High Energy Physics: from Quantum Field Theory to Higgs Coupling Measurement

by

Evan Ranken

Submitted in Partial Fulfillment of the
Requirements for the Degree
Doctor of Philosophy

Supervised by
Professor Regina Demina

Department of Physics and Astronomy
Arts, Sciences and Engineering

University of Rochester
Rochester, New York
2020

"For Mom, Dad, and Lana"

Table of Contents

Biographical Sketch

The author grew up in the town of Bernalillo, New Mexico, in the United States. He obtained his Bachelors degree at Colorado College in 2012, majoring in Physics & Mathematics. In 2013, he enrolled in the PhD program at the University of Rochester, supported by a Sproull Fellowship from the University. From 2014 until 2017, he worked under the guidance of Professor S.G. Rajeev, studying a variety of research topics related to quantum field theory in theoretical and mathematical physics. In the spring of 2014 he received the department's AAPT Teaching Prize for his work as a teaching assistant, and went on to serve as instructor for an introductory physics course that summer. In 2015 he was awarded a fellowship under the NSF Graduate Research Fellowship Program (GRFP).

After inadvertently attending a conference filled with well-funded experimentalists in late 2016, the author was swayed to the dark side and decided to join an experiment. Following a trip to co-instruct a course at Colorado College in spring 2017, he began working in earnest with his second advisor, Professor Regina Demina. in collaboration with the CMS experiment. The author then relocated to work on-site at CERN, where he finished his final doctoral research project measuring the interaction between the top quark and Higgs boson.

Publications during doctoral study:

- "Measurement of the top quark Yukawa coupling using kinematics of $\mathrm{t\bar{t}}$ dilepton decays at $\sqrt{s}$= 13 TeV " (forthcoming)

  with Regina Demina and the CMS collaboration
  technical note CMS-PAS-TOP-19-008 available at cds.cern.ch
  partially reprinted as Appendix A

- "Highly nonlinear wave solutions in a dual to the chiral model" (2016)

  with S.G. Rajeev, Phys. Rev. D **93** 105016 [arXiv:1603.08256]
  partially reprinted as Appendix B

- "Continuity of scalar fields with logarithmic correlations" (2015)

  with S.G. Rajeev, Phys. Rev. D **92** 045021 [arXiv:1508.04655]
  partially reprinted as Appendix C

Acknowledgements

I would like to thank all of the teachers and professors who inspired and helped me, starting during my youth in New Mexico. To name a few: Stuart Lipkowitz, Danny Packer, Dr. Vraspir, Dr. Buchanan, Dr. Julian, Mr. Montaño... and so many more! Thanks as well to everyone in the Physics & Mathematics (& Computer Science) departments at Colorado College circa 2012, with extra special thanks to Professors Stefan Erickson, Shane Burns, Stephanie Dicenzo, Marlow Anderson, Mike Siddoway, David Brown, Phil Cervantes, Dick Hilt, and of course to Marita Beckert! At the University of Rochester, thanks to Laura Blumkin, without whom the department would fall apart. Thanks, of course, to those who advised me during my time at Rochester: to Professor S.G. Rajeev for leading me on a wonderful journey through mathematical physics, and to Professor Regina Demina for offering to take me into a different world of physics research. Thanks as well to Aran Garcia-Bellido for all his help and detailed reviews of our paper with CMS. Thanks as well to the best housemates ever, Joe Murphree and Amanda Davis, and to everyone else who helped with the self-paced curriculum. Thanks to Joseph Dulemba for triple-checking this document, and to Rhys Taus for his valuable help exploring the Galar region.

Thank you to the many friends from outside of physics who supported me while I worked towards this dissertation. Special thanks to Kaila Ryan for uncountably many things–without your friendship I would have found myself very alone in 2019. Thanks as well as John Derksen, Griffin Gorsky, Aku Antikainen and Meghan Dorn, all of whom helped me greatly as I tried to maneuver the American medical system. Thanks to everyone else who was supportive during that stressful time.

Thanks to Kyle Leggott and Robbie Martin for visiting me in France and for taking me on a pleasant trip to the Sword Coast. Thanks to Abigail Kinem for listening to me ramble. Thanks to Colin Martz for the music recommendations. And thanks to everyone else who put up with me over the years, including everyone who I studied alongside at UR and Colorado College. By now I have fallen out of touch with many of those whose academic paths briefly crossed my own, but I know I wouldn't have finished this degree if I hadn't met so many great people along the way.

And lastly but most importantly, thanks to my parents Bill & Bev Ranken, who have supported me through every smart and stupid thing I have ever done. Love you!

Abstract

This document summarizes work in theoretical and experimental high-energy physics completed as part of the author's doctoral studies, while providing context and motivation for these projects. Beginning with an overview of the relevant backgrounds, the motivation for this doctoral research is explored by discussing the current state of particle physics and many unanswered questions in this field. Emphasis is placed on providing broad context to readers outside of the particle physics community. Theoretical work from early PhD studies is briefly summarized, exploring the continuity of scalar quantum fields and studying a novel toy field theory with intriguing nonlinear wave solutions. The discussion then turns to the experimental side of particle physics, covering relevant conceptual material and experimental techniques. The author's research in collaboration with the CMS experiment is presented, detailing the motivation and methodology. This project produces a measurement of the strength of interaction between the top quark and Higgs boson, obtaining a Yukawa coupling of $Y_t = 1.16^{+0.24}_{-0.35}$ relative to the standard model value.

Contributors and Funding Sources

This work was supervised by a dissertation committee consisting of Professors Regina Demina (advisor), S.G. Rajeev, and Arie Bodek of the Physics & Astronomy Department, as well as Professor Alan Grossfield of the Department of Biochemistry and Biophysics.

All theoretical research from 2014-2016 was done in close collaboration with Professor S.G. Rajeev. Experimental research was performed from 2017-2020 with the CMS collaboration under the guidance of Professor Regina Demina. Research with CMS used a reconstruction framework developed by Otto Hindrichs, who also provided useful guidance and tools to assess uncertainties and model detector effects within that framework. All other work for the dissertation was completed by the author independently.

This work was made possible by a wide variety of software, including: Mathematica, The CERN ROOT framework with RooFit and RootStat; as well as free and open source software like the GNU compiler collection, the Linux OS family, and the Python programming language. Among these, the CERN ROOT framework caused the author more mental anguish than the others.

Specific particle physics simulation tools are cited individually.

Experimental work also relied on computing resources located at Fermi National Accelerator Laboratory (Fermilab), and the European Center for Nuclear Research (CERN).

Graduate study was supported by the U.S. Department of Energy Office of Science, under Award Number DE-SC0008475 as well as the NSF Graduate Research Fellowship Program under Grant No. DGE-141911, and by teaching assistantships and the Robert L. and Mary L. Sproull University Fellowship at the University of Rochester.

List of Tables

List of Figures

List of Symbols and Abbreviations

Symbols and abbreviations are defined as they appear in the text, and a few symbols are used differently in multiple contexts as we survey different fields and eras. We collect some of the most commonly used notation here, for convenience.

E	Energy
x	position (usually as a Lorentz 4-vector)
∂_μ	$\frac{\partial}{\partial x^\mu}$
p	momentum
m_f	mass of fermion f
g_f	Yukawa coupling of fermion f
α	electromagnetic coupling
α_S	strong interaction coupling
ϕ	scalar field
ψ	fermionic field
QFT	Quantum Field Theory
QED	Quantum Electrodynamics
QