact that it here appears as the scale of a pole in a propagator makes the connection to the hierarchy problem captivating, but asking that this toy model — where Lorentz violation is at the fore — literally solve the problem for us would be too much. We proceed without further hindrance in exploring NCFT so as to learn more about the appearance and effects of UV/IR mixing here.

7.3.1   Dimensional Regularization

A good question to ask is whether, or to what extent, these effects are an artifact of our choice of regularization. To demonstrate their physicality, we repeat the calculation of the one-loop correction to the two-point function now in dimensional regularization. We set up our integral in $d = 4 - \epsilon$ dimensions, having defined $g^2 = \tilde{g}^2 \tilde{\mu}^\epsilon$, and we again

---

[41]This dispersion relation means that χ only propagates in noncommutative directions, and so attempts to use hidden extra-dimensional noncommutativity to avoid four-dimensional Lorentz violation constraints seem a phenomenological nonstarter.

go to Schwinger space:

$$\Gamma^{(2)}_{1,\text{planar}} = \frac{\tilde{g}^2 \tilde{\mu}^\epsilon}{3\,(2\pi)^d} \int \mathrm{d}^d k \; \mathrm{d}\alpha \; e^{-\alpha(k^2+m^2)}$$

$$\Gamma^{(2)}_{1,\text{nonplanar}} = \frac{\tilde{g}^2 \tilde{\mu}^\epsilon}{6\,(2\pi)^d} \int \mathrm{d}^d k \; \mathrm{d}\alpha \; e^{-\alpha(k^2+m^2)+ik^\mu \theta_{\mu\nu} p^\nu}. \tag{7.22}$$

After completing the square in the nonplanar integral, the momentum integral and the Schwinger integral may then be performed analytically, with the results:

$$\Gamma^{(2)}_{1,\text{planar}} = \frac{\tilde{g}^2 \tilde{\mu}^\epsilon}{3\,(4\pi)^{d/2}} (m^2)^{\frac{d}{2}-1} \Gamma(1-\frac{d}{2})$$

$$\Gamma^{(2)}_{1,\text{nonplanar}} = \frac{\tilde{g}^2 \tilde{\mu}^\epsilon}{6\,(4\pi)^{d/2}} 2^{\frac{d}{2}} (m^2)^{\frac{1}{2}(\frac{d}{2}-1)} \left(\sqrt{p \circ p}\right)^{1-\frac{d}{2}} K_{\frac{d}{2}-1}\left(m\sqrt{p \circ p}\right). \tag{7.23}$$

If we expand the planar graph in the limit $\epsilon \to 0$, which should be thought of as probing the ultraviolet, we recover

$$\Gamma^{(2)}_{1,\text{planar}} = -\frac{\tilde{g}^2 m^2}{3(4\pi)^2} \left[\frac{2}{\epsilon} + \ln\frac{\mu^2}{m^2}\right], \tag{7.24}$$

where in $\overline{\text{MS}}$ we would subtract off the pole and find the renormalization group evolution of m from the logarithmic term, as usual.

The question of dimensional regularization for the nonplanar diagram is a subtle one [518]. If we first take the $\epsilon \to 0$ limit of Equation 7.23, we see this manifestly has no divergences, and we are simply left with the finite, ϵ^0 term

$$\Gamma^{(2)}_{1,\text{nonplanar}} = \frac{g^2 m^2}{6(4\pi)^2} \left[\frac{4}{m^2 p \circ p} - \ln\frac{4}{m^2 p \circ p} - 1 + 2\gamma\right], \tag{7.25}$$

which we have expanded near $p \circ p \to 0$ to manifest the IR divergence. We have again transmogrified our UV divergence into an IR pole. We now expect to see that the IR

limit does not commute with the above UV limit. To do so, we expand Equation 7.23
around $p \circ p \to 0$ to find

$$\Gamma^{(2)}_{1,\text{nonplanar}} = \frac{\tilde{g}^2 m^2}{6(4\pi)^2} \frac{\pi^{\epsilon/2} \tilde{\mu}^\epsilon}{m^\epsilon} \Gamma\left(-1 + \frac{\epsilon}{2}\right) + \frac{\tilde{g}^2}{24\pi^2} \tilde{\mu}^\epsilon \pi^{\epsilon/2} \Gamma\left(1 - \frac{\epsilon}{2}\right) p \circ p^{-1+\epsilon/2} + \mathcal{O}(p \circ p).$$
$$(7.26)$$

If we were to now blindly take the $\epsilon \to 0$ limit of this expression, we would again get
Equation 7.25, contrary to our expectations. However, we notice that if the dimension of
spacetime over which we had performed the integral was particularly low $\epsilon > 2$, then we
have incorrectly kept the second term in Equation 7.26, as that term would be at least
$\mathcal{O}(p \circ p)$. If we were to work in $d < 2$, expand in $p \circ p \to 0$ and so ignore that term, and
then analytically continue back to $d = 4$, we would instead find the ϵ^{-1} pole

$$\Gamma^{(2)}_{1,\text{nonplanar}} = -\frac{\tilde{g}^2 m^2}{6(4\pi)^2} \left[\frac{2}{\epsilon} + \ln \frac{\mu^2}{m^2}\right],$$
$$(7.27)$$

and now we recover the UV divergence that was present in the commutative theory, so
that once again we find the UV and IR limits don't commute.

The key to understanding clearly this seemingly ambiguous dimensional regularization
procedure is that while $\Gamma^{(2)}_{1,\text{nonplanar}}(p \circ p) \sim \int \mathrm{d}^d q \, \mathrm{d}\alpha \, e^{-\alpha(q^2 + m^2) - \frac{p \circ p}{4\alpha}}$ is convergent in $d > 2$
for $p \circ p > 0$, at $p \circ p = 0$ it is only convergent for $d < 2$. Since it is a property of
dimensional regularization that if an integral converges in δ dimensions, it converges to
the same value in $d < \delta$ dimensions [66], we may thus perform the integral at $d < 2$ for
all $p \circ p$ and correctly find Equation 7.23. It is only when taking the IR limit that we
must remember the integral was performed in $d < 2$ dimensions, and so our expansion
to get Equation 7.27 is unambiguously correct. Thus our conclusion that the UV and IR
limits of the two-point function do not commute here is robust.

It is thus clear that the UV/IR mixing we have observed in this model is not an artifact

of a choice of regularization, and is in fact a physical feature of this noncommutative field theory.

7.4   Yukawa Theory

7.4.1   Motivation: Strong UV/IR Duality

We observed in our first example that the UV divergences of the real ϕ^4 commutative theory are transmogrified into infrared poles in the noncommutative theory.[42]   It is natural to ask whether this "strong UV/IR duality" [519] is a common feature of *all* noncommutative theories.

The answer is no, and the simplest counterexample is provided in the case of a complex scalar field with global $U(1)$ symmetry and self-interaction [519]. In the quantization of the scalar potential we have two quartic terms which are noncommutatively-inequivalent due to the ordering non-invariance, so the general noncommutative potential is

$$V = m^2|\phi|^2 + \frac{\lambda_1}{4}\phi^* \star \phi \star \phi^* \star \phi + \frac{\lambda_2}{4}\phi^* \star \phi^* \star \phi \star \phi, \qquad (7.28)$$

where λ_1 and λ_2 are now different couplings. By doodling some directed graphs, one sees simply that the one-loop correction to the scalar two-point function contains planar graphs with each of the λ_1, λ_2 vertices, but the only nonplanar graph has a λ_2 vertex. There is thus no necessary connection of the ensuing nonplanar IR singularity to the UV divergence in the $\theta \to 0$ limit, as the coefficients are unrelated (and in particular, we are free to turn off the IR singularity at one loop by setting $\lambda_2 = 0$).

Another important counterexample is that of charged scalars, the simplest example

---

[42]While we only presented the calculation of the one-loop correction to the two-point function, [464] goes through corrections to the two- and n-point functions for ϕ^n with $n = 3, 4$ and finds the same features in all cases.

of which is noncommutative scalar QED, which was first constructed in [520]. There is a very rich and interesting structure of gauge theories on noncommutative spaces, a full discussion of which is far beyond the scope of this paper. We refer the reader to [521, 522, 523, 524, 525, 470, 526] for discussions of some features relevant to SM model-building. We here satisfy ourselves with the simplest case, for which we have the noncommutative Lagrangian[43]

$$\mathcal{L} = \frac{1}{4g^2} F_{\mu\nu} \star F^{\mu\nu} + (D_\mu \phi)^* \star (D^\mu \phi) + V(\phi, \phi^*), \tag{7.29}$$

where even though we're quantizing $U(1)$ we have $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu - i[A_\mu \overset{\star}{,} A_\nu]$ due to the noncommutativity, where $[\cdot \overset{\star}{,} \cdot]$ is the commutator in our noncommutative algebra. The vector fields transform as $A_\mu \mapsto U \star A_\mu \star U^\dagger + i\partial_\mu U \star U^\dagger$, where $U(x)$ is an element of the noncommutative $U(1)$ group, which consists of functions $U(x) = \left(e^{i\theta(x)}\right)_\star$, which is the exponential constructed via power series with the star-product.

The potential and the covariant derivative both depend on the representation we choose for the scalar. In contrast to commutative $U(1)$ gauge theory, where we merely assign ϕ a charge, our only choices now are to put ϕ in either the fundamental or the adjoint of the gauge group. Note that an adjoint field smoothly becomes uncharged in the commutative limit. Such a field ϕ transforms as $\phi \mapsto U \star \phi \star U^\dagger$. The covariant derivative is thus $D^\mu \phi = \partial_\mu \phi - ig[A_\mu \overset{\star}{,} \phi]$. The gauge-invariant potential then includes both quartic terms in Equation 7.28, in addition to others such as $\phi^* \star \phi \star \phi \star \phi$, since the adjoint complex scalar is uncharged at the level of the global part of the gauge symmetry. Strong UV/IR duality then should not hold here either.

The situation is even worse if ϕ is in the fundamental, where it transforms as $\phi \mapsto U \star \phi$

---

[43]It is important to note that many fundamental concepts which one normally thinks of as depending upon Lorentz invariance still hold on noncommutative spaces, due to a 'twisted Poincaré symmetry' [527, 528, 529, 530]. This includes the unitary irreducible representations, so it is sensible to speak of a vector field.

and $\phi^* \mapsto \phi^* \star U^{-1}$ with covariant derivative $D^\mu \phi = \partial^\mu \phi - iA_\mu \star \phi$. It is easy to see in this case that the λ_2 interaction term is no longer gauge invariant, and a charged scalar may only self-interact through $V = \lambda_1 \phi^* \star \phi \star \phi^* \star \phi$. Purely from gauge invariance we thus see that a fundamental scalar has no nonplanar self-interaction diagrams in the one-loop correction to its two-point function, and so there is no remnant of strong UV/IR duality to speak of.[44]

The question is then whether there are other examples where this strong UV/IR duality *does* occur, or whether it is perhaps a peculiar feature of real ϕ^n theories on noncommutative spaces. To answer this, we will study in detail another case of especial phenomenological significance: Yukawa theory. Noncommutative Yukawa theory was first studied in [533].[45] Our result on the presence of strong UV/IR mixing differs, for reasons we will explain henceforth.

7.4.2   Setup

For reasons that will soon become clear, we will now work directly in Minkowski space, and begin with a commutative theory of a real scalar φ and a Dirac fermion ψ with Yukawa interaction:

$$\mathcal{L}^{(\mathrm{C})} = -\frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi - \frac{1}{2}m^2\varphi^2 + i\overline{\psi}\slashed{\partial}\psi - \overline{\psi}M\psi + g\varphi\overline{\psi}\psi. \tag{7.30}$$

When constructing a noncommutative version of this theory, the quadratic part of

---

[44]Noncommutative QED also has strange behavior in the gauge sector that runs counter to strong UV/IR duality — the photon self-energy correction gains an infrared singularity from nonplanar one-loop diagrams, even though the commutative quadratic power-counting divergence is forbidden by gauge-invariance. The theory is constructed in detail in [531], while more physical interpretation is given in [532], and the possible relation to geometric dynamics in the context of matrix models is discussed in [517].

[45]Aspects of noncommutative Yukawa theory have also been studied recently in d=3 in [534], and with a modified form of noncommuativity in [535].

the action does not change. However, ordering ambiguities appear for the interaction term, and we in fact find two noncommutatively-inequivalent interaction terms which generically appear:

$$\mathcal{L}_{\text{int}}^{\text{(NC)}} = g_1 \varphi \star \overline{\psi} \star \psi + g_2 \overline{\psi} \star \varphi \star \psi. \tag{7.31}$$

These terms are inequivalent because the star product is only cyclically invariant. In the analysis of [533], only the g_2 interaction was included. As a result, it was concluded that this theory contains no nonplanar diagrams at one loop, and the first appear at two loops as in Figure 7.1. This immediately tells us that the one-loop quadratic divergence of the scalar self-energy will not appear with a one-loop IR singularity, and so rules out the putative strong UV/IR duality of the theory they studied.
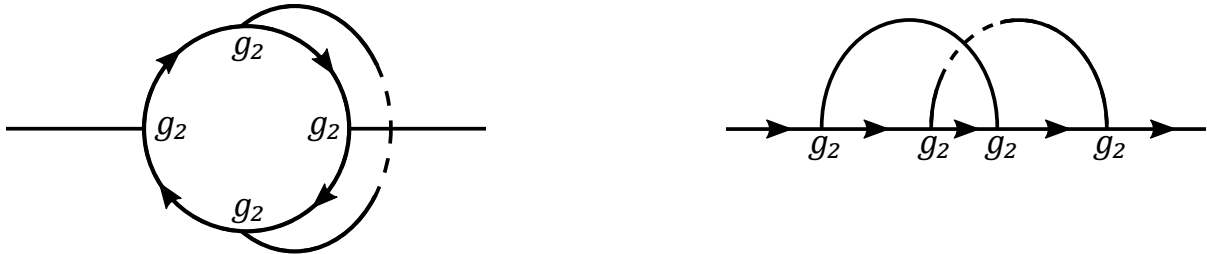
Figure 7.1: Representative leading nonplanar corrections to the self-energies in the non-commutative Yukawa theory of [533]. Fermion lines have arrows and dashing denotes nonintersection.

However, we must ask whether we actually have the freedom to choose g_1 and g_2 independently. To address that question, we must understand the role of discrete symmetries in noncommutative theories. For ease of reference we here repeat our definition of the noncommutativity parameter

$$[x_\mu, x_\nu] = i\theta_{\mu\nu} \tag{7.1}$$

It is manifest that the noncommutativity tensor does not transform homogeneously under either parity or time-reversal, but only under their product: $PT : x_\mu \to -x_\mu \Rightarrow PT : \theta_{\mu\nu} \to \theta_{\mu\nu}$. So while any Lagrangian with full-rank noncommutativity unavoidably

violates both P and T, it may preserve PT.

Since both φ and the scalar fermion bilinear are invariant under all discrete symmetries, these symmetries naïvely play no further role in this theory. However, the time-reversal operator is anti-unitary, and thus negates the phase in the star-product:

$$(PT)^{-1} \left(f(x) \star g(x) \right) PT = g(x) \star f(x). \tag{7.32}$$

Armed with this, we may now apply CPT to our interaction Lagrangian, to find

$$(CPT)^{-1} \mathcal{L}_{\text{int}}^{\text{(NC)}} CPT = g_1 \overline{\psi} \star \varphi \star \psi + g_2 \varphi \star \overline{\psi} \star \psi. \tag{7.33}$$

Comparing with Equation 7.31, we see that our interactions have been re-cycled! Requiring that our interactions preserve CPT amounts to imposing

$$(CPT)^{-1} \mathcal{L}_{\text{int}}^{\text{(NC)}} CPT = \mathcal{L}_{\text{int}}^{\text{(NC)}} \quad \implies \quad g_1 = g_2 \tag{7.34}$$

And so the theory of [533] appears to violate CPT.[46] When we instead include both orderings of interactions the nonplanar diagrams now occur at the first loop order. Furthermore, with both couplings set equal the planar and nonplanar diagrams will have the same coefficients, which reopens the question of strong UV/IR duality for this theory. In the following we will keep g_1 and g_2 distinguished merely to evince how the different vertices appear, but in drawing conclusions about the theory we will set them equal.[47]

---

[46]We note that while the CPT theorem has only been proven in NCFT without space-time noncommutativity [536, 537, 538, 539], the difficulty in the general case is related to the issues with unitarity discussed in Section 7.2, and we expect it should hold in a sensible formulation of the space-time case as well.

[47]We should note that in the construction of noncommutative QED it has been argued that it is sensible to assign θ the anomalous charge conjugation transformation $C : \theta^{\mu\nu} \to -\theta^{\mu\nu}$ ([540] and many others since). The argument is that charged particles in noncommutative space act in some senses like dipoles whose dipole moment is proportional to θ, and so charge conjugation should naturally reverse these dipole moments. Here, however, our particles are uncharged, and thus we have no basis for arguing in

7.4.3   Scalar Two-Point Function

First we consider the planar diagrams, of which there are two:

The 'symmetrization' of the momenta of the internal propagators is an important calculational simplification. This calculation is textbook save for our Schwinger-space regularization, so we will be brief and merely point out the salient features. The sum of these diagrams gives

$$\Gamma_1^{(2),s,p}(p) = i(-1)\left((ig_1)^2 + (ig_2)^2\right) \int \frac{\mathrm{d}^4 k}{(2\pi)^4} \frac{(-i)^2 \mathrm{Tr}\left[\left(M - \slashed{k} - \slashed{p}/2\right)\left(M - \slashed{k} + \slashed{p}/2\right)\right]}{\left((k+p/2)^2 + M^2\right)\left((k-p/2)^2 + M^2\right)}.$$

(7.35)

To evaluate this, we must now introduce two Schwinger parameters α_1, α_2 and then switch to 'lightcone Schwinger coordinates' which effects the change $\int_0^\infty \mathrm{d}\alpha_1 \int_0^\infty \mathrm{d}\alpha_2 \to \int_0^\infty \mathrm{d}\alpha_+ \int_{-\alpha_+}^{+\alpha_+} \mathrm{d}\alpha_-$. Regulating the integral by $\exp\left[-1/\sqrt{2}\alpha_+\Lambda^2\right]$, we may then evaluate and isolate the divergences as $\Lambda \to \infty$ to find

$$\Gamma_1^{(2),s,p}(p) = -\frac{(g_1^2 + g_2^2)}{2\pi^2}\left[\Lambda^2 - \frac{6M^2 + p^2}{4} \log\left(\frac{\Lambda^2}{M^2 + p^2/4}\right) + \dots\right]$$

(7.36)

---

this manner. Furthermore, such an anomalous transformation makes charge conjugation relate theories on *different* noncommutative spaces $\mathcal{M}_\theta \to \mathcal{M}_{-\theta}$. The heuristic picture of the CPT theorem (that is, the reason we care about CPT being a symmetry of our physical theories) is that after Wick rotating to Euclidean space, such a transformation belongs to the connected component of the Euclidean rotation group [541], and so is effectively a symmetry of spacetime. So it is at the least not clear that defining a CPT transformation that takes one to a different space accords with the reason CPT should be satisfied in the first place.

Turning now to the nonplanar diagrams, there are again two

Each now has one g_1 vertex and one g_2 vertex, which makes it clear why the analysis of [533] found no such diagrams. The two diagrams will come with opposite phase factors, $e^{ip\wedge k}$ and $e^{ik\wedge p}$, so we can compute one and then find the other by taking $p \mapsto -p$. In this case it's obvious that after completing the square we will only be left with terms which are quadratic in p, and so the two diagrams give the same contribution. We can thus compute both terms at the same time.

The phase factor in the integrand will modify our change of variables, as it did in the ϕ^4 case, to give again an effective cutoff for this diagram due to the noncommutativity. We find

$$\Gamma_1^{(2),s,np}(p) = \frac{g_1 g_2}{\pi^2} \int \mathrm{d}q \mathrm{d}\alpha_1 \mathrm{d}\alpha_2 q^3 \left(M^2 - q^2 + \frac{\alpha_1 \alpha_2}{(\alpha_1 + \alpha_2)^2} p^2 + \frac{p \circ p}{4(\alpha_1 + \alpha_2)^2} \right)$$
$$\times e^{-(\alpha_1 + \alpha_2)\left(q^2 + M^2\right) - \frac{\alpha_1 \alpha_2}{\alpha_1 + \alpha_2} p^2 - \frac{p \circ p}{4(\alpha_1 + \alpha_2)}}. \quad (7.37)$$

We can now follow the same steps to regulate and integrate this, and again find a closed-form expression for the pieces which contain divergences. Note that unlike the ϕ^4 calculation, we can already see that the nonplanar expression will not merely be given by $\Lambda \to \Lambda_{\mathrm{eff}}$, as the change of variables has here modified the numerator of the integrand to give an extra piece to the momentum polynomial multiplying the exponential. And

so integration gives us

$$\Gamma_1^{(2),s,np}(p) = \frac{g_1 g_2}{1920\pi^2} \left[3 \left(640M^2 + p^4 p \circ p + 40(4M^2 + p^2)p \circ p \Lambda_{\text{eff}}^2 \right) K_0 \left(\frac{\sqrt{4M^2 + p^2}}{\Lambda_{\text{eff}}} \right) \right.$$
$$\left. + 20 \sqrt{4M^2 + p^2} \Lambda_{\text{eff}} \left(-96 + p^2 p \circ p + 12 p \circ p \Lambda_{\text{eff}}^2 \right) K_1 \left(\frac{\sqrt{4M^2 + p^2}}{\Lambda_{\text{eff}}} \right) \right]. \quad (7.38)$$

We must now think slightly more carefully about what we want to add to the quadratic effective action to find a Wilsonian interpretation of this theory. We may isolate the IR divergence that appears when the cutoff is removed by first taking the limit $\Lambda \to \infty$ with $p \circ p$ held fixed, and then expanding around $p \circ p = 0$. We may then ask that this same divergence appears at any value of Λ. To account for this IR divergence, we must add to our effective action

$$\Delta S_{1\text{PI}}(\Lambda) = -\frac{1}{2} \int \frac{\mathrm{d}^4 p}{(2\pi)^4} \frac{g_1 g_2}{2\pi^2} \left(\Lambda_{\text{eff}}^2 - \frac{4}{p \circ p} \right) \varphi(p)\varphi(-p), \quad (7.39)$$

which can easily be done through the addition of an auxiliary scalar field as was done in Section 7.3 and is discussed in more generality in Appendix C. After having added this to our action, for small $p \circ p$ the scalar two-point function now behaves as $\Gamma_1^s(p) = -\frac{2g_1 g_2}{\pi^2 p \circ p} + \ldots$ for any value of Λ. The new pole in this case has the opposite sign as that in 7.19, and so will be accessible in Euclidean signature, clearly signaling a tachyonic instability. While this puts the violation of unitarity in this theory on prime display, it also means that this pole will be accessible in the s-channel in the Lorentzian theory if we allow for timelike noncommutativity.

We emphasize that any conclusions about the Lorentzian theory with timelike non-commutativity are speculative and dependent upon a solid theoretical understanding of a unitary formulation of the field theory, and in principle such a formulation could find radically different IR effects than this naïve approach. However, it was found in [495] that

a modification of time-ordering to explicitly make the theory unitary (at the expense of microcausality violation) leaves the one-loop correction to the self-energy unchanged in ϕ^4 theory, and the same might be expected to hold true for Yukawa theory. This makes it worthwhile to at least briefly consider the potential phenomenological consequences of the new pole.

At low energies, the propagator is here modified to $m^2 + (p_i + p_j)^2 - \frac{2g_1 g_2}{\pi^2} \frac{1}{(p_i+p_j)\circ(p_i+p_j)}$. If we consider scattering of fermions through an s-channel φ and take the simple case of a noncommutativity tensor which in the lab frame has one eigenvalue $1/\Lambda_\theta^2$ with $m^2 \gg \Lambda_\theta^2$, then the emergent pole appears at $s = \frac{2g_1 g_2}{\pi^2} \frac{1-\beta^2}{1+\beta^2} \frac{\Lambda_\theta^4}{m^2}$. Here $s = -(p_i + p_j)^2$ is the invariant momentum routed through the propagator, and β is the boost of the $(p_i + p_j)$ system with respect to the lab frame. The Lorentz-violation here then has the novel effect of smearing out the resonance corresponding to the light pole for a particle which is produced at a variety of boosts. This is in contrast to the pole at m^2, which gives a conventional resonance at leading order. Of course, we have not constructed a fully realistic theory in any respect, and ultimately it may well be that other Lorentz-violating effects provide the leading constraint. Nonetheless, the lineshape of resonances may be an interesting observable in this framework.

A further feature of this opposite sign of the new pole compared to that in the ϕ^4 theory is that the unusual momentum-dependence of the two-point function will lead to ordered phases which break translational invariance [542, 464, 543, 544, 545]. While a Lorentz-violating background field may possibly be very well constrained, the detailed constraint depends on its wavelength and the ways in which it interacts with the SM. But this is another obvious line of exploration for constraining realistic NCFTs.

7.4.4   Fermion Two-Point Function

There are again two planar diagrams:

$$-i\Gamma_1^{(2),f,p} \quad = \quad$$

No new features appear in the evaluation of these diagrams, so we merely quote the final result:

$$\Gamma_1^{(2),f,p} = -\frac{g_1^2 + g_2^2}{16\pi^2}\left(M - \frac{\not{p}}{2}\right)\log\frac{4p^2\Lambda^2}{m^4 + 2m^2(p^2 - M^2) + (M^2 + p^2)^2} + \ldots \qquad (7.40)$$

We also have two nonplanar diagrams, which again mix the two vertices

$$-i\Gamma_1^{(2),f,np} \quad = \quad$$

Here we find that the different phase factors for each diagram, which we saw were inconsequential for the nonplanar corrections to the scalar, have an important role. When we complete the square in each of the two cases, we find that one of the diagrams has an integrand proportional to $\left(M - \not{p}\frac{\alpha_2}{\alpha_1+\alpha_2} - \frac{1}{2}\frac{p_\mu\theta^{\mu\nu}\gamma_\nu}{\alpha_1+\alpha_2}\right)$ and the other is proportional to $\left(M - \not{p}\frac{\alpha_2}{\alpha_1+\alpha_2} + \frac{1}{2}\frac{p_\mu\theta^{\mu\nu}\gamma_\nu}{\alpha_1+\alpha_2}\right)$, so the would-be divergence in $p\theta$ will cancel manifestly between the two diagrams. After this everything proceeds as before, and we find

$$\Gamma_1^{(2),f,np} = -\frac{g_1 g_2}{8\pi^2}\left(M - \frac{\not{p}}{2}\right)\log\frac{4p^2\Lambda_{\text{eff}}^2}{m^4 + 2m^2(p^2 - M^2) + (M^2 + p^2)^2} + \ldots \qquad (7.41)$$

We see that with $g_1 = g_2 \equiv g$, the fermion quadratic effective action also behaves as expected from 'strong UV/IR duality'. The logarithmic divergence of the commutative theory has been transmogrified in the nonplanar diagrams into IR dynamics via the

284

simple replacement $\Lambda \to \Lambda_{\text{eff}}$, and so a $p \circ p \to 0$ pole will emerge when we remove the cutoff. We discuss the use of an auxiliary field to restore a Wilsonian interpretation here in Appendix C.2.

7.4.5 Three-Point Function

The correction to the vertex function constitutes further theoretical data toward the Wilsonian interpretation of the noncommutative corrections. We calculate the one-loop correction in this section and delay the discussion of the use of auxiliary fields to account for them until Appendix C.3. We will find that while we can use the same fields to account for the modifications to both the propagators and the vertices, the physical interpretation of such fields is unclear.

We can compute corrections for each fixed ordering of external lines separately since they're coming from different operators. For simplicity we'll compute the g_1 ordering, which we will denote $\Gamma_3^{\varphi\overline{\psi}\psi}(r, p, \ell)$. There are four diagrams in total: one planar diagram with two insertions of the g_2 vertex, one nonplanar diagram with two insertions of the g_1 vertex, and two nonplanar diagrams with one insertion of each. It is easy to see by looking at the diagrams that the same expressions with $g_1 \leftrightarrow g_2$ compute the correction to the other ordering, $\Gamma_3^{\overline{\psi}\varphi\psi}(r, p, l)$.

The new feature of this computation is that we now need *three* Schwinger parameters, and this presents a problem for our previous computational approach. We won't be able to perform the two finite integrals before expanding in a variable which isolates the divergences when $\alpha_1 + \alpha_2 + \alpha_3 \to 0$, analogously to what we did in $2d$ Schwinger space. Instead we slice $3d$ Schwinger space such that we can perform the integral which isolates the leading divergences first, and then — as long as we're content only to understand this divergence — we can discard the rest without having to worry about performing the

other two integrals.

The planar diagram is

$$i\Gamma_{3,p}^{\varphi\bar{\psi}\psi}(p,\ell) \quad = \quad$$

and corresponds to the expression

$$\Gamma_{3,p}^{\varphi\bar{\psi}\psi}(p,\ell) = -i(ig_1)(ig_2)^2 \times \tag{7.42}$$

$$\int \frac{\mathrm{d}^4k}{(2\pi)^4} \frac{(-i)^3 \left(M - (\not{k} + \frac{\not{p}}{2} + \frac{\not{\ell}}{2})\right)\left(M - (\not{k} - \frac{\not{p}}{2} - \frac{\not{\ell}}{2})\right)}{\left((k + \frac{p}{2} + \frac{\ell}{2})^2 + M^2\right)\left((k - \frac{p}{2} - \frac{\ell}{2})^2 + M^2\right)\left((k + \frac{p}{2} - \frac{\ell}{2})^2 + m^2\right)}.$$

After moving to Schwinger space, integrating over the loop momentum, and introducing a cutoff $\exp\left(-1/\left(\Lambda^2(\alpha_1 + \alpha_2 + \alpha_3)\right)\right)$, we switch variables to

$$\alpha_1 = \xi_1\eta, \qquad \alpha_2 = \xi_2\eta, \qquad \alpha_3 = (1 - \xi_1 - \xi_2)\eta, \tag{7.43}$$

under which $\int_0^\infty \mathrm{d}\alpha_1 \int_0^\infty \mathrm{d}\alpha_2 \int_0^\infty \mathrm{d}\alpha_3 \to \int_0^1 \mathrm{d}\xi_1 \int_0^{1-\xi_1} \mathrm{d}\xi_2 \int_0^\infty \mathrm{d}\eta \; \eta^2$. Performing the momentum integral transfers the divergence for large k to a divergence in small $\alpha_1 + \alpha_2 + \alpha_3 = \eta$. This will allow us to find the leading divergent behavior immediately by carrying out the η integral and then expanding in $\Lambda \to \infty$. This yields

$$\Gamma_{3,p}^{\varphi\bar{\psi}\psi}(p,\ell) = \frac{g_1 g_2^2}{16\pi^2} \log\left(\Lambda^2\right) + \text{ finite}, \tag{7.44}$$

where we are unable to determine the IR cutoff of the logarithm, but this suffices for our purposes.

The three nonplanar graphs now each receive a different phase corresponding to which

external line crosses the internal line

$$
i\Gamma_{3,np}^{\varphi\bar{\psi}\psi}(p,\ell) \quad = \quad \text{} \quad + \quad \text{}
$$

$$
+ \quad \text{} \quad , \tag{7.45}
$$

where the first gets $\exp\left[-i(k\wedge p + k\wedge\ell + p\wedge\ell)\right]$, the second $\exp\left[-i(k\wedge p + p\wedge\ell/2)\right]$, and the third $\exp\left[-i(k\wedge\ell + p\wedge\ell/2)\right]$. The evaluation of these diagrams proceeds as in the previous examples. If we take the IR limit $p,\ell\to 0$ of the nonplanar contributions to this ordering of the three-point function and then expand in large Λ we find

$$
\lim_{p,\ell\to 0}\Gamma_{3,np}^{\varphi\bar{\psi}\psi}(p,\ell) = \frac{g_1^2(g_1+2g_2)}{16\pi^2}\log\left(\Lambda^2\right) + \text{ finite.} \tag{7.46}
$$

However, if we first take the UV limit $\Lambda\to 0$, and then expand in small momenta, we find

$$
\lim_{\Lambda\to\infty}\Gamma_{3,np}^{\varphi\bar{\psi}\psi}(p,\ell) = \frac{g_1^2}{16\pi^2}\left[g_1\log\left(\frac{4}{(p+\ell)\circ(p+\ell)}\right)\right. \tag{7.47}
$$

$$
\left. + g_2\log\left(\frac{4}{p\circ p}\right) + g_2\log\left(\frac{4}{\ell\circ\ell}\right)\right] + \text{ finite,}
$$

where we again see UV/IR mixing, and we note that each nonplanar diagram has been effectively cutoff by the momenta which cross the internal line. We discuss the use of auxiliary fields to restore a Wilsonian interpretation to this vertex correction in Appendix C.3.

7.5   Softly-broken Wess-Zumino Model

We now turn our attention to the softly-broken noncommutative Wess-Zumino model as a controllable example of the interplay between UV/IR mixing and the finiteness of the field theory. We will restrict ourselves to calculating the one-loop correction to the scalar two-point function. Since the new poles appearing in the quadratic effective action in the scalar and Yukawa theories are intimately related to the quadratic divergences of the commutative theories, we will not be surprised to find that this feature will disappear when both the scalar and the fermion are present in the EFT below the cutoff. By studying the softly-broken theory we can take the fermion above or below the cutoff to smoothly see the relation between the finiteness of the field theory and the effects of UV/IR mixing. The exactly supersymmetric noncommutative Wess-Zumino model was first discussed in detail in [546], and the absence of an infrared pole in a softly-broken theory was first noted in [532]. The softly-broken Wess-Zumino model was first considered in [469].[48]

The noncommutative Wess-Zumino theory can be suitably formulated in off-shell superspace as

$$\mathcal{L} = \int \mathrm{d}^4\theta \; Z\Phi^\dagger\Phi + \int \mathrm{d}^2\theta \; \left(\frac{1}{2}M\Phi^2 + \frac{1}{6}y \; \Phi \star \Phi \star \Phi\right) + \; \text{h.c.}, \qquad (7.48)$$

where Φ is a chiral superfield and we have included a wavefunction renormalization factor in the Kähler potential $Z = 1 + \mathcal{O}(y^2)$. We can introduce soft supersymmetry breaking by promoting this factor to a spurion $Z = 1 + (|M|^2 - m^2)\theta^2\theta^{\dagger 2}$, the only effect of which is to modify the scalar mass spectrum.

---

[48] Our one-loop results agree with those of [469] save for their claim that logarithmic IR divergences are absent in the exactly supersymmetric theory, which contradicts [546]. We will below find a logarithmic IR divergence in the wavefunction renormalization which is independent of the soft-breaking, which is consistent with the expectations of strong UV/IR duality.

Formulating the noncommutative theory including the auxiliary F fields makes it manifest that we have preserved supersymmetry off-shell. This procedure is in fact precisely the same as quantizing after integrating out F, and so we end up with a star-product version of the familiar Lagrangian:

$$
\begin{aligned}
-\mathcal{L}_{\mathrm{NCWZ}} = {} & Z\partial^\mu\phi^*\partial_\mu\phi - iZ\psi^\dagger\bar{\sigma}^\mu\partial_\mu\psi \\
& + Z^{-1}m^2\phi^*\phi + \frac{1}{2}M\psi\psi + \frac{1}{2}M^*\psi^\dagger\psi^\dagger \\
& + \frac{1}{2}Z^{-1}y\phi\star\psi\star\psi + \frac{1}{2}Z^{-1}y^*\phi^*\star\psi^\dagger\star\psi^\dagger \\
& + \frac{1}{2}Z^{-1}yM^*\phi\star\phi\star\phi^* + \frac{1}{2}Z^{-1}y^*M\phi^*\star\phi^*\star\phi \\
& + \frac{1}{4}Z^{-1}\left|y\right|^2\phi\star\phi\star\phi^*\star\phi^*
\end{aligned}
\tag{7.49}
$$

where ϕ is a complex scalar and ψ is a Weyl fermion. Of course, now that we've introduced supersymmetry breaking we expect to find that there is further renormalization beyond that associated with Z, but keeping the manifest factors of Z will allow us to easily compare to our expectations for the supersymmetric limit.

The calculation of the one-loop correction to the two-point function goes much as the previously-demonstrated examples. The presence of the three-scalar interaction gives a new class of diagrams, whose evaluation is routine. The two-component fermions yield slightly different factors than did the Dirac fermions [113]. Finally, it is important to note that the results for the diagrams computed in Section 7.3 cannot be used here, as we must here regulate uniformly using $\exp(-1/(\Lambda^2(\alpha_1 + \alpha_2)))$ like we did in Section 7.4. This may be easily accommodated by writing the integrand in the quartic diagrams as $\frac{1}{k^2+m^2}\frac{k^2+m^2}{k^2+m^2}$.

Adding up all these diagrams and taking the limit where $\Lambda, \Lambda_{\mathrm{eff}}$ are large, we find

that the one-loop scalar two-point function may be organized as

$$\Gamma^{(2),s} \equiv Z p^2 + Z^{-1}(m^2 + \delta m^2) \tag{7.50}$$

$$Z = 1 + \frac{y^2}{32\pi^2} \log\left[\frac{\Lambda\Lambda_{\text{eff}}}{M^2}\right] + \dots \tag{7.51}$$

$$\delta m^2 = \frac{y^2}{32\pi^2}\left(M^2 - m^2\right)\log\left[\frac{\Lambda\Lambda_{\text{eff}}}{M^2}\right] + \dots, \tag{7.52}$$

where we make manifest the presence of supersymmetric nonrenormalization in the limit $m \to M$, which acts as a non-trivial check. As expected, the absence of the quadratic UV divergence in the Wess-Zumino model has led to the absence of an infrared pole from the noncommutativity, even as the fermion is made arbitrarily heavy relative to the scalar. However, logarithmic UV/IR mixing still occurs.

We may repeat this calculation using dimensional regularization and taking note of the issues which arose in Section 7.3.1. Using the same parametrization of the one-loop two-point function as above, the planar diagrams contribute

$$Z_{\text{planar}} = 1 + \frac{y^2}{64\pi^2}\left(\frac{2}{\epsilon} + \log\frac{\mu^2}{M^2}\right) + \dots \tag{7.53}$$

$$\delta m^2_{\text{planar}} = \frac{y^2}{64\pi^2}(M^2 - m^2)\left(\frac{2}{\epsilon} + \log\frac{\mu^2}{M^2}\right) + \dots, \tag{7.54}$$

as expected. The full form of the nonplanar diagrams is unenlightening, but if we take the IR limit $p \circ p \to 0$ first, they give precisely the same contribution as the planar diagrams, since the diagram degeneracies are all the same in this case. Taking the UV limit $\epsilon \to 0$ first (and staying in $d < 2$), we instead find

$$Z_{\text{nonplanar}} = 1 + \frac{y^2}{64\pi^2} \log\frac{4}{M^2 p \circ p} + \dots \tag{7.55}$$

$$\delta m^2_{\text{nonplanar}} = \frac{y^2}{64\pi^2}(M^2 - m^2)\log\frac{4}{M^2 p \circ p} + \dots, \tag{7.56}$$

which has precisely the same correspondence with the Schwinger-space regularization as we saw for the ϕ^4 case.

We thus see clearly the conflict between supersymmetry and the use of UV/IR mixing to explain low-energy puzzles. UV/IR mixing transmogrified UV momentum dependence into IR momentum dependence, and so depended crucially on the sensitivity of our field theory to UV modes. For a theory which is finite as a field theory, the dependence on the UV physics has been removed, and so we see no interesting IR effects.

Of course, in the presence of a cutoff Λ it is also possible to study the behavior of the scalar two-point function when $M^2 \gg \Lambda^2 \gg |M^2 - m^2|$ as the fermion is taken above the cutoff while keeping the scalar light. This corresponds to taking $M/\Lambda, M/\Lambda_{\text{eff}} > 1$ and then expanding in the limit where $\Lambda, \Lambda_{\text{eff}}$ are large. This gets rid of the nonplanar Yukawa-type diagrams and, as one might expect, results in a return of UV sensitivity in the scalar EFT below the cutoff, foreshadowing a return of the UV/IR mixing effects. The scalar mass-squared in this limit becomes

$$\delta m^2 = \frac{y^2}{256\pi^2} \left(6M^2 + 16\Lambda^2 + 8\Lambda_{\text{eff}}^2 \right) + \dots. \tag{7.57}$$

and UV/IR mixing reappears at the quadratic level. So our EFT intuition isn't totally out the window; it's been broken in a controlled way, and we can smoothly interpolate between theories with and without UV/IR mixing by taking the states responsible for finiteness above the cutoff. This sharpens the sense in which UV/IR mixing can do something interesting in the IR as long as the field-theoretic description of our universe is never finite.

Ultimately, this highlights a central challenge for approaching the hierarchy problem via UV/IR mixing. The hierarchy problem is particularly sharp when the full theory is finite and scale separation is large, in which case the sensitivity of the Higgs mass to

underlying scales is unambiguous. But UV/IR mixing effects potentially relevant to the hierarchy problem are absent in this case, and emerge only when finiteness is lost. This tension is not necessarily fatal to UV/IR approaches to the hierarchy problem – ultimately the UV sensitive degrees of freedom are not the ones we would wish to identify with the Higgs – but it bears emphasizing.

Moreover, there is a possible loophole in the general argument that finiteness must be surrendered in order to generate a scale from UV/IR mixing. The presence of interesting effects in the IR here depends solely on the UV sensitivity of the nonplanar diagrams. The 'orbifold correspondence' [286, 285, 284] provides non-supersymmetric field theories constructed via orbifold truncation of $\mathcal{N} > 0$ theories whose planar diagrams agree with those of the supersymmetric theory and so are finite. A noncommutative orbifold field theory [547] may then provide a theory which is fully predictive, yet which still generates an infrared scale via UV/IR mixing. Generally, it may be possible that UV/IR mixing appears in such a way that it is the sole effect sensitive to short distances.

7.6   Whence UV/IR Mixing?

To attempt to formulate a realistic theory which uses UV/IR mixing to solve extant theoretical puzzles, it would be useful to have an understanding of which features of NCFT were responsible for the curious infrared effects discussed above. This would be helpful whether one wishes to test out these ideas in any of the many proposed modifications of NCFT, or to write down other toy models which share some features of NCFT but are based upon different principles.

Qualitatively, the two unusual features involved in the formulation of NCFT are Lorentz invariance violation and nonlocality. However, it is obvious that one may have theories with one or both of these features without the interesting effects we have seen.

The answer then is not so simple as pointing to one axiom or another of EFT which has been broken, but depends sensitively on the way in which they are broken. We briefly explore two ways we may better understand the interplay here between nonlocality and Lorentz-violation and how they come together to cause surprising low-energy effects. We first give a general argument based on the way nonlocality appears to postdict the form of the violation of EFT expectations. We then phenomenologically examine the loop integration appearing in our NCFT calculations to diagnose what caused the appearance of the IR pole. This will lead us to discuss an avenue toward investigating (or manufacturing) such effects in nonlocal, Lorentz-invariant theories.

To see how EFT expectations may be violated, consider the peculiar way in which the noncommutative effects in the one-loop action (e.g. Equation 7.16) induce nonlocality. In Wilsonian EFT, integrating out momentum modes $p \gtrsim \Lambda$ produces a nonlocal theory at those scales, or equivalently on distances $x \lesssim 1/\Lambda$. However, particles on a noncommutative space can be thought of as rods of size $L \sim p\theta$ [480, 481, 482, 483, 484]. This tells us that in a NCFT we should expect nonlocality to be present for scales $x \lesssim p\theta$. Comparing the two scales, we see that we should find nonlocal effects past those expected in Wilsonian EFT for $\frac{1}{\Lambda} < p\theta$. Here this momentum-dependent nonlocality occurs in a Lorentz-violating way. This expectation was exactly borne out in the examples above, where we saw that the one-loop effective action in momentum space is nonlocal for $p \circ p \gg 1/\Lambda^2$ [464].

Purely from this analysis of the form of nonlocality, we may conclude there will be a breakdown of Wilsonian renormalization. After we remove the cutoff, the theory should be nonlocal on all scales $p \circ p > 0$. But if we compute a correlation function at a large-but-finite Λ, the theory will still be local for momenta $p \circ p < 1/\Lambda^2$, and so will greatly differ from the continuum result. So our surprising discovery of the non-uniform convergence of correlation functions in the examples above is understood easily from this picture.

While this sort of momentum-dependent nonlocality may seem *ad hoc*, it has been suggested previously for separate purposes. It has been argued [548] that quantum gravity should obey a 'Generalized Uncertainty Principle' $\Delta x \gtrsim \frac{\hbar}{\Delta p} + \ell_p^2 \Delta p$, with ℓ_p the Planck length, based on the use of Hawking radiation to measure the horizon area of a black hole. This gives precisely the same sort of momentum-dependent nonlocality as we saw above. We refer the reader to [549] for a review of the Generalized Uncertainty Principle, [550, 551] for similar conclusions within string theory, and [552] for a more general review of the appearance of an effective minimal length in quantum gravity. It would be interesting to investigate other field theories which obey such uncertainty principles and determine whether UV/IR mixing causes similar features as appear in NCFT. For theories which violate Lorentz invariance, care must be taken to avoid arguments that even Planck-scale Lorentz violation is empirically ruled out [465, 553].

We may also attempt to phenomenologically diagnose what caused the appearance of the IR pole from the form of the loop integration. The presence of an exponential of momenta was clearly crucial, and this implies a necessity of nonlocality. It's also clear that the modification of the cutoff in the nonplanar diagrams $\Lambda \mapsto \Lambda_{\text{eff}}$, which rendered the diagrams UV finite in a way that brought UV/IR mixing, was a result of the contraction between the loop momentum and the external momentum. Less obviously, one may see that any quadratic term in loop momentum in the exponential would have erased this feature, as after momentum integration one would find an integrand $\sim \frac{1}{1+\alpha_+}$, and any divergence will have disappeared. Heuristically, the quadratic suppression in loop momentum is too strong and regulates the UV divergence entirely independently of the cutoff, so no UV/IR mixing appears. NCFT disallows such terms as a result of momentum contractions being performed with an antisymmetric tensor, and this particular mechanism seems to imply the necessity of Lorentz invariance violation. However, this argument only considers small deviations from the form of the integral in NCFT. Further

discussions of the form of loop integrals with generalizations of the star-product may be found in [554, 555].

Likely a better approach to understand the prospect for finding features similar to that of NCFT in a Lorentz invariant theory is to back up and study formulations of Lorentz invariant extensions of NCFT. This is accomplished by upgrading the noncommutativity tensor $\theta^{\mu\nu}$ from a c-number to an operator. This was proposed already by Snyder in 1947 [311], and this approach has been revived a number of times more recently (e.g. [556, 557, 558, 559, 560]). Schematically, this results in an action containing an integral over $\theta^{\mu\nu}$

$$ S = \int \mathrm{d}^4 x \, \mathrm{d}^6 \theta \, W(\theta) \, \mathcal{L}(\phi, \partial\phi), \tag{7.58} $$

where $W(\theta)$ is a 'weighting function', and the Lagrangian is still defined using the star-product. The challenge in this approach for our purposes is in devising a method for nonperturbative calculations in θ, which as we saw above was necessary to preserve the features of UV/IR mixing.

Searching more generally for Lorentz invariant theories which contain UV/IR mixing will likely allow more promising phenomenological applications. That such theories should exist can be broadly motivated by quantum gravity, as any gravitational theory is expected both to be nonlocal and to have UV/IR mixing. That Lorentz violation should be present is less clear. A particularly interesting line of development is to then understand in detail the class of nonlocal theories that would have UV/IR mixing of a sort similar to that discussed here. Recent work toward placing nonlocal quantum field theories on solid theoretical ground [471, 505] is clearly of sharp interest here, though the larger goal is quite distinct. The nonlocality studied in these works is designed to render the field theory UV-finite, and so the nonlocal vertex kernels are chosen precisely to avoid the introduction of new poles by ensuring these are momentum-space entire

functions which vanish rapidly in Euclidean directions. The nonlocal vertices of NCFT manage to introduce new poles by oscillating as $p \to \infty$, which presumably allows for the appearance of new 'endpoint singularities' [561, 562], though a full examination of the Landau equations in NCFT has not (to our knowledge) been performed. Our interest is thus in a disjoint class of nonlocal theories, where new poles can appear in interesting ways. Classifying the space of such theories and developing an approach to systematically understand their unitarity properties seems well motivated.

7.7 Conclusions

The lack of evidence for conventional solutions to the hierarchy problem has placed particle physics at a crossroads. While it is possible that the answer ultimately lies further down the well-trodden path of existing paradigms, the appeal of less-travelled paths grows greater with every inverse femtobarn of LHC data.

In this work we have ventured to take seriously the apparent failure of expectations from Wilsonian effective field theory regarding the hierarchy problem by investigating a concrete framework — noncommutative field theory — in which Wilsonian EFT itself breaks down. Not only does noncommutative field theory violate Wilsonian expectations, it provides a sharp instance of UV/IR mixing: ultraviolet modes of noncommutative theories can generate an infrared scale whose origin is opaque to effective field theory. To the extent that UV/IR mixing has any relevance to the hierarchy problem, the emergence of an infrared scale seems to be among the most promising effects. Although the real-world applicability of these theories is likely limited by their Lorentz violation, they nonetheless provide valuable toy models for exploring the potential relevance of UV/IR mixing to problems of the Standard Model.

To this end, we have surveyed existing results on noncommutative theories with an

eye towards 'strong UV/IR duality' — the transmogrification of UV divergences into infrared poles at the same order. This led us to a detailed analysis of noncommutative Yukawa theory, perhaps the most useful toy model for thinking about the hierarchy problem (insofar as the Yukawa sector of the Standard Model is responsible for the largest UV sensitivity of the Higgs mass, and highlights the relative UV *insensitivity* of the fermion masses). In the noncommutative theory, the presence of both inequivalent Yukawa couplings implies the same strong UV/IR duality exhibited by real ϕ^4 theory: a quadratic divergence in the one-loop correction to the scalar mass from fermion loops gives rise to a simple IR pole, while a logarithmic UV divergence in the one-loop correction to the fermion mass from scalar loops give rise to only a logarithmic IR divergence. Intriguingly, the infrared pole in the scalar two-point function appears accessible in the s-channel in the Lorentzian theory, a feature which gives it particular phenomenological relevance.

We then introduced softly-broken supersymmetry as a way to explore the interplay between (in)finiteness and UV/IR mixing. Choosing soft terms in order to keep the scalar light as the fermion mass is varied concretely illustrates several expected features. Strong UV/IR duality is preserved in the sense that both UV and IR divergences are absent at quadratic order (and persist at logarithmic order) when both the scalar and the fermion are in the spectrum. However, infrared structure reappears as the fermion mass is raised above a fixed cutoff and (quadratic) finiteness is lost. This underlines the sense in which UV/IR mixing may only ever play an interesting role when the field theory is quadratically UV sensitive at all scales, a scenario in which the hierarchy problem is less concrete.

Finally, building on the lessons from the toy models considered here, we have highlighted a variety of interesting lines of exploration in theories featuring nonlocality with or without Lorentz violation that may be of relevance to the hierarchy problem.

While the prospect that UV/IR mixing will solve outstanding theoretical problems in the low-energy universe is possibly fanciful, now is the time for such reveries. The paradigms of the past few decades of particle theory are under considerable empirical pressure, and innovative approaches are needed. At the very least, by pushing the limits of EFT we stand to learn more about the broad spectrum of phenomena possible within quantum field theory.

Chapter 8

Conclusion

We end the way we began: Declaring it to be an exciting time in particle physics. The picture we have painted above on the state of the field is one of uncertainty— and indeed we have barely even touched on many of the important problems of the Standard Model. Dark matter and neutrino masses, while having had canonical, obvious, beautiful solutions in the context of supersymmetric grand unified theories, have likewise seen a resurgence of interest in alternative mechanisms in the past few years as a result of the lack of observational evidence for these standard solutions. But these facts all make the universe a more exciting place to study. Imagine if we had found weak-scale supersymmetry at the LHC, and our job now was simply to interpret the data in terms of which of the supersymmetric extensions proposed and well-studied in the past decades was correct. Or even worse, if technicolor had really been the answer and we had to watch Nature repeat the same trick she used at the strong scale again at the weak scale.

How dreadfully boring!

Yes, yes, this attitude is selfish and a bit flippant, but what we now have is the chance to learn *more* about the universe, about the spectrum of possibilities in physics, and about new, radical ideas. I look forward to what we will discover.

Appendix A

Kinetic Mixing in the Mirror Twin Higgs

Since kinetic mixing plays a central role in freeze-twin dark matter, we discuss here at some length the order at which it is expected in the low-energy EFT. Of course, there may always be UV contributions which set ϵ to the value needed for freeze-in. However, if the UV completion of the MTH disallows such terms - for example, via supersymmetry, an absence of fields charged under both sectors, and eventually grand unification in each sector (see e.g. [564, 565, 270, 271, 566, 567])- then the natural expectation is for mixing of order these irreducible IR contributions.

To be concrete, we imagine that $\epsilon = 0$ at the UV cutoff of the MTH, $\Lambda \lesssim 4\pi f$. To find the kinetic mixing in the regime of relevance, at momenta $\mu \lesssim 1$ GeV, we must run down to this scale. As we do not have the technology to easily calculate high-loop-order diagrams, our analysis is limited to whether we can prove diagrams at some loop order are vanishing or finite, and so do not generate mixing. Thus our conclusions are strictly always 'we know no argument that kinetic mixing of this order is not generated', and there is always the possibility that further hidden cancellations appear. With that caveat

divulged, we proceed and consider diagrammatic arguments in both the unbroken and broken phases of electroweak symmetry.

Starting in the unbroken phase, we compute the mixing between the hypercharge gauge bosons. Two- and three-loop diagrams with Higgs loops containing one gauge vertex and one quartic insertion vanish. By charge conjugation in scalar QED, the three-leg amplitude of a gauge boson and a complex scalar pair must be antisymmetric under exchange of the scalars. However, the quartic coupling of the external legs ensures that their momenta enter symmetrically. As this holds off-shell, the presence of a loop which looks like
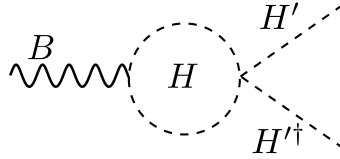

causes the diagram to vanish. However, at four loops the following diagram can be drawn which avoids this issue:
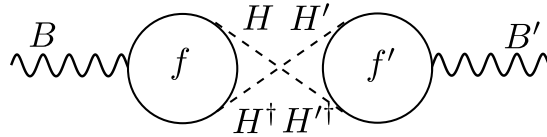
where the two hypercharges are connected by charged fermion loops in their respective sectors and the Higgs doublets' quartic interaction. This diagram contributes at least from the MTH cutoff $\Lambda \lesssim 4\pi f$ down to f, the scale at which twin and electroweak symmetries are broken. We have no argument that this vanishes nor that its unitarity cuts vanish. We thus expect a contribution to kinetic mixing of $\epsilon \sim g_1^2 c_W^2/(4\pi)^8$, with g_1 the twin and SM hypercharge coupling and $c_W = \cos\theta_W$ appearing as the contribution to the photon mixing operator. In this estimate we have omitted any logarithmic dependence on mass scales, as it is subleading.

In the broken phase, we find it easiest to perform this analysis in unitary gauge. The Higgs radial modes now mass-mix, but the emergent charge conjugation symmetries in the two QED sectors allow us to argue vanishing to higher-loop order. The implications of the formal statement of charge conjugation symmetry are subtle because we have two QED sectors, so whether charge conjugation violation is required in both sectors seems unclear. However, similarly to the above case, there is a symmetry argument which holds off-shell. The result we rely on here is that in a vector-like gauge theory, diagrams with any fermion loops with an odd number of gauge bosons cancel pairwise. Thus, each fermion loop must be sensitive to the chiral nature of the theory, so the first non-vanishing contribution is at five loops as in:
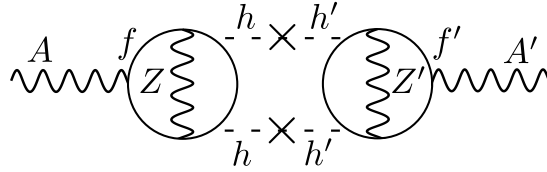
where the crosses indicate mass-mixing insertions between the two Higgs radial modes which each contribute $\sim v/f$. Thus, both the running down to low energies and the finite contributions are five-loop suppressed. From such diagrams, one expects a contribution $\epsilon \sim e^2 g_A^2 g_V^2 (v/f)^2/(4\pi)^{10}$, where with g_V and g_A we denote the vector and axial-vector couplings of the Z, respectively. We note there are other five loop diagrams in which Higgses couple to massive vectors which are of similar size or smaller.

Depending on the relative sizes of these contributions, one then naturally expects kinetic mixing of order $\epsilon \sim 10^{-13} - 10^{-10}$. If ϵ is indeed generated at these loop-levels, then mixing on the smaller end of this range likely requires that it becomes disallowed not far above the scale f. However, we note that our ability to argue for higher-loop order vanishing in the broken versus unbroken phase is suggestive of the possibility that there may be further cancellations. We note also the possibility that these diagrams,

even if nonzero, generate only higher-dimensional operators. Further investigation of the generation of kinetic mixing through a scalar portal is certainly warranted.

Appendix B

How to Formulate Field Theory on a Noncommutative Space

In this appendix we provide detail on how to formulate field theories on a space which is defined by Equation 7.1, which we repeat here for convenience:

$$[\hat{x}^\mu, \hat{x}^\nu] = i\theta^{\mu\nu}. \tag{B.1}$$

To construct a field theory on this space we must specify the algebra of observables. First we briefly recall the familiar, commutative case. For simplicity, we consider a scalar field theory on flat Euclidean space. We denote by $\mathrm{Alg}\left(\mathbb{R}^d[x], \cdot\right)$ the commutative, C^*-algebra of Schwartz functions of d-dimensional Euclidean space with the standard point-wise product, and this constitutes our algebra of observables. A convenient basis for the vector space is that of plane waves $e^{ip\cdot x}$.

The case of interest here is noncommutative flat Euclidean space, on which we define $\mathrm{Alg}\left(\mathbb{R}^d_\theta[\hat{x}], \cdot\right)$. This now consists of such functions of d variables $\hat{x}^\mu$ related by Equation 7.1, and so is a noncommutative algebra, although we've specified again the normal

'point-wise' product. A useful basis will again be that of plane waves, which we may define as the eigenfunctions of appropriately-defined derivatives on the noncommutative space, and which look familiar $e^{ip \cdot \hat{x}}$. To can get a sense for this algebra it is useful to carry out the simple exercise of multiplying two plane waves by simply applying Baker-Campbell-Hausdorff

$$e^{ik \cdot \hat{x}} \cdot e^{ik' \cdot \hat{x}} = \exp\left(ik \cdot \hat{x} + ik' \cdot \hat{x} - \frac{1}{2} k_\mu k'_\nu \left[\hat{x}^\mu, \hat{x}^\nu \right] \right) = e^{-\frac{i}{2} \theta^{\mu\nu} k_\mu k'_\nu} e^{i(k+k') \cdot \hat{x}}. \qquad \text{(B.2)}$$

As in quantum mechanics, we will wish to study noncommutative versions of familiar commutative theories, and so it will be useful to view $\mathbb{R}^d_\theta$ as a 'deformation' of $\mathbb{R}^d$. We then wish to construct a map from our commutative algebra to our noncommutative one which returns smoothly to the identity as $\theta^{\mu\nu} \to 0$. The standard such choice is the Weyl-Wigner map $\hat{\mathcal{W}}$, which one may roughly think of as merely replacing xs with $\hat{x}$s.

The procedure is simply to Fourier transform from commutative space to momenta, and then inverse Fourier transform to noncommutative space. Given a commutative space Schwartz function f, we may compose the two operations and write

$$\hat{\mathcal{W}}[f] = \int \mathrm{d}^d x\, f(x) \hat{\Delta}(x), \qquad \hat{\Delta}(x) = \int \frac{\mathrm{d}^d k}{(2\pi)^d} e^{ik_i \cdot \hat{x}^i} e^{-ik_i \cdot x^i}. \qquad \text{(B.3)}$$

Note that this is an injective map of Schwartz functions on $\mathbb{R}^d$ to those on $\mathbb{R}^d_\theta$ which respects the vector space structure but not the structure of the algebra. This property is familiar from quantum mechanics.

We may now construct noncommutative versions of field variables, but we still don't know how to do physics on these spaces. That is, we can write down the Lagrangian for noncommutative ϕ^4 theory, and we could even determine an action after we formulate a notion of an integral over a noncommutative space. But our familiar results about how

to go from the action of a field theory to a calculation for a physical observable most certainly depended implicitly on living on a commutative space, and so it seems we must re-formulate physics from the bottom up.

Fortunately, such a drastic measure may not be necessary, as one may formulate QFT on noncommutative spaces as a simple modification of our normal field theory structure. The core idea is to find an algebra of functions on $\mathbb{R}^d$ which is isomorphic to $\mathrm{Alg}\left(\mathbb{R}^d_\theta[\hat{x}],\cdot\right)$ by pushing the noncommutativity into a new field product, known as a Groenewold-Moyal product (or star-product). We diagram the structure we wish to look for in Figure B.1.

$$\mathrm{Alg}\left(\mathbb{R}^d[x],\cdot\right)$$

$$\hat{\mathcal{W}}{:}\mathbb{R}^d[x]{\to}\mathbb{R}^d_\theta[\hat{x}] \qquad\qquad \mathrm{Id}_{\mathbb{R}^d[x]}$$

$$\mathrm{Alg}\left(\mathbb{R}^d_\theta[\hat{x}],\cdot\right) \xleftarrow{\hspace{1cm}} \hat{\mathcal{W}} \text{ an isomorphism of algebras} \xrightarrow{\hspace{1cm}} \mathrm{Alg}\left(\mathbb{R}^d[x],\star_\theta\right)$$
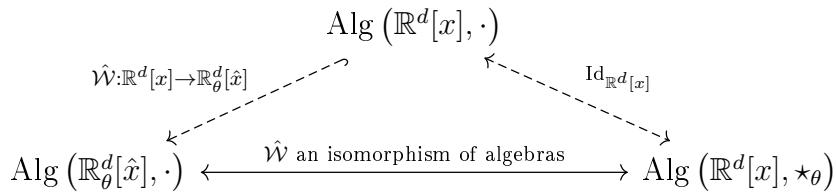
Figure B.1: The relations between the algebra of a commutative field theory, the noncommutative algebra one finds by applying the Weyl-Wigner map, and the noncommutative algebra most useful for field theory making use of the Groenewold-Moyal product.The vertical arrows respect only the vector space structure, and one should think of constructing a new algebra by applying a vector space map and then endowing the vectors with a multiplication operation.

In particular, we may do this by demanding that our quantization map $\hat{\mathcal{W}}$ is upgraded to an isomorphism between $\mathrm{Alg}\left(\mathbb{R}^d_\theta[\hat{x}],\cdot\right)$ and an algebra on the vector space of functions of commutative Euclidean space, with a multiplication operation which is chosen to preserve the algebraic structure. That is, we must satisfy

$$\hat{\mathcal{W}}\left[f \star_\theta g\right] = \hat{\mathcal{W}}\left[f\right] \cdot \hat{\mathcal{W}}\left[g\right], \tag{B.4}$$

for any Schwartz functions f, g on commutative Euclidean space. But we can guarantee this by ensuring it for plane waves, the calculation of which we've essentially already

done above in Equation B.2:

$$
\begin{aligned}
e^{ikx} \star_\theta e^{ik'x} &= \hat{\mathcal{W}}^{-1} \left[\hat{\mathcal{W}} \left[e^{ikx} \right] \cdot \hat{\mathcal{W}} \left[e^{ik'x} \right] \right] \\
&= \hat{\mathcal{W}}^{-1} \left[e^{-\frac{i}{2} \theta^{ij} k_i k'_j} e^{i(k+k')\cdot \hat{x}} \right] \\
e^{ikx} \star_\theta e^{ik'x} &\equiv e^{-\frac{i}{2} \theta^{\mu\nu} k_\mu k'_\nu} e^{i(k+k')\cdot x},
\end{aligned}
\tag{B.5}
$$

where the θ subscript merely tells us the star-product will depend on the noncommutativity tensor, and this will henceforth be dropped. This gives a position-space representation of the star-product,

$$
f(x) \star g(x) = \exp \left(\frac{i}{2} \theta_{\mu\nu} \partial_y^\mu \partial_z^\nu \right) f(y) g(z) \Big|_{y=z=x} = f(x) \exp \left(\frac{i}{2} \overleftarrow{\partial}^\mu \theta_{\mu\nu} \overrightarrow{\partial}^\nu \right) g(x). \tag{B.6}
$$

The general procedure to construct a noncommutative field theory from a commutative one is then by application of the Weyl-Wigner map. As an example, for a simple ϕ^n theory we find

$$
\mathcal{L}_{\text{int}}^{(NC)} = \frac{\lambda}{n!} \hat{\mathcal{W}}^{-1} \left[\hat{\mathcal{W}} \left(\phi(x) \right)^n \right] = \frac{\lambda}{n!} \overbrace{\phi(x) \star \phi(x) \star \cdots \star \phi(x)}^{\text{n copies}}. \tag{B.7}
$$

Appendix C

Wilsonian Interpretations of NCFTs from Auxiliary Fields

In this appendix we discuss various generalizations of the procedure introduced in [464, 488] to account for the new structures appearing in the noncommutative quantum effective action via the introduction of additional auxiliary fields.

C.1 Scalar Two-Point Function

It is simple to generalize the procedure discussed in Section 7.3 to add to the quadratic effective action of ϕ any function we wish through judicious choice of the two-point function for an auxiliary field σ which linearly mixes with it. In position space, if we wish to add to our effective Lagrangian

$$\Delta\mathcal{L}_{\text{eff}} = \frac{1}{2}c^2\phi(x)f(-i\partial)\phi(x), \tag{C.1}$$

where $f(-i\partial)$ is any function of momenta, and c is a coupling we've taken out for convenience, then we simply add to our tree-level Lagrangian

$$\Delta\mathcal{L} = \frac{1}{2}\sigma(x)f^{-1}(-i\partial)\sigma(x) + ic\sigma(x)\phi(x), \tag{C.2}$$

where f^{-1} is the operator inverse of f. It should be obvious that this procedure is entirely general. As applied to the Euclidean ϕ^4 model, we may use this procedure to add a second auxiliary field to account for the logarithmic term in the quadratic effective action as

$$\Delta\mathcal{L} = \frac{1}{2}\sigma(x)\frac{1}{\log\left[1 - \frac{4}{\Lambda^2\partial\circ\partial}\right]}\sigma(x) - \frac{gM}{\sqrt{96\pi^2}}\sigma(x)\phi(x), \tag{C.3}$$

where we point out that the argument of the log is just $4/(\Lambda_{\text{eff}}^2 p\circ p)$ in position space. We may then try to interpret σ also as a new particle. As discussed in [488], its logarithmic propagator may be interpreted as propagation in an additional dimension of spacetime.

Alternatively, we may simply add a single auxiliary field which accounts for both the quadratic and logarithmic IR singularities by formally applying the above procedure. But having assigned them an exotic propagator, it then becomes all the more difficult to interpret such particles as quanta of elementary fields.

C.2 Fermion Two-Point Function

To account for the IR structure in the fermion two-point function, we must add an auxiliary fermion ξ. If we wish to find a contribution to our effective Lagrangian of

$$\Delta\mathcal{L}_{\text{eff}} = c^2\bar{\psi}\mathcal{O}\psi, \tag{C.4}$$

where $\mathcal{O}$ is any operator on Dirac fields, then we should add to our tree-level Lagrangian

$$\Delta\mathcal{L} = -\bar{\xi}\mathcal{O}^{-1}\xi + c\left(\bar{\xi}\psi + \bar{\psi}\xi\right), \tag{C.5}$$

with $\mathcal{O}^{-1}$ the operator inverse of $\mathcal{O}$. In the Lorentzian Yukawa theory of Section 7.4, if we add to the Lagrangian

$$\Delta\mathcal{L} = -\bar{\xi}\frac{M - i\slashed{\partial}/2}{M^2 - \partial^2/4}\left[\log\left(1 - \frac{4}{\Lambda^2\partial\circ\partial}\right)\right]^{-1}\xi + \frac{g}{2\sqrt{2}\pi}\left(\bar{\xi}\psi + \overline{\psi}\xi\right). \tag{C.6}$$

we again find a one-loop quadratic effective Lagrangian which is equal to the $\Lambda \to \infty$ value of the original, but now for any value of Λ.

C.3   Three-Point Function

We may further generalize the procedure for introducing auxiliary fields to account for IR poles to the case of poles in the three-point effective action. It's clear from the form of the IR divergences in Equation ?? that they 'belong' to each leg, and so naïvely one might think this means that the divergences we've already found in the two point functions already fix them. However those corrections only appear in the internal lines and were already proportional to g^2, and so they will be higher order corrections. Instead we must generate a correction to the vertex function itself which only corrects one of the legs.

To do this we must introduce auxiliary fields connecting each possible partition of the interaction operator. However, while an auxiliary scalar χ coupled as $\chi\varphi + \chi\overline{\psi}\psi$ would generate a contribution to the vertex which includes the χ propagator with the φ momentum flowing through it, it would also generate a new $(\overline{\psi}\psi)^2$ contact operator, which we don't want. To avoid this we introduce *two* auxiliary fields with off-diagonal

311

two-point functions, a trick used for similar purposes in [488]. By abandoning minimality, we can essentially use an auxiliary sector to surgically introduce insertions of functions of momenta wherever we want them.

We can first see how this works on the scalar leg. We add to our tree-level Lagrangian

$$\Delta \mathcal{L} = -\chi_1(x) f^{-1}(-i\partial)\chi_2(x) + \kappa_1 \chi_1(x)\varphi(x) + \kappa_2 \chi_2(x) \star \overline{\psi}(x) \star \psi(x). \qquad (C.7)$$

Now to integrate out the auxiliary fields we note that for a three point vertex, one may use momentum conservation to put all the noncommutativity between two of the fields. That is, $\chi_2(x) \star \overline{\psi}(x) \star \psi(x) = \chi_2(x)(\overline{\psi}(x) \star \psi(x)) = (\overline{\psi}(x) \star \psi(x))\chi_2(x)$ as long as this is not being multiplied by any other functions of x. So we may use this form of the interaction to simply integrate out the auxiliary fields. We end up with

$$\Delta \mathcal{L}_{\text{eff}} = \kappa_1 \kappa_2 \overline{\psi} \star \psi \star f(-i\partial)\varphi \qquad (C.8)$$

which is exactly of the right form to account for an IR divergence in the three-point function which only depends on the φ momentum.

For the fermionic legs, we need to add fermionic auxiliary fields which split the Yukawa operator in the other possible ways. We introduce Dirac fields ξ, ξ' and a differential operator on such fields $\mathcal{O}^{-1}(-i\partial)$. Then if we add to the Lagrangian

$$\Delta \mathcal{L} = -\overline{\xi}\mathcal{O}^{-1}\xi' - \overline{\xi'}\mathcal{O}^{-1}\xi + c_1(\overline{\xi}\star\psi\star\varphi + \overline{\psi}\star\xi\star\varphi) + c_2(\overline{\xi}\star\varphi\star\psi + \overline{\psi}\star\varphi\star\xi) + c_3(\overline{\xi'}\psi + \overline{\psi}\xi'), \ (C.9)$$

we now end up with a contribution to the effective Lagrangian

$$\Delta \mathcal{L}_{\text{eff}} = c_1 c_3 \left(\bar{\psi} \star \mathcal{O}(\psi) \star \varphi + \bar{\psi} \star \mathcal{O}(\psi \star \varphi) \right) + c_2 c_3 \left(\bar{\psi} \star \varphi \star \mathcal{O}(\psi) + \bar{\psi} \star \mathcal{O}(\varphi \star \psi) \right),$$
$$(C.10)$$

where we have abused notation and now the argument of $\mathcal{O}$ specifies which fields it acts on. These terms have the right form to correct both vertex orderings.

Now that we've introduced interactions between auxiliary fields and our original fields, the obvious question to ask is whether we can utilize the *same* auxiliary fields to correct both the two-point and three-point actions. In fact, using two auxiliary fields with off-diagonal propagators per particle we may insert any corrections we wish. The new trick is to endow the auxiliary field interactions with extra momentum dependence.

For a first example with a scalar, consider differential operators f, Φ, and add to the Lagrangian

$$\Delta\mathcal{L} = -\chi_1 f^{-1}(-i\partial)\chi_2 + \kappa_1\chi_1\varphi + \kappa_2\chi_2\overline{\psi}\star\psi + g\varphi\Phi(-i\partial)\chi_2. \tag{C.11}$$

We may now integrate out the auxiliary fields and find

$$\Delta\mathcal{L}_{\text{eff}} = g\kappa_1\varphi f(\Phi(\varphi)) + \kappa_1\kappa_2\overline{\psi}\star\psi\star f(\varphi) \tag{C.12}$$

where we've assumed that f and Φ commute. If we take $\Phi = \mathbb{1}$ then we have the interpretation of merely inserting the χ two-point function in both the two-and three-point functions. But we are also free to use some nontrivial Φ, and thus to make the corrections to the two- and three-point functions have whatever momentum dependence we wish. It should be obvious how to generalize this to insert momentum dependence into the scalar lines of arbitrary n−point functions.

The case of a fermion is no more challenging in principle. For differential operators

313

$\mathcal{O}, \mathcal{F}$, we add

$$\Delta\mathcal{L} = -\overline{\xi}\mathcal{O}^{-1}\xi' - \overline{\xi'}\mathcal{O}^{-1}\xi + c_1(\overline{\xi} \star \psi \star \varphi + \overline{\psi} \star \xi \star \varphi) + c_2(\overline{\xi} \star \varphi \star \psi + \overline{\psi} \star \varphi \star \xi)$$
$$+ c_3(\overline{\xi'}\psi + \overline{\psi}\xi') + \frac{g}{2}\left(\bar{\xi}\mathcal{O}^{-1}\mathcal{F}\psi + \bar{\psi}\mathcal{O}^{-1}\mathcal{F}\xi\right), \quad \text{(C.13)}$$

and upon integrating out the auxiliary fields we find

$$\Delta\mathcal{L}_{\text{eff}} = gc_3\bar{\psi}\mathcal{F}\psi + c_1 c_3\left(\bar{\psi} \star \mathcal{O}\left(\psi\right) \star \varphi + \bar{\psi} \star \mathcal{O}\left(\psi \star \varphi\right)\right) + c_2 c_3\left(\bar{\psi} \star \varphi \star \mathcal{O}\left(\psi\right) + \bar{\psi} \star \mathcal{O}\left(\varphi \star \psi\right)\right),$$
$$\text{(C.14)}$$

where the generalization to n-points is again clear. Note that in the fermionic case it's crucial that we be allowed to insert different momentum dependence in the corrections to the two- and three-point functions, as these have different Lorentz structures.

Now we cannot quite implement this for the two- and three-point functions calculated in Section 7.4, for the simple reason that we regulated these quantities differently. That is, we have abused notation and the symbol 'Λ' means different things in the results for the two- and three-point functions. In order to carry out this procedure, we could simply regulate the two-point functions in $3d$ Schwinger space, though we run into the technical obstruction that the integration method above only calculates the leading divergence, which is not good enough for the scalar case.

Bibliography

[1] L. Carroll and J. Tenniel, *Through the Looking-Glass, and What Alice Found There*. Macmillan, 1871.

[2] T. Cohen, N. Craig, S. Koren, M. McCullough and J. Tooby-Smith, *Supersoft Stops*, 2002.12630.

[3] N. Craig and S. Koren, *IR Dynamics from UV Divergences: UV/IR Mixing, NCFT, and the Hierarchy Problem*, *JHEP* **03** (2020) 037 [1909.01365].

[4] S. Koren and R. McGehee, *Freezing-in twin dark matter*, *Phys. Rev. D* **101** (2020) 055024 [1908.03559].

[5] N. Craig, I. Garcia Garcia and S. Koren, *The Weak Scale from Weak Gravity*, *JHEP* **09** (2019) 081 [1904.08426].

[6] S. B. Giddings, S. Koren and G. Treviño, *Exploring strong-field deviations from general relativity via gravitational waves*, *Phys. Rev. D* **100** (2019) 044005 [1904.04258].

[7] S. Koren, *Neutrino – Dark Matter Scattering and Coincident Detections of UHE Neutrinos with EM Sources*, *JCAP* **09** (2019) 013 [1903.05096].

[8] A. Herderschee, S. Koren and T. Trott, *Constructing $\mathcal{N} = 4$ Coulomb branch superamplitudes*, *JHEP* **08** (2019) 107 [1902.07205].

[9] A. Herderschee, S. Koren and T. Trott, *Massive On-Shell Supersymmetric Scattering Amplitudes*, *JHEP* **10** (2019) 092 [1902.07204].

[10] S. Alipour-Fard, N. Craig, S. Gori, S. Koren and D. Redigolo, *The second Higgs at the lifetime frontier*, *JHEP* **07** (2020) 029 [1812.09315].

[11] N. Craig, I. Garcia Garcia and S. Koren, *Discrete Gauge Symmetries and the Weak Gravity Conjecture*, *JHEP* **05** (2019) 140 [1812.08181].

[12] S. Alipour-Fard, N. Craig, M. Jiang and S. Koren, *Long Live the Higgs Factory: Higgs Decays to Long-Lived Particles at Future Lepton Colliders*, *Chin. Phys. C* **43** (2019) 053101 [1812.05588].

[13] N. Craig, S. Koren and T. Trott, *Cosmological Signals of a Mirror Twin Higgs*, *JHEP* **05** (2017) 038 [1611.07977].

[14] S. C. Koren, C. H. Blake, C. C. Dahn and H. C. Harris, *The Low-mass Astrometric Binary LSR 1610-0040*, *The Astronomical Journal* **151** (2016) 57 [1511.02234].

[15] S. C. Koren, E. L. Wright and A. Mainzer, *Characterizing asteroids multiply-observed at infrared wavelengths*, *Icarus* **258** (2015) 82 [1506.04751].

[16] T. Cohen, *As Scales Become Separated: Lectures on Effective Field Theory*, *PoS* **TASI2018** (2019) 011 [1903.03622].

[17] H. Georgi, *Effective field theory*, *Ann. Rev. Nucl. Part. Sci.* **43** (1993) 209.

[18] J. P. Society, *The Jewish Bible Tanakhh: The New Jps Translation According to the Traditional Hebrew Text*. Jewish Publication Society, 2007.

[19] P. W. Anderson, *Plasmons, Gauge Invariance, and Mass*, *Phys. Rev.* **130** (1963) 439.

[20] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, *Phys. Rev. Lett.* **13** (1964) 321.

[21] G. Guralnik, C. Hagen and T. Kibble, *Global Conservation Laws and Massless Particles*, *Phys. Rev. Lett.* **13** (1964) 585.

[22] P. W. Higgs, *Broken symmetries, massless particles and gauge fields*, *Phys. Lett.* **12** (1964) 132.

[23] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, *Phys. Rev. Lett.* **13** (1964) 508.

[24] G. 't Hooft, *Renormalizable Lagrangians for Massive Yang-Mills Fields*, *Nucl. Phys. B* **35** (1971) 167.

[25] M. Srednicki, *Quantum field theory*. Cambridge University Press, 2007.

[26] M. D. Schwartz, *Quantum Field Theory and the Standard Model*. Cambridge University Press, 3, 2014.

[27] P. Langacker, *The standard model and beyond*. 2, 2010.

[28] C. Burgess and G. Moore, *The standard model: A primer*. Cambridge University Press, 12, 2006.

[29] B. Grzadkowski, M. Iskrzynski, M. Misiak and J. Rosiek, *Dimension-Six Terms in the Standard Model Lagrangian*, *JHEP* **10** (2010) 085 [1008.4884].

[30] B. Gripaios and D. Sutherland, *DEFT: A program for operators in EFT*, *JHEP* **01** (2019) 128 [1807.07546].

[31] J. de Blas, J. Criado, M. Perez-Victoria and J. Santiago, *Effective description of general extensions of the Standard Model: the complete tree-level dictionary*, *JHEP* **03** (2018) 109 [1711.10391].

[32] I. Brivio, Y. Jiang and M. Trott, *The SMEFTsim package, theory and tools*, *JHEP* **12** (2017) 070 [1709.06492].

[33] A. Falkowski, M. González-Alonso and K. Mimouni, *Compilation of low-energy constraints on 4-fermion operators in the SMEFT*, *JHEP* **08** (2017) 123 [1706.03783].

[34] J. Ellis, C. W. Murphy, V. Sanz and T. You, *Updated Global SMEFT Fit to Higgs, Diboson and Electroweak Data*, *JHEP* **06** (2018) 146 [1803.03252].

[35] J. de Blas, M. Ciuchini, E. Franco, S. Mishima, M. Pierini, L. Reina et al., *The Global Electroweak and Higgs Fits in the LHC era*, *PoS* **EPS-HEP2017** (2017) 467 [1710.05402].

[36] I. Brivio and M. Trott, *The Standard Model as an Effective Field Theory*, *Phys. Rept.* **793** (2019) 1 [1706.08945].

[37] E. Fermi, *An attempt of a theory of beta radiation. 1.*, *Z. Phys.* **88** (1934) 161.

[38] J. C. Pati and A. Salam, *Lepton Number as the Fourth Color*, *Phys. Rev. D* **10** (1974) 275.

[39] H. Georgi and S. Glashow, *Unity of All Elementary Particle Forces*, *Phys. Rev. Lett.* **32** (1974) 438.

[40] A. Einstein, *Considerations concerning the fundaments of theoretical physics*, *Science* **91** (1940) 487 [https://science.sciencemag.org/content/91/2369/487.full.pdf].

[41] C. Kounnas, A. Masiero, D. V. Nanopoulos and K. A. Olive, *Grand Unification With and Without Supersymmetry and Cosmological Implications*. 1985.

[42] G. G. Ross, *Grand Unified Theories*. 1985.

[43] S. Raby, *Supersymmetric Grand Unified Theories: From Quarks to Strings via SUSY GUTs*, vol. 939. Springer, 2017, 10.1007/978-3-319-55255-2.

[44] M. Sher, *Grand unification, Higgs bosons, and baryogenesis*, in *Strings, Branes and Extra Dimensions: TASI 2001: Proceedings*, 2001.

[45] PARTICLE DATA GROUP collaboration, *Review of Particle Physics, Phys. Rev.* **D98** (2018) 030001.

[46] A. B. Goncharov, *A simple construction of Grassmannian polylogarithms, arXiv e-prints* (2009) arXiv:0908.2238 [`0908.2238`].

[47] A. B. Goncharov, M. Spradlin, C. Vergu and A. Volovich, *Classical Polylogarithms for Amplitudes and Wilson Loops, Phys. Rev. Lett.* **105** (2010) 151605 [`1006.5703`].

[48] C. Duhr, H. Gangl and J. R. Rhodes, *From polygons and symbols to polylogarithmic functions, JHEP* **10** (2012) 075 [`1110.0458`].

[49] D. Gaiotto, J. Maldacena, A. Sever and P. Vieira, *Pulling the straps of polygons, JHEP* **12** (2011) 011 [`1102.0062`].

[50] C. Duhr, *Mathematical aspects of scattering amplitudes*, in *Theoretical Advanced Study Institute in Elementary Particle Physics: Journeys Through the Precision Frontier: Amplitudes for Colliders*, pp. 419–476, 2015, DOI [`1411.7538`].

[51] N. Arkani-Hamed, J. L. Bourjaily, F. Cachazo, A. B. Goncharov, A. Postnikov and J. Trnka, *Grassmannian Geometry of Scattering Amplitudes*. Cambridge University Press, 4, 2016, 10.1017/CBO9781316091548, [`1212.5605`].

[52] A. Zee, *Quantum field theory in a nutshell*. 11, 2003.

[53] S. Mandelstam, *Determination of the pion - nucleon scattering amplitude from dispersion relations and unitarity. General theory, Phys. Rev.* **112** (1958) 1344.

[54] K. G. Wilson, *The Renormalization Group: Critical Phenomena and the Kondo Problem, Rev. Mod. Phys.* **47** (1975) 773.

[55] M. E. Peskin and D. V. Schroeder, *An Introduction to quantum field theory*. Addison-Wesley, Reading, USA, 1995.

[56] T. Appelquist and J. Carazzone, *Infrared Singularities and Massive Fields, Phys. Rev. D* **11** (1975) 2856.

[57] B. Delamotte, *A Hint of renormalization, Am. J. Phys.* **72** (2004) 170 [`hep-th/0212049`].

[58] A. Neumaier, *Renormalization without infinities – an elementary tutorial*, Jun, 2015.

[59] S. Weinberg, *The Quantum theory of fields. Vol. 1: Foundations*. Cambridge University Press, 6, 2005.

[60] T. Schäfer and E. V. Shuryak, *Instantons in QCD, Rev. Mod. Phys.* **70** (1998) 323 [`hep-ph/9610451`].

[61] S. Vandoren and P. van Nieuwenhuizen, *Lectures on instantons*, `0802.1862`.

[62] G. V. Dunne and M. Unsal, *Resurgence and Trans-series in Quantum Field Theory: The CP(N-1) Model, JHEP* **11** (2012) 170 [`1210.2423`].

[63] G. V. Dunne and M. Unsal, *Uniform WKB, Multi-instantons, and Resurgent Trans-Series, Phys. Rev. D* **89** (2014) 105009 [`1401.5202`].

[64] D. Dorigoni, *An Introduction to Resurgence, Trans-Series and Alien Calculus, Annals Phys.* **409** (2019) 167914 [`1411.3585`].

[65] M. Mariño, *Instantons and Large N: An Introduction to Non-Perturbative Methods in Quantum Field Theory*. Cambridge University Press, 9, 2015.

[66] J. C. Collins, *Renormalization*, vol. 26 of *Cambridge Monographs on Mathematical Physics*. Cambridge University Press, Cambridge, 1986.

[67] S. R. Coleman and J. Mandula, *All Possible Symmetries of the S Matrix, Phys. Rev.* **159** (1967) 1251.

[68] J. Polchinski, *Scale and Conformal Invariance in Quantum Field Theory, Nucl. Phys. B* **303** (1988) 226.

[69] Y. Nakayama, *Scale invariance vs conformal invariance, Phys. Rept.* **569** (2015) 1 [`1302.0884`].

[70] D. J. Gross and F. Wilczek, *Ultraviolet Behavior of Nonabelian Gauge Theories, Phys. Rev. Lett.* **30** (1973) 1343.

[71] H. Politzer, *Reliable Perturbative Results for Strong Interactions?, Phys. Rev. Lett.* **30** (1973) 1346.

[72] *Manuscripts 78-81*, in *Collected Papers of L.D. Landau*, D. T. HAAR], ed., pp. 607 – 610, Pergamon, (1965), DOI.

[73] J. Polchinski, *Renormalization and Effective Lagrangians, Nucl. Phys. B* **231** (1984) 269.

[74] K. von Meyenn, *Über das relativitätsprinzip und die aus demselben gezogenen folgerungen*, in *Albert Einsteins Relativitätstheorie: Die grundlegenden Arbeiten*, K. von Meyenn, ed., (Wiesbaden), pp. 160–214, Vieweg+Teubner Verlag, (1990), DOI.

[75] S. Weinberg, *Photons and Gravitons in S-Matrix Theory: Derivation of Charge Conservation and Equality of Gravitational and Inertial Mass, Phys. Rev.* **135** (1964) B1049.

[76] P. A. Dirac, *New basis for cosmology, Proc. Roy. Soc. Lond. A* **A165** (1938) 199.

[77] G. 't Hooft, *Naturalness, chiral symmetry, and spontaneous chiral symmetry breaking, NATO Sci. Ser. B* **59** (1980) 135.

[78] K. Zhou, *Cosmological relaxation*, 2019.

[79] G. D'Ambrosio, G. Giudice, G. Isidori and A. Strumia, *Minimal flavor violation: An Effective field theory approach, Nucl. Phys. B* **645** (2002) 155 [hep-ph/0207036].

[80] Y. Grossman and P. Tanedo, *Just a Taste: Lectures on Flavor Physics*, in *Theoretical Advanced Study Institute in Elementary Particle Physics: Anticipating the Next Discoveries in Particle Physics*, pp. 109–295, 2018, DOI [1711.03624].

[81] J. Zupan, *Introduction to flavour physics, CERN Yellow Rep. School Proc.* **6** (2019) 181 [1903.05062].

[82] S. Gori, *TASI lectures on flavor physics, PoS* **TASI2018** (2019) 013.

[83] N. Arkani-Hamed, *The Future of Fundamental Physics, Daedalus* **141** (2012) 53.

[84] S. Samuel, *The Standard model in its other phase, Nucl. Phys. B* **597** (2001) 70 [hep-ph/9910559].

[85] C. Quigg and R. Shrock, *Gedanken Worlds without Higgs: QCD-Induced Electroweak Symmetry Breaking, Phys. Rev. D* **79** (2009) 096002 [0901.3958].

[86] B. R. Holstein and J. F. Donoghue, *Classical physics and quantum loops, Phys. Rev. Lett.* **93** (2004) 201602 [hep-th/0405239].

[87] G. W. Anderson and D. J. Castano, *Measures of fine tuning, Phys. Lett. B* **347** (1995) 300 [hep-ph/9409419].

[88] S. Fichet, *Quantified naturalness from Bayesian statistics, Phys. Rev. D* **86** (2012) 125029 [1204.4940].

[89] R. Barbieri and G. F. Giudice, *Upper Bounds on Supersymmetric Particle Masses, Nucl. Phys.* **B306** (1988) 63.

[90] P. Athron and D. Miller, *A New Measure of Fine Tuning, Phys. Rev. D* **76** (2007) 075010 [0705.2241].

[91] C. Cheung and G. N. Remmen, *Naturalness and the Weak Gravity Conjecture*, *Phys. Rev. Lett.* **113** (2014) 051601 [`1402.2287`].

[92] H. Ooguri and C. Vafa, *Non-supersymmetric AdS and the Swampland*, *Adv. Theor. Math. Phys.* **21** (2017) 1787 [`1610.01533`].

[93] L. E. Ibanez, V. Martin-Lozano and I. Valenzuela, *Constraining Neutrino Masses, the Cosmological Constant and BSM Physics from the Weak Gravity Conjecture*, *JHEP* **11** (2017) 066 [`1706.05392`].

[94] L. E. Ibanez, V. Martin-Lozano and I. Valenzuela, *Constraining the EW Hierarchy from the Weak Gravity Conjecture*, `1707.05811`.

[95] Y. Hamada and G. Shiu, *Weak Gravity Conjecture, Multiple Point Principle and the Standard Model Landscape*, *JHEP* **11** (2017) 043 [`1707.06326`].

[96] D. Lust and E. Palti, *Scalar Fields, Hierarchical UV/IR Mixing and The Weak Gravity Conjecture*, *JHEP* **02** (2018) 040 [`1709.01790`].

[97] E. Gonzalo, A. Herráez and L. E. Ibáñez, *AdS-phobia, the WGC, the Standard Model and Supersymmetry*, *JHEP* **06** (2018) 051 [`1803.08455`].

[98] E. Gonzalo and L. E. Ibáñez, *The Fundamental Need for a SM Higgs and the Weak Gravity Conjecture*, *Phys. Lett.* **B786** (2018) 272 [`1806.09647`].

[99] J. March-Russell and R. Petrossian-Byrne, *QCD, Flavor, and the de Sitter Swampland*, `2006.01144`.

[100] G. Box, *Robustness in the strategy of scientific model building*, in *Robustness in Statistics*, R. L. LAUNER and G. N. WILKINSON, eds., pp. 201 – 236, Academic Press, (1979), DOI.

[101] A. de Gouvea, D. Hernandez and T. M. P. Tait, *Criteria for Natural Hierarchies*, *Phys. Rev. D* **89** (2014) 115005 [`1402.2658`].

[102] G. Harrison, *The inner light*, 1968.

[103] M. Sendak and S. Inc, *Chicken Soup with Rice: A Book of Months*. Scholastic Incorporated, 1962.

[104] R. Haag, J. T. Lopuszanski and M. Sohnius, *All Possible Generators of Supersymmetries of the s Matrix*, *Nucl. Phys. B* **88** (1975) 257.

[105] J. Terning, *Modern supersymmetry: Dynamics and duality*. 4, 2006, 10.1093/acprof:oso/9780198567639.001.0001.

[106] S. P. Martin, *A Supersymmetry primer*, in *Perspectives on supersymmetry. Vol.2*, G. L. Kane, ed., vol. 21, pp. 1–153, (2010), `hep-ph/9709356`, DOI.

[107] D. Shih, *Supersymmetry*, Jan, 2014.

[108] D. Bertolini, J. Thaler and Z. Thomas, *Super-Tricks for Superspace*, in *Theoretical Advanced Study Institute in Elementary Particle Physics: Searching for New Physics at Small and Large Scales*, pp. 421–496, 2013, DOI [`1302.6229`].

[109] V. Nair, *A Current Algebra for Some Gauge Theory Amplitudes*, Phys. Lett. B **214** (1988) 215.

[110] N. Arkani-Hamed, F. Cachazo and J. Kaplan, *What is the Simplest Quantum Field Theory?*, JHEP **09** (2010) 016 [`0808.1446`].

[111] H. Elvang and Y.-t. Huang, *Scattering Amplitudes in Gauge Theory and Gravity*. Cambridge University Press, 4, 2015.

[112] N. Arkani-Hamed, T.-C. Huang and Y.-t. Huang, *Scattering Amplitudes For All Masses and Spins*, `1709.04891`.

[113] H. K. Dreiner, H. E. Haber and S. P. Martin, *Two-component spinor techniques and Feynman rules for quantum field theory and supersymmetry*, Phys. Rept. **494** (2010) 1 [`0812.1594`].

[114] S. Dimopoulos and H. Georgi, *Softly Broken Supersymmetry and SU(5)*, Nucl. Phys. B **193** (1981) 150.

[115] R. Davies, J. March-Russell and M. McCullough, *A Supersymmetric One Higgs Doublet Model*, JHEP **04** (2011) 108 [`1103.1647`].

[116] D. Chung, L. Everett, G. Kane, S. King, J. D. Lykken and L.-T. Wang, *The Soft supersymmetry breaking Lagrangian: Theory and applications*, Phys. Rept. **407** (2005) 1 [`hep-ph/0312378`].

[117] A. H. Chamseddine, R. L. Arnowitt and P. Nath, *Locally Supersymmetric Grand Unification*, Phys. Rev. Lett. **49** (1982) 970.

[118] R. Barbieri, S. Ferrara and C. A. Savoy, *Gauge Models with Spontaneously Broken Local Supersymmetry*, Phys. Lett. B **119** (1982) 343.

[119] L. E. Ibanez, *Locally Supersymmetric SU(5) Grand Unification*, Phys. Lett. B **118** (1982) 73.

[120] L. J. Hall, J. D. Lykken and S. Weinberg, *Supergravity as the Messenger of Supersymmetry Breaking*, Phys. Rev. D **27** (1983) 2359.

[121] N. Ohta, *GRAND UNIFIED THEORIES BASED ON LOCAL SUPERSYMMETRY*, Prog. Theor. Phys. **70** (1983) 542.

[122] J. R. Ellis, D. V. Nanopoulos and K. Tamvakis, *Grand Unification in Simple Supergravity, Phys. Lett. B* **121** (1983) 123.

[123] L. Alvarez-Gaume, J. Polchinski and M. B. Wise, *Minimal Low-Energy Supergravity, Nucl. Phys. B* **221** (1983) 495.

[124] M. Dine and W. Fischler, *A Phenomenological Model of Particle Physics Based on Supersymmetry, Phys. Lett. B* **110** (1982) 227.

[125] C. R. Nappi and B. A. Ovrut, *Supersymmetric Extension of the SU(3) x SU(2) x U(1) Model, Phys. Lett. B* **113** (1982) 175.

[126] L. Alvarez-Gaume, M. Claudson and M. B. Wise, *Low-Energy Supersymmetry, Nucl. Phys. B* **207** (1982) 96.

[127] M. Dine and A. E. Nelson, *Dynamical supersymmetry breaking at low-energies, Phys. Rev. D* **48** (1993) 1277 [`hep-ph/9303230`].

[128] M. Dine, A. E. Nelson and Y. Shirman, *Low-energy dynamical supersymmetry breaking simplified, Phys. Rev. D* **51** (1995) 1362 [`hep-ph/9408384`].

[129] M. Dine, A. E. Nelson, Y. Nir and Y. Shirman, *New tools for low-energy dynamical supersymmetry breaking, Phys. Rev. D* **53** (1996) 2658 [`hep-ph/9507378`].

[130] E. A. Mirabelli and M. E. Peskin, *Transmission of supersymmetry breaking from a four-dimensional boundary, Phys. Rev. D* **58** (1998) 065002 [`hep-th/9712214`].

[131] D. Kaplan, G. D. Kribs and M. Schmaltz, *Supersymmetry breaking through transparent extra dimensions, Phys. Rev. D* **62** (2000) 035010 [`hep-ph/9911293`].

[132] Z. Chacko, M. A. Luty, A. E. Nelson and E. Ponton, *Gaugino mediated supersymmetry breaking, JHEP* **01** (2000) 003 [`hep-ph/9911323`].

[133] M. Schmaltz and W. Skiba, *The Superpartner spectrum of gaugino mediation, Phys. Rev. D* **62** (2000) 095004 [`hep-ph/0004210`].

[134] M. Schmaltz and W. Skiba, *Minimal gaugino mediation, Phys. Rev. D* **62** (2000) 095005 [`hep-ph/0001172`].

[135] C. Csaki, J. Erlich, C. Grojean and G. D. Kribs, *4-D constructions of supersymmetric extra dimensions and gaugino mediation, Phys. Rev. D* **65** (2002) 015003 [`hep-ph/0106044`].

[136] H. Cheng, D. Kaplan, M. Schmaltz and W. Skiba, *Deconstructing Gaugino Mediation, Phys. Lett. B* **515** (2001) 395 [`hep-ph/0106098`].

[137] L. Randall and R. Sundrum, *Out of this world supersymmetry breaking*, *Nucl. Phys. B* **557** (1999) 79 [`hep-th/9810155`].

[138] G. F. Giudice, R. Rattazzi and J. D. Wells, *Quantum gravity and extra dimensions at high-energy colliders*, *Nucl. Phys. B* **544** (1999) 3 [`hep-ph/9811291`].

[139] J. A. Bagger, T. Moroi and E. Poppitz, *Anomaly mediation in supergravity theories*, *JHEP* **04** (2000) 009 [`hep-th/9911029`].

[140] K. A. Intriligator and N. Seiberg, *Lectures on Supersymmetry Breaking*, *Class. Quant. Grav.* **24** (2007) S741 [`hep-ph/0702069`].

[141] E. Witten, *Dynamical Breaking of Supersymmetry*, *Nucl. Phys. B* **188** (1981) 513.

[142] I. Affleck, M. Dine and N. Seiberg, *Dynamical Supersymmetry Breaking in Supersymmetric QCD*, *Nucl. Phys. B* **241** (1984) 493.

[143] I. Affleck, M. Dine and N. Seiberg, *Supersymmetry Breaking by Instantons*, *Phys. Rev. Lett.* **51** (1983) 1026.

[144] I. Affleck, M. Dine and N. Seiberg, *Calculable Nonperturbative Supersymmetry Breaking*, *Phys. Rev. Lett.* **52** (1984) 1677.

[145] I. Affleck, M. Dine and N. Seiberg, *Dynamical Supersymmetry Breaking in Four-Dimensions and Its Phenomenological Implications*, *Nucl. Phys. B* **256** (1985) 557.

[146] N. Seiberg, *Exact results on the space of vacua of four-dimensional SUSY gauge theories*, *Phys. Rev. D* **49** (1994) 6857 [`hep-th/9402044`].

[147] K. A. Intriligator and N. Seiberg, *Lectures on supersymmetric gauge theories and electric-magnetic duality*, *Subnucl. Ser.* **34** (1997) 237 [`hep-th/9509066`].

[148] M. A. Shifman, *Nonperturbative dynamics in supersymmetric gauge theories*, *Prog. Part. Nucl. Phys.* **39** (1997) 1 [`hep-th/9704114`].

[149] M. E. Peskin, *Duality in supersymmetric Yang-Mills theory*, in *Theoretical Advanced Study Institute in Elementary Particle Physics (TASI 96): Fields, Strings, and Duality*, pp. 729–809, 2, 1997, `hep-th/9702094`.

[150] E. Poppitz and S. P. Trivedi, *Dynamical supersymmetry breaking*, *Ann. Rev. Nucl. Part. Sci.* **48** (1998) 307 [`hep-th/9803107`].

[151] K. A. Intriligator, N. Seiberg and D. Shih, *Dynamical SUSY breaking in meta-stable vacua*, *JHEP* **04** (2006) 021 [`hep-th/0602239`].

[152] M. Shifman, *Advanced topics in quantum field theory.: A lecture course*. Cambridge Univ. Press, Cambridge, UK, 2, 2012.

[153] M. J. Strassler, *An Unorthodox introduction to supersymmetric gauge theory*, in *Theoretical Advanced Study Institute in Elementary Particle Physics (TASI 2001): Strings, Branes and EXTRA Dimensions*, pp. 561–638, 9, 2003, DOI [hep-th/0309149].

[154] A. E. Nelson and N. Seiberg, *R symmetry breaking versus supersymmetry breaking, Nucl. Phys. B* **416** (1994) 46 [hep-ph/9309299].

[155] M. A. Vasiliev, *Consistent equation for interacting gauge fields of all spins in (3+1)-dimensions, Phys. Lett. B* **243** (1990) 378.

[156] J. Wess and J. Bagger, *Supersymmetry and supergravity.* Princeton University Press, Princeton, NJ, USA, 1992.

[157] D. Z. Freedman and A. Van Proeyen, *Supergravity.* Cambridge Univ. Press, Cambridge, UK, 5, 2012.

[158] E. Abbott, *Flatland: A Romance of Many Dimensions.* Roberts Brothers, 1885.

[159] G. Nordstrom, *On the possibility of unifying the electromagnetic and the gravitational fields, Phys. Z.* **15** (1914) 504 [physics/0702221].

[160] T. Kaluza, *Zum Unitätsproblem der Physik, Int. J. Mod. Phys. D* **27** (2018) 1870001 [1803.08616].

[161] O. Klein, *Quantum Theory and Five-Dimensional Theory of Relativity. (In German and English), Z. Phys.* **37** (1926) 895.

[162] N. Arkani-Hamed, S. Dimopoulos and G. Dvali, *The Hierarchy problem and new dimensions at a millimeter, Phys. Lett. B* **429** (1998) 263 [hep-ph/9803315].

[163] C. Hoyle, U. Schmidt, B. R. Heckel, E. Adelberger, J. Gundlach, D. Kapner et al., *Submillimeter tests of the gravitational inverse square law: a search for 'large' extra dimensions, Phys. Rev. Lett.* **86** (2001) 1418 [hep-ph/0011014].

[164] E. Adelberger, J. Gundlach, B. Heckel, S. Hoedl and S. Schlamminger, *Torsion balance experiments: A low-energy frontier of particle physics, Prog. Part. Nucl. Phys.* **62** (2009) 102.

[165] J. Lee, E. Adelberger, T. Cook, S. Fleischer and B. Heckel, *New Test of the Gravitational $1/r^2$ Law at Separations down to 52 μm, Phys. Rev. Lett.* **124** (2020) 101101 [2002.11761].

[166] J. Polchinski, S. Chaudhuri and C. V. Johnson, *Notes on D-branes,* hep-th/9602052.

[167] J. Polchinski, *Tasi lectures on D-branes*, in *Theoretical Advanced Study Institute in Elementary Particle Physics (TASI 96): Fields, Strings, and Duality*, pp. 293–356, 11, 1996, `hep-th/9611050`.

[168] C. V. Johnson, *D-brane primer*, in *Theoretical Advanced Study Institute in Elementary Particle Physics (TASI 99): Strings, Branes, and Gravity*, pp. 129–350, 7, 2000, DOI [`hep-th/0007170`].

[169] I. Antoniadis, N. Arkani-Hamed, S. Dimopoulos and G. Dvali, *New dimensions at a millimeter to a Fermi and superstrings at a TeV*, *Phys. Lett. B* **436** (1998) 257 [`hep-ph/9804398`].

[170] G. Shiu and S. Tye, *TeV scale superstring and extra dimensions*, *Phys. Rev. D* **58** (1998) 106007 [`hep-th/9805157`].

[171] L. E. Ibanez, F. Marchesano and R. Rabadan, *Getting just the standard model at intersecting branes*, *JHEP* **11** (2001) 002 [`hep-th/0105155`].

[172] N. Arkani-Hamed, S. Dimopoulos and G. Dvali, *Phenomenology, astrophysics and cosmology of theories with submillimeter dimensions and TeV scale quantum gravity*, *Phys. Rev. D* **59** (1999) 086004 [`hep-ph/9807344`].

[173] T. G. Rizzo and J. D. Wells, *Electroweak precision measurements and collider probes of the standard model with large extra dimensions*, *Phys. Rev. D* **61** (2000) 016007 [`hep-ph/9906234`].

[174] S. Cullen and M. Perelstein, *SN1987A constraints on large compact dimensions*, *Phys. Rev. Lett.* **83** (1999) 268 [`hep-ph/9903422`].

[175] K. R. Dienes, E. Dudas and T. Gherghetta, *Grand unification at intermediate mass scales through extra dimensions*, *Nucl. Phys. B* **537** (1999) 47 [`hep-ph/9806292`].

[176] K. R. Dienes, E. Dudas and T. Gherghetta, *Extra space-time dimensions and unification*, *Phys. Lett. B* **436** (1998) 55 [`hep-ph/9803466`].

[177] N. Arkani-Hamed and M. Schmaltz, *Hierarchies without symmetries from extra dimensions*, *Phys. Rev. D* **61** (2000) 033005 [`hep-ph/9903417`].

[178] J. L. Hewett, *Indirect collider signals for extra dimensions*, *Phys. Rev. Lett.* **82** (1999) 4765 [`hep-ph/9811356`].

[179] P. C. Argyres, S. Dimopoulos and J. March-Russell, *Black holes and submillimeter dimensions*, *Phys. Lett. B* **441** (1998) 96 [`hep-th/9808138`].

[180] R. Emparan, G. T. Horowitz and R. C. Myers, *Black holes radiate mainly on the brane*, *Phys. Rev. Lett.* **85** (2000) 499 [`hep-th/0003118`].

[181] S. B. Giddings and S. D. Thomas, *High-energy colliders as black hole factories: The End of short distance physics*, *Phys. Rev. D* **65** (2002) 056010 [hep-ph/0106219].

[182] D. M. Eardley and S. B. Giddings, *Classical black hole production in high-energy collisions*, *Phys. Rev. D* **66** (2002) 044011 [gr-qc/0201034].

[183] V. Rubakov, *Large and infinite extra dimensions: An Introduction*, *Phys. Usp.* **44** (2001) 871 [hep-ph/0104152].

[184] R. Maartens and K. Koyama, *Brane-World Gravity*, *Living Rev. Rel.* **13** (2010) 5 [1004.3962].

[185] C. Csaki, *TASI lectures on extra dimensions and branes*, in *Theoretical Advanced Study Institute in Elementary Particle Physics (TASI 2002): Particle Physics and Cosmology: The Quest for Physics Beyond the Standard Model(s)*, pp. 605–698, 4, 2004, hep-ph/0404096.

[186] G. D. Kribs, *TASI 2004 lectures on the phenomenology of extra dimensions*, in *Theoretical Advanced Study Institute in Elementary Particle Physics: Physics in $D \geqq 4$*, pp. 633–699, 5, 2006, hep-ph/0605325.

[187] A. Perez-Lorenzana, *An Introduction to extra dimensions*, *J. Phys. Conf. Ser.* **18** (2005) 224 [hep-ph/0503177].

[188] H.-C. Cheng, *Introduction to Extra Dimensions*, in *Theoretical Advanced Study Institute in Elementary Particle Physics: Physics of the Large and the Small*, pp. 125–162, 2011, DOI [1003.1162].

[189] C. Csaki, J. Hubisz and P. Meade, *TASI lectures on electroweak symmetry breaking from extra dimensions*, in *Theoretical Advanced Study Institute in Elementary Particle Physics: Physics in $D \geqq 4$*, pp. 703–776, 10, 2005, hep-ph/0510275.

[190] M. Quiros, *New ideas in symmetry breaking*, in *Theoretical Advanced Study Institute in Elementary Particle Physics (TASI 2002): Particle Physics and Cosmology: The Quest for Physics Beyond the Standard Model(s)*, pp. 549–601, 2, 2003, hep-ph/0302189.

[191] A. Delgado, A. Pomarol and M. Quiros, *Supersymmetry and electroweak breaking from extra dimensions at the TeV scale*, *Phys. Rev. D* **60** (1999) 095008 [hep-ph/9812489].

[192] R. Barbieri, L. J. Hall and Y. Nomura, *Models of Scherk-Schwarz symmetry breaking in 5-D: Classification and calculability*, *Nucl. Phys. B* **624** (2002) 63 [hep-th/0107004].

[193] L. J. Hall and Y. Nomura, *Gauge unification in higher dimensions, Phys. Rev. D* **64** (2001) 055003 [`hep-ph/0103125`].

[194] J. Scherk and J. H. Schwarz, *Spontaneous Breaking of Supersymmetry Through Dimensional Reduction, Phys. Lett. B* **82** (1979) 60.

[195] J. Scherk and J. H. Schwarz, *How to Get Masses from Extra Dimensions, Nucl. Phys. B* **153** (1979) 61.

[196] I. Antoniadis and K. Benakli, *Limits on extra dimensions in orbifold compactifications of superstrings, Phys. Lett. B* **326** (1994) 69 [`hep-th/9310151`].

[197] I. Antoniadis, S. Dimopoulos, A. Pomarol and M. Quiros, *Soft masses in theories with supersymmetry breaking by TeV compactification, Nucl. Phys. B* **544** (1999) 503 [`hep-ph/9810410`].

[198] A. Pomarol and M. Quiros, *The Standard model from extra dimensions, Phys. Lett. B* **438** (1998) 255 [`hep-ph/9806263`].

[199] R. Barbieri, L. J. Hall and Y. Nomura, *A Constrained standard model from a compact extra dimension, Phys. Rev. D* **63** (2001) 105007 [`hep-ph/0011311`].

[200] R. Barbieri, L. J. Hall and Y. Nomura, *Softly broken supersymmetric desert from orbifold compactification, Phys. Rev. D* **66** (2002) 045025 [`hep-ph/0106190`].

[201] N. Manton, *A New Six-Dimensional Approach to the Weinberg-Salam Model, Nucl. Phys. B* **158** (1979) 141.

[202] P. Forgacs and N. Manton, *Space-Time Symmetries in Gauge Theories, Commun. Math. Phys.* **72** (1980) 15.

[203] D. Fairlie, *Higgs' Fields and the Determination of the Weinberg Angle, Phys. Lett. B* **82** (1979) 97.

[204] D. Fairlie, *Two Consistent Calculations of the Weinberg Angle, J. Phys. G* **5** (1979) L55.

[205] Y. Hosotani, *Dynamical Gauge Symmetry Breaking as the Casimir Effect, Phys. Lett. B* **129** (1983) 193.

[206] Y. Hosotani, *Dynamical Mass Generation by Compact Extra Dimensions, Phys. Lett. B* **126** (1983) 309.

[207] Y. Hosotani, *Dynamics of Nonintegrable Phases and Gauge Symmetry Breaking, Annals Phys.* **190** (1989) 233.

[208] I. Antoniadis, *A Possible new dimension at a few TeV, Phys. Lett. B* **246** (1990) 377.

[209] H. Hatanaka, T. Inami and C. Lim, *The Gauge hierarchy problem and higher dimensional gauge theories*, *Mod. Phys. Lett. A* **13** (1998) 2601 [`hep-th/9805067`].

[210] L. J. Hall, H. Murayama and Y. Nomura, *Wilson lines and symmetry breaking on orbifolds*, *Nucl. Phys. B* **645** (2002) 85 [`hep-th/0107245`].

[211] G. Burdman and Y. Nomura, *Unification of Higgs and Gauge Fields in Five Dimensions*, *Nucl. Phys. B* **656** (2003) 3 [`hep-ph/0210257`].

[212] C. Csaki, C. Grojean and H. Murayama, *Standard model Higgs from higher dimensional gauge fields*, *Phys. Rev. D* **67** (2003) 085012 [`hep-ph/0210133`].

[213] C. A. Scrucca, M. Serone and L. Silvestrini, *Electroweak symmetry breaking and fermion masses from extra dimensions*, *Nucl. Phys. B* **669** (2003) 128 [`hep-ph/0304220`].

[214] K.-w. Choi, N.-y. Haba, K.-S. Jeong, K.-i. Okumura, Y. Shimizu and M. Yamaguchi, *Electroweak symmetry breaking in supersymmetric gauge Higgs unification models*, *JHEP* **02** (2004) 037 [`hep-ph/0312178`].

[215] Y. Hosotani, S. Noda and K. Takenaga, *Dynamical gauge symmetry breaking and mass generation on the orbifold T**2 / Z(2)*, *Phys. Rev. D* **69** (2004) 125014 [`hep-ph/0403106`].

[216] Y. Hosotani, S. Noda and K. Takenaga, *Dynamical gauge-Higgs unification in the electroweak theory*, *Phys. Lett. B* **607** (2005) 276 [`hep-ph/0410193`].

[217] I. Antoniadis, K. Benakli and M. Quiros, *Finite Higgs mass without supersymmetry*, *New J. Phys.* **3** (2001) 20 [`hep-th/0108005`].

[218] K. Hasegawa, C. Lim and N. Maru, *An Attempt to solve the hierarchy problem based on gravity-gauge-Higgs unification scenario*, *Phys. Lett. B* **604** (2004) 133 [`hep-ph/0408028`].

[219] G. Cacciapaglia, C. Csaki and S. C. Park, *Fully radiative electroweak symmetry breaking*, *JHEP* **03** (2006) 099 [`hep-ph/0510366`].

[220] T. Han, I. Lewis and Z. Liu, *Colored Resonant Signals at the LHC: Largest Rate and Simplest Topology*, *JHEP* **12** (2010) 085 [`1010.4309`].

[221] K. Nishiwaki, K.-y. Oda, N. Okuda and R. Watanabe, *A Bound on Universal Extra Dimension Models from up to 2fb^{-1} of LHC Data at 7TeV*, *Phys. Lett. B* **707** (2012) 506 [`1108.1764`].

[222] G. Belanger, A. Belyaev, M. Brown, M. Kakizaki and A. Pukhov, *Testing Minimal Universal Extra Dimensions Using Higgs Boson Searches at the LHC*, *Phys. Rev. D* **87** (2013) 016008 [`1207.0798`].

[223] A. Belyaev, M. Brown, J. Moreno and C. Papineau, *Discovering Minimal Universal Extra Dimensions (MUED) at the LHC*, *JHEP* **06** (2013) 080 [1212.4858].

[224] S. Weinberg, *Implications of Dynamical Symmetry Breaking*, *Phys. Rev. D* **13** (1976) 974.

[225] L. Susskind, *Dynamics of Spontaneous Symmetry Breaking in the Weinberg-Salam Theory*, *Phys. Rev. D* **20** (1979) 2619.

[226] R. Chivukula, *Lectures on technicolor and compositeness*, in *Theoretical Advanced Study Institute in Elementary Particle Physics (TASI 2000): Flavor Physics for the Millennium*, pp. 731–772, 6, 2000, hep-ph/0011264.

[227] R. Contino, *The Higgs as a Composite Nambu-Goldstone Boson*, in *Theoretical Advanced Study Institute in Elementary Particle Physics: Physics of the Large and the Small*, pp. 235–306, 2011, DOI [1005.4269].

[228] C. T. Hill and E. H. Simmons, *Strong Dynamics and Electroweak Symmetry Breaking*, *Phys. Rept.* **381** (2003) 235 [hep-ph/0203079].

[229] E. Eichten and K. D. Lane, *Dynamical Breaking of Weak Interaction Symmetries*, *Phys. Lett. B* **90** (1980) 125.

[230] S. Dimopoulos and L. Susskind, *Mass Without Scalars*, .

[231] S. Dimopoulos and J. R. Ellis, *Challenges for Extended Technicolor Theories*, *Nucl. Phys. B* **182** (1982) 505.

[232] M. Dine, W. Fischler and M. Srednicki, *Supersymmetric Technicolor*, *Nucl. Phys. B* **189** (1981) 575.

[233] M. E. Peskin and T. Takeuchi, *A New constraint on a strongly interacting Higgs sector*, *Phys. Rev. Lett.* **65** (1990) 964.

[234] M. E. Peskin and T. Takeuchi, *Estimation of oblique electroweak corrections*, *Phys. Rev. D* **46** (1992) 381.

[235] B. Holdom, *Raising the Sideways Scale*, *Phys. Rev. D* **24** (1981) 1441.

[236] B. Holdom, *Techniodor*, *Phys. Lett. B* **150** (1985) 301.

[237] K. Yamawaki, M. Bando and K.-i. Matumoto, *Scale Invariant Technicolor Model and a Technidilaton*, *Phys. Rev. Lett.* **56** (1986) 1335.

[238] T. Akiba and T. Yanagida, *Hierarchic Chiral Condensate*, *Phys. Lett. B* **169** (1986) 432.

[239] T. W. Appelquist, D. Karabali and L. Wijewardhana, *Chiral Hierarchies and the Flavor Changing Neutral Current Problem in Technicolor*, Phys. Rev. Lett. **57** (1986) 957.

[240] T. Appelquist and L. Wijewardhana, *Chiral Hierarchies and Chiral Perturbations in Technicolor*, Phys. Rev. D **35** (1987) 774.

[241] T. Appelquist and L. Wijewardhana, *Chiral Hierarchies from Slowly Running Couplings in Technicolor Theories*, Phys. Rev. D **36** (1987) 568.

[242] V. Gorbenko, S. Rychkov and B. Zan, *Walking, Weak first-order transitions, and Complex CFTs II. Two-dimensional Potts model at $Q > 4$*, SciPost Phys. **5** (2018) 050 [`1808.04380`].

[243] V. Gorbenko, S. Rychkov and B. Zan, *Walking, Weak first-order transitions, and Complex CFTs*, JHEP **10** (2018) 108 [`1807.11512`].

[244] F. Benini, C. Iossa and M. Serone, *Conformality Loss, Walking, and 4D Complex Conformal Field Theories at Weak Coupling*, Phys. Rev. Lett. **124** (2020) 051602 [`1908.04325`].

[245] A. F. Faedo, C. Hoyos, D. Mateos and J. G. Subils, *Holographic Complex Conformal Field Theories*, Phys. Rev. Lett. **124** (2020) 161601 [`1909.04008`].

[246] O. Antipin, J. Bersini, F. Sannino, Z.-W. Wang and C. Zhang, *Charging the Walking U(N)×U(N) Higgs Theory as a Complex CFT*, `2006.10078`.

[247] K. Agashe, R. Contino and A. Pomarol, *The Minimal composite Higgs model*, Nucl. Phys. B **719** (2005) 165 [`hep-ph/0412089`].

[248] G. Panico and A. Wulzer, *The Composite Nambu-Goldstone Higgs*, vol. 913. Springer, 2016, 10.1007/978-3-319-22617-0, [`1506.01961`].

[249] S. R. Coleman, J. Wess and B. Zumino, *Structure of phenomenological Lagrangians. 1.*, Phys. Rev. **177** (1969) 2239.

[250] J. Callan, Curtis G., S. R. Coleman, J. Wess and B. Zumino, *Structure of phenomenological Lagrangians. 2.*, Phys. Rev. **177** (1969) 2247.

[251] M. Schmaltz and D. Tucker-Smith, *Little Higgs review*, Ann. Rev. Nucl. Part. Sci. **55** (2005) 229 [`hep-ph/0502182`].

[252] S. Scherer, *Introduction to chiral perturbation theory*, Adv. Nucl. Phys. **27** (2003) 277 [`hep-ph/0210398`].

[253] S. R. Coleman and E. J. Weinberg, *Radiative Corrections as the Origin of Spontaneous Symmetry Breaking*, Phys. Rev. D **7** (1973) 1888.

[254] N. Arkani-Hamed, A. Cohen, E. Katz, A. Nelson, T. Gregoire and J. G. Wacker, *The Minimal moose for a little Higgs*, *JHEP* **08** (2002) 021 [hep-ph/0206020].

[255] N. Arkani-Hamed, A. Cohen, E. Katz and A. Nelson, *The Littlest Higgs*, *JHEP* **07** (2002) 034 [hep-ph/0206021].

[256] D. E. Kaplan and M. Schmaltz, *The Little Higgs from a simple group*, *JHEP* **10** (2003) 039 [hep-ph/0302049].

[257] M. Schmaltz, *Physics beyond the standard model (theory): Introducing the little Higgs*, *Nucl. Phys. B Proc. Suppl.* **117** (2003) 40 [hep-ph/0210415].

[258] M. Perelstein, *Little Higgs models and their phenomenology*, *Prog. Part. Nucl. Phys.* **58** (2007) 247 [hep-ph/0512128].

[259] N. Arkani-Hamed, A. G. Cohen and H. Georgi, *(De)constructing dimensions*, *Phys. Rev. Lett.* **86** (2001) 4757 [hep-th/0104005].

[260] N. Arkani-Hamed, A. G. Cohen and H. Georgi, *Electroweak symmetry breaking from dimensional deconstruction*, *Phys. Lett. B* **513** (2001) 232 [hep-ph/0105239].

[261] T. H. Huxley, *Biogenesis and abiogenesis [1870]*, in *Collected Essays*, vol. 8 of *Cambridge Library Collection - Philosophy*, p. 229–271, Cambridge University Press, (2011), DOI.

[262] R. Barbieri and A. Strumia, *The 'LEP paradox'*, in *4th Rencontres du Vietnam: Physics at Extreme Energies (Particle Physics and Astrophysics)*, 7, 2000, hep-ph/0007265.

[263] H.-C. Cheng and I. Low, *TeV symmetry and the little hierarchy problem*, *JHEP* **09** (2003) 051 [hep-ph/0308199].

[264] J. D. Wells, *Implications of supersymmetry breaking with a little hierarchy between gauginos and scalars*, in *11th International Conference on Supersymmetry and the Unification of Fundamental Interactions*, 6, 2003, hep-ph/0306127.

[265] N. Arkani-Hamed and S. Dimopoulos, *Supersymmetric unification without low energy supersymmetry and signatures for fine-tuning at the LHC*, *JHEP* **06** (2005) 073 [hep-th/0405159].

[266] G. Giudice and A. Romanino, *Split supersymmetry*, *Nucl. Phys. B* **699** (2004) 65 [hep-ph/0406088].

[267] 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].

[268] N. Craig, A. Katz, M. Strassler and R. Sundrum, *Naturalness in the Dark at the LHC*, *JHEP* **07** (2015) 105 [1501.05310].

[269] G. Burdman, Z. Chacko, R. Harnik, L. de Lima and C. B. Verhaaren, *Colorless Top Partners, a 125 GeV Higgs, and the Limits on Naturalness*, *Phys. Rev.* **D91** (2015) 055007 [1411.3310].

[270] S. Chang, L. J. Hall and N. Weiner, *A Supersymmetric twin Higgs*, *Phys. Rev.* **D75** (2007) 035009 [hep-ph/0604076].

[271] N. Craig and K. Howe, *Doubling down on naturalness with a supersymmetric twin Higgs*, *JHEP* **03** (2014) 140 [1312.1341].

[272] Z. Chacko, H.-S. Goh and R. Harnik, *A Twin Higgs model from left-right symmetry*, *JHEP* **01** (2006) 108 [hep-ph/0512088].

[273] G. Burdman, Z. Chacko, H.-S. Goh and R. Harnik, *Folded supersymmetry and the LEP paradox*, *JHEP* **02** (2007) 009 [hep-ph/0609152].

[274] D. Poland and J. Thaler, *The Dark Top*, *JHEP* **11** (2008) 083 [0808.1290].

[275] H. Cai, H.-C. Cheng and J. Terning, *A Quirky Little Higgs Model*, *JHEP* **05** (2009) 045 [0812.0843].

[276] N. Craig, S. Knapen and P. Longhi, *Neutral Naturalness from Orbifold Higgs Models*, *Phys. Rev. Lett.* **114** (2015) 061803 [1410.6808].

[277] B. Batell and M. McCullough, *Neutrino Masses from Neutral Top Partners*, *Phys. Rev.* **D92** (2015) 073018 [1504.04016].

[278] D. Curtin and P. Saraswat, *Towards a No-Lose Theorem for Naturalness*, *Phys. Rev.* **D93** (2016) 055044 [1509.04284].

[279] H.-C. Cheng, S. Jung, E. Salvioni and Y. Tsai, *Exotic Quarks in Twin Higgs Models*, *JHEP* **03** (2016) 074 [1512.02647].

[280] N. Craig, S. Knapen, P. Longhi and M. Strassler, *The Vector-like Twin Higgs*, *JHEP* **07** (2016) 002 [1601.07181].

[281] T. Cohen, N. Craig, G. F. Giudice and M. McCullough, *The Hyperbolic Higgs*, *JHEP* **05** (2018) 091 [1803.03647].

[282] H.-C. Cheng, L. Li, E. Salvioni and C. B. Verhaaren, *Singlet Scalar Top Partners from Accidental Supersymmetry*, *JHEP* **05** (2018) 057 [1803.03651].

[283] J. Serra, S. Stelzl, R. Torre and A. Weiler, *Hypercharged Naturalness*, 1905.02203.

[284] M. Schmaltz, *Duality of nonsupersymmetric large N gauge theories*, *Phys. Rev.* **D59** (1999) 105018 [`hep-th/9805218`].

[285] M. Bershadsky and A. Johansen, *Large N limit of orbifold field theories*, *Nucl. Phys.* **B536** (1998) 141 [`hep-th/9803249`].

[286] S. Kachru and E. Silverstein, *4-D conformal theories and strings on orbifolds*, *Phys. Rev. Lett.* **80** (1998) 4855 [`hep-th/9802183`].

[287] N. Craig, S. Knapen and P. Longhi, *The Orbifold Higgs*, *JHEP* **03** (2015) 106 [`1411.7393`].

[288] W. Shakespeare, *The Tragedie of Hamlet, Prince of Denmarke*, Mr. William Shakespeare's Comedies, Histories, Tragedies. Edward Blount & William and Isaac Jaggard, 1623.

[289] S. R. Coleman, *Why There Is Nothing Rather Than Something: A Theory of the Cosmological Constant*, *Nucl. Phys. B* **310** (1988) 643.

[290] A. G. Cohen, D. B. Kaplan and A. E. Nelson, *Effective field theory, black holes, and the cosmological constant*, *Phys. Rev. Lett.* **82** (1999) 4971 [`hep-th/9803132`].

[291] J. F. Donoghue, *General relativity as an effective field theory: The leading quantum corrections*, *Phys. Rev. D* **50** (1994) 3874 [`gr-qc/9405057`].

[292] R. Feynman, *Quantum theory of gravitation*, *Acta Phys. Polon.* **24** (1963) 697.

[293] B. S. DeWitt, *Quantum Theory of Gravity. 2. The Manifestly Covariant Theory*, *Phys. Rev.* **162** (1967) 1195.

[294] S. B. Giddings, *The gravitational S-matrix: Erice lectures*, *Subnucl. Ser.* **48** (2013) 93 [`1105.2036`].

[295] S. W. Hawking, *Particle Creation by Black Holes*, *Commun. Math. Phys.* **43** (1975) 199.

[296] C. G. Torre, *Gravitational observables and local symmetries*, *Phys. Rev.* **D48** (1993) R2373 [`gr-qc/9306030`].

[297] S. B. Giddings, D. Marolf and J. B. Hartle, *Observables in effective gravity*, *Phys. Rev.* **D74** (2006) 064018 [`hep-th/0512200`].

[298] W. Donnelly and S. B. Giddings, *Diffeomorphism-invariant observables and their nonlocal algebra*, *Phys. Rev.* **D93** (2016) 024030 [`1507.07921`].

[299] W. Donnelly and S. B. Giddings, *Observables, gravitational dressing, and obstructions to locality and subsystems*, *Phys. Rev.* **D94** (2016) 104038 [`1607.01025`].

[300] S. B. Giddings and A. Kinsella, *Gauge-invariant observables, gravitational dressings, and holography in AdS*, *JHEP* **11** (2018) 074 [`1802.01602`].

[301] A. Almheiri, D. Marolf, J. Polchinski and J. Sully, *Black Holes: Complementarity or Firewalls?*, *JHEP* **02** (2013) 062 [`1207.3123`].

[302] G. Penington, *Entanglement Wedge Reconstruction and the Information Paradox*, `1905.08255`.

[303] A. Almheiri, R. Mahajan, J. Maldacena and Y. Zhao, *The Page curve of Hawking radiation from semiclassical geometry*, *JHEP* **03** (2020) 149 [`1908.10996`].

[304] A. Almheiri, N. Engelhardt, D. Marolf and H. Maxfield, *The entropy of bulk quantum fields and the entanglement wedge of an evaporating black hole*, *JHEP* **12** (2019) 063 [`1905.08762`].

[305] A. Almheiri, T. Hartman, J. Maldacena, E. Shaghoulian and A. Tajdini, *Replica Wormholes and the Entropy of Hawking Radiation*, *JHEP* **05** (2020) 013 [`1911.12333`].

[306] G. Penington, S. H. Shenker, D. Stanford and Z. Yang, *Replica wormholes and the black hole interior*, `1911.11977`.

[307] D. Marolf and H. Maxfield, *Transcending the ensemble: baby universes, spacetime wormholes, and the order and disorder of black hole information*, `2002.08950`.

[308] C. Vafa, *The String landscape and the swampland*, `hep-th/0509212`.

[309] N. Arkani-Hamed, L. Motl, A. Nicolis and C. Vafa, *The String landscape, black holes and gravity as the weakest force*, *JHEP* **06** (2007) 060 [`hep-th/0601001`].

[310] D. Marolf and R. D. Sorkin, *On the status of highly entropic objects*, *Phys. Rev. D* **69** (2004) 024014 [`hep-th/0309218`].

[311] H. S. Snyder, *Quantized space-time*, *Phys. Rev.* **71** (1947) 38.

[312] A. Connes, *Noncommutative geometry*. Academic Press, 1994.

[313] A. Connes, M. R. Douglas and A. S. Schwarz, *Noncommutative geometry and matrix theory: Compactification on tori*, *JHEP* **02** (1998) 003 [`hep-th/9711162`].

[314] M. R. Douglas and C. M. Hull, *D-branes and the noncommutative torus*, *JHEP* **02** (1998) 008 [`hep-th/9711165`].

[315] N. Seiberg and E. Witten, *String theory and noncommutative geometry*, *JHEP* **09** (1999) 032 [`hep-th/9908142`].

[316] R. C. Myers, *Dielectric branes*, *JHEP* **12** (1999) 022 [`hep-th/9910053`].

[317] D. Adams, *The Restaurant at the End of the Universe*, Adams, Douglas: +The hitch hiker's guide to the galaxy. Pan, 1980.

[318] PLANCK collaboration, *Planck 2015 results. XIII. Cosmological parameters*, Astron. Astrophys. **594** (2016) A13 [1502.01589].

[319] R. H. Cyburt, B. D. Fields, K. A. Olive and T.-H. Yeh, *Big Bang Nucleosynthesis: 2015*, Rev. Mod. Phys. **88** (2016) 015004 [1505.01076].

[320] M. Geller and O. Telem, *Holographic Twin Higgs Model*, Phys. Rev. Lett. **114** (2015) 191801 [1411.2974].

[321] R. Barbieri, D. Greco, R. Rattazzi and A. Wulzer, *The Composite Twin Higgs scenario*, JHEP **08** (2015) 161 [1501.07803].

[322] M. Low, A. Tesi and L.-T. Wang, *Twin Higgs mechanism and a composite Higgs boson*, Phys. Rev. D **91** (2015) 095012 [1501.07890].

[323] R. Barbieri, L. J. Hall and K. Harigaya, *Minimal Mirror Twin Higgs*, JHEP **11** (2016) 172 [1609.05589].

[324] C. Csaki, E. Kuflik and S. Lombardo, *Viable Twin Cosmology from Neutrino Mixing*, Phys. Rev. **D96** (2017) 055013 [1703.06884].

[325] D. Liu and N. Weiner, *A Portalino to the Twin Sector*, 1905.00861.

[326] K. Harigaya, R. McGehee, H. Murayama and K. Schutz, *A Predictive Mirror Twin Higgs with Small* $\mathbf{Z}_2$ *Breaking*, 1905.08798.

[327] J. D. Clarke, *Constraining portals with displaced Higgs decay searches at the LHC*, JHEP **10** (2015) 061 [1505.00063].

[328] D. Curtin and C. B. Verhaaren, *Discovering Uncolored Naturalness in Exotic Higgs Decays*, JHEP **12** (2015) 072 [1506.06141].

[329] C. Csaki, E. Kuflik, S. Lombardo and O. Slone, *Searching for displaced Higgs boson decays*, Phys. Rev. **D92** (2015) 073008 [1508.01522].

[330] A. Pierce, B. Shakya, Y. Tsai and Y. Zhao, *Searching for confining hidden valleys at LHCb, ATLAS, and CMS*, Phys. Rev. **D97** (2018) 095033 [1708.05389].

[331] C. Kilic, S. Najjari and C. B. Verhaaren, *Discovering the Twin Higgs Boson with Displaced Decays*, Phys. Rev. **D99** (2019) 075029 [1812.08173].

[332] L. Li and Y. Tsai, *Detector-size Upper Bounds on Dark Hadron Lifetime from Cosmology*, JHEP **05** (2019) 072 [1901.09936].

[333] I. Garcia Garcia, R. Lasenby and J. March-Russell, *Twin Higgs WIMP Dark Matter*, *Phys. Rev.* **D92** (2015) 055034 [`1505.07109`].

[334] N. Craig and A. Katz, *The Fraternal WIMP Miracle*, *JCAP* **1510** (2015) 054 [`1505.07113`].

[335] I. Garcia Garcia, R. Lasenby and J. March-Russell, *Twin Higgs Asymmetric Dark Matter*, *Phys. Rev. Lett.* **115** (2015) 121801 [`1505.07410`].

[336] M. Farina, *Asymmetric Twin Dark Matter*, *JCAP* **1511** (2015) 017 [`1506.03520`].

[337] M. Freytsis, S. Knapen, D. J. Robinson and Y. Tsai, *Gamma-rays from Dark Showers with Twin Higgs Models*, *JHEP* **05** (2016) 018 [`1601.07556`].

[338] M. Farina, A. Monteux and C. S. Shin, *Twin mechanism for baryon and dark matter asymmetries*, *Phys. Rev.* **D94** (2016) 035017 [`1604.08211`].

[339] V. Prilepina and Y. Tsai, *Reconciling Large And Small-Scale Structure In Twin Higgs Models*, *JHEP* **09** (2017) 033 [`1611.05879`].

[340] R. Barbieri, L. J. Hall and K. Harigaya, *Effective Theory of Flavor for Minimal Mirror Twin Higgs*, *JHEP* **10** (2017) 015 [`1706.05548`].

[341] Y. Hochberg, E. Kuflik and H. Murayama, *Twin Higgs model with strongly interacting massive particle dark matter*, *Phys. Rev.* **D99** (2019) 015005 [`1805.09345`].

[342] H.-C. Cheng, L. Li and R. Zheng, *Coscattering/Coannihilation Dark Matter in a Fraternal Twin Higgs Model*, `1805.12139`.

[343] J. Terning, C. B. Verhaaren and K. Zora, *Composite Twin Dark Matter*, *Phys. Rev.* **D99** (2019) 095020 [`1902.08211`].

[344] K. Earl, C. S. Fong, T. Gregoire and A. Tonero, *Mirror Dirac leptogenesis*, `1903.12192`.

[345] M. Drees, F. Hajkarim and E. R. Schmitz, *The Effects of QCD Equation of State on the Relic Density of WIMP Dark Matter*, *JCAP* **1506** (2015) 025 [`1503.03513`].

[346] J. Edsjo and P. Gondolo, *Neutralino relic density including coannihilations*, *Phys. Rev. D* **56** (1997) 1879 [`hep-ph/9704361`].

[347] P. Adshead, Y. Cui and J. Shelton, *Chilly Dark Sectors and Asymmetric Reheating*, *JHEP* **06** (2016) 016 [`1604.02458`].

[348] G. Mangano, G. Miele, S. Pastor, T. Pinto, O. Pisanti and P. D. Serpico, *Relic neutrino decoupling including flavor oscillations*, *Nucl. Phys. B* **729** (2005) 221 [hep-ph/0506164].

[349] Z. Hou, R. Keisler, L. Knox, M. Millea and C. Reichardt, *How Massless Neutrinos Affect the Cosmic Microwave Background Damping Tail*, *Phys. Rev. D* **87** (2013) 083008 [1104.2333].

[350] CMB-S4 collaboration, *CMB-S4 Science Book, First Edition*, 1610.02743.

[351] C. Brust, D. E. Kaplan and M. T. Walters, *New Light Species and the CMB*, *JHEP* **12** (2013) 058 [1303.5379].

[352] S. Dodelson, A. Melchiorri and A. Slosar, *Is cosmology compatible with sterile neutrinos?*, *Phys. Rev. Lett.* **97** (2006) 041301 [astro-ph/0511500].

[353] D. Baumann, D. Green, J. Meyers and B. Wallisch, *Phases of New Physics in the CMB*, *JCAP* **01** (2016) 007 [1508.06342].

[354] M. Archidiacono, S. Hannestad, A. Mirizzi, G. Raffelt and Y. Y. Wong, *Axion hot dark matter bounds after Planck*, *JCAP* **10** (2013) 020 [1307.0615].

[355] T. D. Jacques, L. M. Krauss and C. Lunardini, *Additional Light Sterile Neutrinos and Cosmology*, *Phys. Rev. D* **87** (2013) 083515 [1301.3119].

[356] J. Lesgourgues and S. Pastor, *Neutrino mass from Cosmology*, *Adv. High Energy Phys.* **2012** (2012) 608515 [1212.6154].

[357] S. Dodelson, E. Gates and A. Stebbins, *Cold + hot dark matter and the cosmic microwave background*, *Astrophys. J.* **467** (1996) 10 [astro-ph/9509147].

[358] E. Di Valentino, S. Gariazzo, M. Gerbino, E. Giusarma and O. Mena, *Dark Radiation and Inflationary Freedom after Planck 2015*, *Phys. Rev. D* **93** (2016) 083523 [1601.07557].

[359] M. Costanzi, B. Sartoris, M. Viel and S. Borgani, *Neutrino constraints: what large-scale structure and CMB data are telling us?*, *JCAP* **10** (2014) 081 [1407.8338].

[360] S. Gariazzo, C. Giunti, M. Laveder, Y. Li and E. Zavanin, *Light sterile neutrinos*, *J. Phys. G* **43** (2016) 033001 [1507.08204].

[361] N. Arkani-Hamed, T. Cohen, R. T. D'Agnolo, A. Hook, H. D. Kim and D. Pinner, *Solving the Hierarchy Problem at Reheating with a Large Number of Degrees of Freedom*, *Phys. Rev. Lett.* **117** (2016) 251801 [1607.06821].

[362] M. Reece and T. Roxlo, *Nonthermal production of dark radiation and dark matter, JHEP* **09** (2016) 096 [1511.06768].

[363] L. Randall, J. Scholtz and J. Unwin, *Flooded Dark Matter and S Level Rise, JHEP* **03** (2016) 011 [1509.08477].

[364] E. W. Kolb and M. S. Turner, *The Early Universe*, vol. 69. 1990.

[365] P. de Salas, M. Lattanzi, G. Mangano, G. Miele, S. Pastor and O. Pisanti, *Bounds on very low reheating scenarios after Planck, Phys. Rev. D* **92** (2015) 123534 [1511.00672].

[366] A. Djouadi, *The Anatomy of electro-weak symmetry breaking. I: The Higgs boson in the standard model, Phys. Rept.* **457** (2008) 1 [hep-ph/0503172].

[367] M. Drees and K.-i. Hikasa, *Heavy Quark Thresholds in Higgs Physics, Phys. Rev. D* **41** (1990) 1547.

[368] M. Baumgart and A. Katz, *Implications of a New Light Scalar Near the Bottomonium Regime, JHEP* **08** (2012) 133 [1204.6032].

[369] J. D. Clarke, R. Foot and R. R. Volkas, *Phenomenology of a very light scalar (100 MeV < m_h < 10 GeV) mixing with the SM Higgs, JHEP* **02** (2014) 123 [1310.8042].

[370] U. Haisch and J. F. Kamenik, *Searching for new spin-0 resonances at LHCb, Phys. Rev. D* **93** (2016) 055047 [1601.05110].

[371] TOPICAL CONVENERS: K.N. ABAZAJIAN, J.E. CARLSTROM, A.T. LEE collaboration, *Neutrino Physics from the Cosmic Microwave Background and Large Scale Structure, Astropart. Phys.* **63** (2015) 66 [1309.5383].

[372] PLANCK collaboration, *Planck 2015 results. XX. Constraints on inflation, Astron. Astrophys.* **594** (2016) A20 [1502.02114].

[373] Z. Berezhiani, A. Dolgov and R. Mohapatra, *Asymmetric inflationary reheating and the nature of mirror universe, Phys. Lett. B* **375** (1996) 26 [hep-ph/9511221].

[374] J. Silk and M. S. Turner, *Double Inflation, Phys. Rev. D* **35** (1987) 419.

[375] D. Polarski and A. A. Starobinsky, *Spectra of perturbations produced by double inflation with an intermediate matter dominated stage, Nucl. Phys. B* **385** (1992) 623.

[376] D. Wands, N. Bartolo, S. Matarrese and A. Riotto, *An Observational test of two-field inflation, Phys. Rev. D* **66** (2002) 043520 [astro-ph/0205253].

[377] S. E. Hong, H.-J. Lee, Y. J. Lee, E. D. Stewart and H. Zoe, *Effects of thermal inflation on small scale density perturbations*, *JCAP* **06** (2015) 002 [1503.08938].

[378] W. Hu and T. Okamoto, *Principal power of the CMB*, *Phys. Rev. D* **69** (2004) 043004 [astro-ph/0308049].

[379] Z. Chacko, N. Craig, P. J. Fox and R. Harnik, *Cosmology in Mirror Twin Higgs and Neutrino Masses*, *JHEP* **07** (2017) 023 [1611.07975].

[380] T. Asaka, K. Ishiwata and T. Moroi, *Right-handed sneutrino as cold dark matter*, *Phys. Rev.* **D73** (2006) 051301 [hep-ph/0512118].

[381] S. Gopalakrishna, A. de Gouvea and W. Porod, *Right-handed sneutrinos as nonthermal dark matter*, *JCAP* **0605** (2006) 005 [hep-ph/0602027].

[382] T. Asaka, K. Ishiwata and T. Moroi, *Right-handed sneutrino as cold dark matter of the universe*, *Phys. Rev.* **D75** (2007) 065001 [hep-ph/0612211].

[383] V. Page, *Non-thermal right-handed sneutrino dark matter and the Omega(DM)/Omega(b) problem*, *JHEP* **04** (2007) 021 [hep-ph/0701266].

[384] L. J. Hall, K. Jedamzik, J. March-Russell and S. M. West, *Freeze-In Production of FIMP Dark Matter*, *JHEP* **03** (2010) 080 [0911.1120].

[385] X. Chu, T. Hambye and M. H. G. Tytgat, *The Four Basic Ways of Creating Dark Matter Through a Portal*, *JCAP* **1205** (2012) 034 [1112.0493].

[386] N. Bernal, M. Heikinheimo, T. Tenkanen, K. Tuominen and V. Vaskonen, *The Dawn of FIMP Dark Matter: A Review of Models and Constraints*, *Int. J. Mod. Phys.* **A32** (2017) 1730023 [1706.07442].

[387] C. Dvorkin, T. Lin and K. Schutz, *Making dark matter out of light: freeze-in from plasma effects*, *Phys. Rev.* **D99** (2019) 115009 [1902.08623].

[388] S. Tulin and H.-B. Yu, *Dark Matter Self-interactions and Small Scale Structure*, *Phys. Rept.* **730** (2018) 1 [1705.02358].

[389] PLANCK collaboration, *Planck 2018 results. VI. Cosmological parameters*, 1807.06209.

[390] CMB-S4 collaboration, *CMB-S4 Science Book, First Edition*, 1610.02743.

[391] Z. Chacko, D. Curtin, M. Geller and Y. Tsai, *Cosmological Signatures of a Mirror Twin Higgs*, 1803.03263.

[392] Z. Chacko, C. Kilic, S. Najjari and C. B. Verhaaren, *Collider Signals of the Mirror Twin Higgs through the Hypercharge Portal*, 1904.11990.

[393] A. S. Goldhaber and M. M. Nieto, *Photon and Graviton Mass Limits*, *Rev. Mod. Phys.* **82** (2010) 939 [`0809.1003`].

[394] M. Reece, *Photon Masses in the Landscape and the Swampland*, `1808.09966`.

[395] N. Craig and I. Garcia Garcia, *Rescuing Massive Photons from the Swampland*, *JHEP* **11** (2018) 067 [`1810.05647`].

[396] B. Batell and C. B. Verhaaren, *Breaking Mirror Twin Hypercharge*, `1904.10468`.

[397] M. Blennow, E. Fernandez-Martinez and B. Zaldivar, *Freeze-in through portals*, *JCAP* **1401** (2014) 003 [`1309.7348`].

[398] J. H. Chang, R. Essig and A. Reinert, *Light(ly)-coupled Dark Matter in the keV Range: Freeze-In and Constraints*, `1911.03389`.

[399] L. Forestell and D. E. Morrissey, *Infrared Effects of Ultraviolet Operators on Dark Matter Freeze-In*, `1811.08905`.

[400] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, *JHEP* **07** (2014) 079 [`1405.0301`].

[401] A. Alloul, N. D. Christensen, C. Degrande, C. Duhr and B. Fuks, *FeynRules 2.0 - A complete toolbox for tree-level phenomenology*, *Comput. Phys. Commun.* **185** (2014) 2250 [`1310.1921`].

[402] J. H. Chang, R. Essig and S. D. McDermott, *Revisiting Supernova 1987A Constraints on Dark Photons*, *JHEP* **01** (2017) 107 [`1611.03864`].

[403] E. Hardy and R. Lasenby, *Stellar cooling bounds on new light particles: plasma mixing effects*, *JHEP* **02** (2017) 033 [`1611.05852`].

[404] D. Clowe, A. Gonzalez and M. Markevitch, *Weak lensing mass reconstruction of the interacting cluster 1E0657-558: Direct evidence for the existence of dark matter*, *Astrophys. J.* **604** (2004) 596 [`astro-ph/0312273`].

[405] M. Markevitch, A. H. Gonzalez, D. Clowe, A. Vikhlinin, L. David, W. Forman et al., *Direct constraints on the dark matter self-interaction cross-section from the merging galaxy cluster 1E0657-56*, *Astrophys. J.* **606** (2004) 819 [`astro-ph/0309303`].

[406] S. W. Randall, M. Markevitch, D. Clowe, A. H. Gonzalez and M. Bradac, *Constraints on the Self-Interaction Cross-Section of Dark Matter from Numerical Simulations of the Merging Galaxy Cluster 1E 0657-56*, *Astrophys. J.* **679** (2008) 1173 [`0704.0261`].

[407] A. Ahmed, *Heavy Higgs of the Twin Higgs Models*, *JHEP* **02** (2018) 048 [1711.03107].

[408] Z. Chacko, C. Kilic, S. Najjari and C. B. Verhaaren, *Testing the Scalar Sector of the Twin Higgs Model at Colliders*, *Phys. Rev.* **D97** (2018) 055031 [1711.05300].

[409] J. Tolkien, *The Hobbit, Or, There and Back Again*, Lord of the Rings. George Allen Unwin, 1937.

[410] ATLAS collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, *Phys. Lett.* **B716** (2012) 1 [1207.7214].

[411] CMS collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, *Phys. Lett.* **B716** (2012) 30 [1207.7235].

[412] CEPC Study Group collaboration, *CEPC Conceptual Design Report: Volume 1 - Accelerator*, 1809.00285.

[413] CEPC Study Group collaboration, *CEPC Conceptual Design Report: Volume 2 - Physics & Detector*, 1811.10545.

[414] TLEP Design Study Working Group collaboration, *First Look at the Physics Case of TLEP*, *JHEP* **01** (2014) 164 [1308.6176].

[415] T. Behnke, J. E. Brau, B. Foster, J. Fuster, M. Harrison, J. M. Paterson et al., *The International Linear Collider Technical Design Report - Volume 1: Executive Summary*, 1306.6327.

[416] H. Baer, T. Barklow, K. Fujii, Y. Gao, A. Hoang, S. Kanemura et al., *The International Linear Collider Technical Design Report - Volume 2: Physics*, 1306.6352.

[417] M. Aicheler, P. Burrows, M. Draper, T. Garvey, P. Lebrun, K. Peach et al., *A Multi-TeV Linear Collider Based on CLIC Technology*, .

[418] J. de Blas et al., *The CLIC Potential for New Physics*, 1812.02093.

[419] D. Curtin et al., *Exotic decays of the 125 GeV Higgs boson*, *Phys. Rev.* **D90** (2014) 075004 [1312.4992].

[420] LHC Higgs Cross Section Working Group collaboration, *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*, 1610.07922.

[421] Z. Liu, L.-T. Wang and H. Zhang, *Exotic decays of the 125 GeV Higgs boson at future e^+e^- lepton colliders*, *Chin. Phys.* **C41** (2017) 063102 [1612.09284].

[422] M. J. Strassler and K. M. Zurek, *Echoes of a hidden valley at hadron colliders*, Phys. Lett. **B651** (2007) 374 [hep-ph/0604261].

[423] M. J. Strassler and K. M. Zurek, *Discovering the Higgs through highly-displaced vertices*, Phys. Lett. **B661** (2008) 263 [hep-ph/0605193].

[424] T. Han, Z. Si, K. M. Zurek and M. J. Strassler, *Phenomenology of hidden valleys at hadron colliders*, JHEP **07** (2008) 008 [0712.2041].

[425] Y. Cui and B. Shuve, *Probing Baryogenesis with Displaced Vertices at the LHC*, JHEP **02** (2015) 049 [1409.6729].

[426] D. Curtin et al., *Long-Lived Particles at the Energy Frontier: The MATHUSLA Physics Case*, 1806.07396.

[427] CMS collaboration, *Search for long-lived particles that decay into final states containing two electrons or two muons in proton-proton collisions at $\sqrt{s} = 8$ TeV*, Phys. Rev. **D91** (2015) 052012 [1411.6977].

[428] ATLAS collaboration, *Search for the production of a long-lived neutral particle decaying within the ATLAS hadronic calorimeter in association with a Z boson from pp collisions at $\sqrt{s} = 13$ TeV*, Submitted to: Phys. Rev. Lett. (2018) [1811.02542].

[429] LHCʙ collaboration, *Updated search for long-lived particles decaying to jet pairs*, Eur. Phys. J. **C77** (2017) 812 [1705.07332].

[430] L. Lee, C. Ohm, A. Soffer and T.-T. Yu, *Collider Searches for Long-Lived Particles Beyond the Standard Model*, 1810.12602.

[431] J. Liu, Z. Liu and L.-T. Wang, *Long-lived particles at the LHC: catching them in time*, 1805.05957.

[432] Y. Gershtein, *CMS Hardware Track Trigger: New Opportunities for Long-Lived Particle Searches at the HL-LHC*, Phys. Rev. **D96** (2017) 035027 [1705.04321].

[433] CMS collaboration, *First Level Track Jet Trigger for Displaced Jets at High Luminosity LHC*, .

[434] LHCʙ collaboration, *Prospects for searches for long-lived particles after the LHCb detector upgrades*, .

[435] M. Kucharczyk and T. Wojton, *Hidden Valley searches at CLIC*, .

[436] S. Antusch, E. Cazzato and O. Fischer, *Displaced vertex searches for sterile neutrinos at future lepton colliders*, JHEP **12** (2016) 007 [1604.02420].

[437] D. Curtin, K. Deshpande, O. Fischer and J. Zurita, *New Physics Opportunities for Long-Lived Particles at Electron-Proton Colliders*, JHEP **07** (2018) 024 [1712.07135].

[438] T. Sjostrand, S. Mrenna and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852 [0710.3820].

[439] DELPHES 3 collaboration, *DELPHES 3, A modular framework for fast simulation of a generic collider experiment*, JHEP **02** (2014) 057 [1307.6346].

[440] R. Brun and F. Rademakers, *Root — an object oriented data analysis framework*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **389** (1997) 81 .

[441] CMS collaboration, *Search for Long-Lived Neutral Particles Decaying to Quark-Antiquark Pairs in Proton-Proton Collisions at $\sqrt{s} = 8$ TeV*, Phys. Rev. **D91** (2015) 012007 [1411.6530].

[442] FCC-EE collaboration, *FCC-ee Machine Parameters*, 2017.

[443] V. Silveira and A. Zee, *SCALAR PHANTOMS*, Phys. Lett. **161B** (1985) 136.

[444] J. McDonald, *Gauge singlet scalars as cold dark matter*, Phys. Rev. **D50** (1994) 3637 [hep-ph/0702143].

[445] C. P. Burgess, M. Pospelov and T. ter Veldhuis, *The Minimal model of nonbaryonic dark matter: A Singlet scalar*, Nucl. Phys. **B619** (2001) 709 [hep-ph/0011335].

[446] X.-G. He, T. Li, X.-Q. Li, J. Tandean and H.-C. Tsai, *Constraints on Scalar Dark Matter from Direct Experimental Searches*, Phys. Rev. **D79** (2009) 023521 [0811.0658].

[447] M. Gonderinger, Y. Li, H. Patel and M. J. Ramsey-Musolf, *Vacuum Stability, Perturbativity, and Scalar Singlet Dark Matter*, JHEP **01** (2010) 053 [0910.3167].

[448] Y. Mambrini, *Higgs searches and singlet scalar dark matter: Combined constraints from XENON 100 and the LHC*, Phys. Rev. **D84** (2011) 115017 [1108.0671].

[449] C. Englert and M. McCullough, *Modified Higgs Sectors and NLO Associated Production*, JHEP **07** (2013) 168 [1303.1526].

[450] N. Craig, C. Englert and M. McCullough, *New Probe of Naturalness*, Phys. Rev. Lett. **111** (2013) 121803 [1305.5251].

[451] J. Cameron, *Is commuting dead?*, The Connecticut Mirror (2020) .

[452] V. O. Rivelles, *Noncommutative field theories and gravity, Phys. Lett.* **B558** (2003) 191 [`hep-th/0212262`].

[453] H. S. Yang, *Exact Seiberg-Witten map and induced gravity from noncommutativity, Mod. Phys. Lett.* **A21** (2006) 2637 [`hep-th/0402002`].

[454] H. S. Yang, *On The Correspondence Between Noncommuative Field Theory And Gravity, Mod. Phys. Lett.* **A22** (2007) 1119 [`hep-th/0612231`].

[455] H. S. Yang, *Instantons and Emergent Geometry, EPL* **88** (2009) 31002 [`hep-th/0608013`].

[456] H. Steinacker, *Emergent Gravity from Noncommutative Gauge Theory, JHEP* **12** (2007) 049 [`0708.2426`].

[457] H. Steinacker, *Emergent Gravity and Noncommutative Branes from Yang-Mills Matrix Models, Nucl. Phys.* **B810** (2009) 1 [`0806.2032`].

[458] H. Steinacker, *Covariant Field Equations, Gauge Fields and Conservation Laws from Yang-Mills Matrix Models, JHEP* **02** (2009) 044 [`0812.3761`].

[459] H. Grosse, H. Steinacker and M. Wohlgenannt, *Emergent Gravity, Matrix Models and UV/IR Mixing, JHEP* **04** (2008) 023 [`0802.0973`].

[460] D. Klammer and H. Steinacker, *Fermions and Emergent Noncommutative Gravity, JHEP* **08** (2008) 074 [`0805.1157`].

[461] H. Steinacker, *On the Newtonian limit of emergent NC gravity and long-distance corrections, JHEP* **12** (2009) 024 [`0909.4621`].

[462] H. Steinacker, *Emergent Geometry and Gravity from Matrix Models: an Introduction, Class. Quant. Grav.* **27** (2010) 133001 [`1003.4134`].

[463] B. S. DeWitt, *The Quantization of geometry*, in *Gravitation: An Introduction to Current Research*, (New York), pp. 266–381, Wiley, (1962).

[464] S. Minwalla, M. Van Raamsdonk and N. Seiberg, *Noncommutative perturbative dynamics, JHEP* **02** (2000) 020 [`hep-th/9912072`].

[465] J. Collins, A. Perez, D. Sudarsky, L. Urrutia and H. Vucetich, *Lorentz invariance and quantum gravity: an additional fine-tuning problem?, Phys. Rev. Lett.* **93** (2004) 191301 [`gr-qc/0403053`].

[466] T. Filk, *Divergencies in a field theory on quantum space, Phys. Lett.* **B376** (1996) 53.

[467] R. J. Szabo, *Quantum field theory on noncommutative spaces, Phys. Rept.* **378** (2003) 207 [`hep-th/0109162`].

345

[468] M. R. Douglas and N. A. Nekrasov, *Noncommutative field theory, Rev. Mod. Phys.* **73** (2001) 977 [`hep-th/0106048`].

[469] G. Amelino-Camelia, G. Mandanici and K. Yoshida, *On the IR / UV mixing and experimental limits on the parameters of canonical noncommutative space-times, JHEP* **01** (2004) 037 [`hep-th/0209254`].

[470] V. V. Khoze and J. Levell, *Noncommutative standard modelling, JHEP* **09** (2004) 019 [`hep-th/0406178`].

[471] E. T. Tomboulis, *Nonlocal and quasilocal field theories, Phys. Rev.* **D92** (2015) 125037 [`1507.00981`].

[472] J. Gomis and T. Mehen, *Space-time noncommutative field theories and unitarity, Nucl. Phys.* **B591** (2000) 265 [`hep-th/0005129`].

[473] A. Bassetto, L. Griguolo, G. Nardelli and F. Vian, *On the unitarity of quantum gauge theories on noncommutative spaces, JHEP* **07** (2001) 008 [`hep-th/0105257`].

[474] O. Aharony, J. Gomis and T. Mehen, *On theories with lightlike noncommutativity, JHEP* **09** (2000) 023 [`hep-th/0006236`].

[475] M. M. Sheikh-Jabbari and A. Tureanu, *Light-Like Noncommutativity, Light-Front Quantization and New Light on UV/IR Mixing, Phys. Lett.* **B697** (2011) 63 [`1010.0317`].

[476] L. Alvarez-Gaume, J. L. F. Barbon and R. Zwicky, *Remarks on time space noncommutative field theories, JHEP* **05** (2001) 057 [`hep-th/0103069`].

[477] N. Seiberg, L. Susskind and N. Toumbas, *Strings in background electric field, space / time noncommutativity and a new noncritical string theory, JHEP* **06** (2000) 021 [`hep-th/0005040`].

[478] R. Gopakumar, J. M. Maldacena, S. Minwalla and A. Strominger, *S duality and noncommutative gauge theory, JHEP* **06** (2000) 036 [`hep-th/0005048`].

[479] A. Torrielli, *Cutting rules and perturbative unitarity of noncommutative electric type field theories from string theory, Phys. Rev.* **D67** (2003) 086010 [`hep-th/0207148`].

[480] M. M. Sheikh-Jabbari, *Open strings in a B field background as electric dipoles, Phys. Lett.* **B455** (1999) 129 [`hep-th/9901080`].

[481] D. Bigatti and L. Susskind, *Magnetic fields, branes and noncommutative geometry, Phys. Rev.* **D62** (2000) 066004 [`hep-th/9908056`].

[482] N. Seiberg, L. Susskind and N. Toumbas, *Space-time noncommutativity and causality*, *JHEP* **06** (2000) 044 [hep-th/0005015].

[483] H. O. Girotti, M. Gomes, A. Yu. Petrov, V. O. Rivelles and A. J. da Silva, *The Low-energy limit of the noncommutative Wess-Zumino model*, *JHEP* **05** (2002) 040 [hep-th/0101159].

[484] C. Acatrinei, *Canonical quantization of noncommutative field theory*, *Phys. Rev.* **D67** (2003) 045020 [hep-th/0204197].

[485] G. Arcioni, J. L. F. Barbon, J. Gomis and M. A. Vazquez-Mozo, *On the stringy nature of winding modes in noncommutative thermal field theories*, *JHEP* **06** (2000) 038 [hep-th/0004080].

[486] A. Rajaraman and M. Rozali, *Noncommutative gauge theory, divergences and closed strings*, *JHEP* **04** (2000) 033 [hep-th/0003227].

[487] W. Fischler, E. Gorbatov, A. Kashani-Poor, S. Paban, P. Pouliot and J. Gomis, *Evidence for winding states in noncommutative quantum field theory*, *JHEP* **05** (2000) 024 [hep-th/0002067].

[488] M. Van Raamsdonk and N. Seiberg, *Comments on noncommutative perturbative dynamics*, *JHEP* **03** (2000) 035 [hep-th/0002186].

[489] Y. Kiem and S. Lee, *UV / IR mixing in noncommutative field theory via open string loops*, *Nucl. Phys.* **B586** (2000) 303 [hep-th/0003145].

[490] A. Armoni and E. Lopez, *UV / IR mixing via closed strings and tachyonic instabilities*, *Nucl. Phys.* **B632** (2002) 240 [hep-th/0110113].

[491] A. Armoni, E. Lopez and A. M. Uranga, *Closed strings tachyons and noncommutative instabilities*, *JHEP* **02** (2003) 020 [hep-th/0301099].

[492] E. Lopez, *From UV / IR mixing to closed strings*, *JHEP* **09** (2003) 033 [hep-th/0307196].

[493] J. Gomis, K. Kamimura and J. Llosa, *Hamiltonian formalism for space-time noncommutative theories*, *Phys. Rev.* **D63** (2001) 045003 [hep-th/0006235].

[494] D. Bahns, S. Doplicher, K. Fredenhagen and G. Piacitelli, *On the Unitarity problem in space-time noncommutative theories*, *Phys. Lett.* **B533** (2002) 178 [hep-th/0201222].

[495] H. Bozkaya, P. Fischer, H. Grosse, M. Pitschmann, V. Putz, M. Schweda et al., *Space-time noncommutative field theories and causality*, *Eur. Phys. J.* **C29** (2003) 133 [hep-th/0209253].

[496] Y. Liao and K. Sibold, *Time-ordered perturbation theory on non-commutative spacetime ii: Unitarity, The European Physical Journal C - Particles and Fields* **25** (2002) 479.

[497] C.-h. Rim and J. H. Yee, *Unitarity in space-time noncommutative field theories, Phys. Lett.* **B574** (2003) 111 [hep-th/0205193].

[498] S. Denk and M. Schweda, *Time ordered perturbation theory for nonlocal interactions: Applications to NCQFT, JHEP* **09** (2003) 032 [hep-th/0306101].

[499] P. Fischer and V. Putz, *No UV / IR mixing in unitary space-time noncommutative field theory, Eur. Phys. J.* **C32** (2004) 269 [hep-th/0306099].

[500] Y. Liao, *On evaluation of nonplanar diagrams in noncommutative field theory, Nucl. Phys.* **B713** (2005) 235 [hep-th/0408047].

[501] N. Barnaby and N. Kamran, *Dynamics with infinitely many derivatives: The Initial value problem, JHEP* **02** (2008) 008 [0709.3968].

[502] T. Salminen and A. Tureanu, *Noncommutative Time in Quantum Field Theory, Phys. Rev.* **D84** (2011) 025009 [1101.4798].

[503] T. Biswas and N. Okada, *Towards LHC physics with nonlocal Standard Model, Nucl. Phys.* **B898** (2015) 113 [1407.3331].

[504] A. Addazi and G. Esposito, *Nonlocal quantum field theory without acausality and nonunitarity at quantum level: is SUSY the key?, Int. J. Mod. Phys.* **A30** (2015) 1550103 [1502.01471].

[505] P. Chin and E. T. Tomboulis, *Nonlocal vertices and analyticity: Landau equations and general Cutkosky rule, JHEP* **06** (2018) 014 [1803.08899].

[506] I. Ya. Aref'eva, D. M. Belov and A. S. Koshelev, *Two loop diagrams in noncommutative phi**4(4) theory, Phys. Lett.* **B476** (2000) 431 [hep-th/9912075].

[507] I. Chepelev and R. Roiban, *Renormalization of quantum field theories on noncommutative R**d. 1. Scalars, JHEP* **05** (2000) 037 [hep-th/9911098].

[508] M. Kontsevich, *Deformation quantization of Poisson manifolds. 1., Lett. Math. Phys.* **66** (2003) 157 [q-alg/9709040].

[509] E. Langmann and R. J. Szabo, *Duality in scalar field theory on noncommutative phase spaces, Phys. Lett.* **B533** (2002) 168 [hep-th/0202039].

[510] H. Grosse and R. Wulkenhaar, *Renormalization of phi**4 theory on noncommutative R**4 in the matrix base, Commun. Math. Phys.* **256** (2005) 305 [hep-th/0401128].

[511] H. Grosse and R. Wulkenhaar, *Self-Dual Noncommutative ϕ^4 -Theory in Four Dimensions is a Non-Perturbatively Solvable and Non-Trivial Quantum Field Theory, Commun. Math. Phys.* **329** (2014) 1069 [`1205.0465`].

[512] R. Gurau, J. Magnen, V. Rivasseau and A. Tanasa, *A Translation-invariant renormalizable non-commutative scalar model, Commun. Math. Phys.* **287** (2009) 275 [`0802.0791`].

[513] H. Grosse and R. Wulkenhaar, *A Solvable Four-Dimensional QFT*, in *Proceedings, Conference on Quantum Mathematical Physics: A Bridge between Mathematics and Physics: Regensburg, Germany, September 29-October 2, 2014*, pp. 137–161, 2016, DOI.

[514] B. Ydri, *Lectures on Matrix Field Theory, Lect. Notes Phys.* **929** (2017) pp.1 [`1603.00924`].

[515] T. D. Lee and G. C. Wick, *Negative Metric and the Unitarity of the S Matrix, Nucl. Phys.* **B9** (1969) 209.

[516] B. Grinstein, D. O'Connell and M. B. Wise, *The Lee-Wick standard model, Phys. Rev.* **D77** (2008) 025012 [`0704.1845`].

[517] M. Van Raamsdonk, *The Meaning of infrared singularities in noncommutative gauge theories, JHEP* **11** (2001) 006 [`hep-th/0110093`].

[518] H. Huffel, *Comment on the decoupling of UV and IR divergencies within dimensional regularization in noncommutative theories*, `hep-th/0210028`.

[519] F. Ruiz Ruiz, *UV / IR mixing and the Goldstone theorem in noncommutative field theory, Nucl. Phys.* **B637** (2002) 143 [`hep-th/0202011`].

[520] I. Ya. Aref'eva, D. M. Belov, A. S. Koshelev and O. A. Rytchkov, *UV / IR mixing for noncommutative complex scalar field theory. 2. (Interaction with gauge fields), Nucl. Phys. Proc. Suppl.* **102** (2001) 11 [`hep-th/0003176`].

[521] A. Armoni, *Comments on perturbative dynamics of noncommutative Yang-Mills theory, Nucl. Phys.* **B593** (2001) 229 [`hep-th/0005208`].

[522] C. P. Martin and F. Ruiz Ruiz, *Paramagnetic dominance, the sign of the beta function and UV / IR mixing in noncommutative U(1), Nucl. Phys.* **B597** (2001) 197 [`hep-th/0007131`].

[523] M. Chaichian, P. Presnajder, M. M. Sheikh-Jabbari and A. Tureanu, *Noncommutative standard model: Model building, Eur. Phys. J.* **C29** (2003) 413 [`hep-th/0107055`].

[524] M. Chaichian, P. Presnajder, M. M. Sheikh-Jabbari and A. Tureanu, *Noncommutative gauge field theories: A No go theorem*, Phys. Lett. **B526** (2002) 132 [hep-th/0107037].

[525] M. Chaichian, A. Kobakhidze and A. Tureanu, *Spontaneous reduction of noncommutative gauge symmetry and model building*, Eur. Phys. J. **C47** (2006) 241 [hep-th/0408065].

[526] M. Arai, S. Saxell, A. Tureanu and N. Uekusa, *Circumventing the No-Go Theorem in Noncommutative Gauge Field Theory*, Phys. Lett. **B661** (2008) 210 [0710.3513].

[527] R. Oeckl, *Untwisting noncommutative R**d and the equivalence of quantum field theories*, Nucl. Phys. **B581** (2000) 559 [hep-th/0003018].

[528] J. Wess, *Deformed coordinate spaces: Derivatives*, in *Proceedings, 2nd Southeastern European Workshop on Mathematical, theoretical and phenomenological challenges beyond the standard model*, pp. 122–128, 2003, DOI [hep-th/0408080].

[529] M. Chaichian, P. Presnajder and A. Tureanu, *New concept of relativistic invariance in NC space-time: Twisted Poincare symmetry and its implications*, Phys. Rev. Lett. **94** (2005) 151602 [hep-th/0409096].

[530] M. Chaichian, P. P. Kulish, K. Nishijima and A. Tureanu, *On a Lorentz-invariant interpretation of noncommutative space-time and its implications on noncommutative QFT*, Phys. Lett. **B604** (2004) 98 [hep-th/0408069].

[531] M. Hayakawa, *Perturbative analysis on infrared and ultraviolet aspects of noncommutative QED on R**4*, hep-th/9912167.

[532] A. Matusis, L. Susskind and N. Toumbas, *The IR / UV connection in the noncommutative gauge theories*, JHEP **12** (2000) 002 [hep-th/0002075].

[533] A. Anisimov, T. Banks, M. Dine and M. Graesser, *Comments on noncommutative phenomenology*, Phys. Rev. **D65** (2002) 085032 [hep-ph/0106356].

[534] R. Bufalo and M. Ghasemkhani, *Three-dimensional noncommutative Yukawa theory: Induced effective action and propagating modes*, Int. J. Mod. Phys. **A32** (2017) 1750019 [1608.02606].

[535] K. Bouchachia, S. Kouadik, M. Hachemane and M. Schweda, *One loop radiative corrections to the translation-invariant noncommutative Yukawa Theory*, J. Phys. **A48** (2015) 365401 [1502.02992].

[536] M. Chaichian, K. Nishijima and A. Tureanu, *Spin statistics and CPT theorems in noncommutative field theory*, Phys. Lett. **B568** (2003) 146 [hep-th/0209008].

[537] D. H. T. Franco and C. M. M. Polito, *A new derivation of the CPT and spin-statistics theorems in non-commutative field theories*, J. Math. Phys. **46** (2005) 083503 [hep-th/0403028].

[538] L. Alvarez-Gaume and M. A. Vazquez-Mozo, *General properties of noncommutative field theories*, Nucl. Phys. **B668** (2003) 293 [hep-th/0305093].

[539] M. A. Soloviev, *Axiomatic formulations of nonlocal and noncommutative field theories*, Theor. Math. Phys. **147** (2006) 660 [hep-th/0605249].

[540] M. M. Sheikh-Jabbari, *C, P, and T invariance of noncommutative gauge theories*, Phys. Rev. Lett. **84** (2000) 5265 [hep-th/0001167].

[541] E. Witten, *APS Medal for Exceptional Achievement in Research: Invited article on entanglement properties of quantum field theory*, Rev. Mod. Phys. **90** (2018) 045003 [1803.04993].

[542] S. S. Gubser and S. L. Sondhi, *Phase structure of noncommutative scalar field theories*, Nucl. Phys. **B605** (2001) 395 [hep-th/0006119].

[543] H. Steinacker, *A Non-perturbative approach to non-commutative scalar field theory*, JHEP **03** (2005) 075 [hep-th/0501174].

[544] G.-H. Chen and Y.-S. Wu, *Renormalization group equations and the Lifshitz point in noncommutative Landau-Ginsburg theory*, Nucl. Phys. **B622** (2002) 189 [hep-th/0110134].

[545] P. Castorina and D. Zappala, *Nonuniform symmetry breaking in noncommutative lambda phi**4 theory*, Phys. Rev. **D68** (2003) 065008 [hep-th/0303030].

[546] H. O. Girotti, M. Gomes, V. O. Rivelles and A. J. da Silva, *A Consistent noncommutative field theory: The Wess-Zumino model*, Nucl. Phys. **B587** (2000) 299 [hep-th/0005272].

[547] A. Armoni, *A Note on noncommutative orbifold field theories*, JHEP **03** (2000) 033 [hep-th/9910031].

[548] M. Maggiore, *A Generalized uncertainty principle in quantum gravity*, Phys. Lett. **B304** (1993) 65 [hep-th/9301067].

[549] A. N. Tawfik and A. M. Diab, *Review on Generalized Uncertainty Principle*, Rept. Prog. Phys. **78** (2015) 126001 [1509.02436].

[550] K. Konishi, G. Paffuti and P. Provero, *Minimum Physical Length and the Generalized Uncertainty Principle in String Theory*, Phys. Lett. **B234** (1990) 276.

[551] T. Yoneya, *String theory and space-time uncertainty principle, Prog. Theor. Phys.* **103** (2000) 1081 [`hep-th/0004074`].

[552] S. Hossenfelder, *Minimal Length Scale Scenarios for Quantum Gravity, Living Rev. Rel.* **16** (2013) 2 [`1203.6191`].

[553] J. Polchinski, *Comment on [arXiv:1106.1417] 'Small Lorentz violations in quantum gravity: do they lead to unacceptably large effects?', Class. Quant. Grav.* **29** (2012) 088001 [`1106.6346`].

[554] S. Galluccio, F. Lizzi and P. Vitale, *Translation Invariance, Commutation Relations and Ultraviolet/Infrared Mixing, JHEP* **09** (2009) 054 [`0907.3640`].

[555] F. Ardalan and N. Sadooghi, *Translational-invariant noncommutative gauge theory, Phys. Rev.* **D83** (2011) 025014 [`1008.5064`].

[556] S. Doplicher, K. Fredenhagen and J. E. Roberts, *The Quantum structure of space-time at the Planck scale and quantum fields, Commun. Math. Phys.* **172** (1995) 187 [`hep-th/0303037`].

[557] H. Kase, K. Morita, Y. Okumura and E. Umezawa, *Lorentz invariant noncommutative space-time based on DFR algebra, Prog. Theor. Phys.* **109** (2003) 663 [`hep-th/0212176`].

[558] C. E. Carlson, C. D. Carone and N. Zobin, *Noncommutative gauge theory without Lorentz violation, Phys. Rev.* **D66** (2002) 075001 [`hep-th/0206035`].

[559] J. Heckman and H. Verlinde, *Covariant non-commutative space–time, Nucl. Phys.* **B894** (2015) 58 [`1401.1810`].

[560] A. Much and J. D. Vergara, *A Poincaré covariant noncommutative spacetime, Int. J. Geom. Meth. Mod. Phys.* **15** (2018) 1850159 [`1704.07932`].

[561] R. J. Eden, P. V. Landshoff, D. I. Olive and J. C. Polkinghorne, *The analytic S-matrix.* Cambridge Univ. Press, Cambridge, 1966.

[562] C. Itzykson and J. B. Zuber, *Quantum Field Theory,* International Series In Pure and Applied Physics. McGraw-Hill, New York, 1980.

[563] J. Achenbach, *Scientists are baffled: What's up with the universe?, Washington Post* (2019) .

[564] Z. Berezhiani, *Through the looking-glass: Alice's adventures in mirror world,* `hep-ph/0508233`.

[565] A. Falkowski, S. Pokorski and M. Schmaltz, *Twin SUSY, Phys. Rev.* **D74** (2006) 035003 [`hep-ph/0604066`].

[566] A. Katz, A. Mariotti, S. Pokorski, D. Redigolo and R. Ziegler, *SUSY Meets Her Twin*, *JHEP* **01** (2017) 142 [1611.08615].

[567] M. Badziak and K. Harigaya, *Supersymmetric D-term Twin Higgs*, *JHEP* **06** (2017) 065 [1703.02122].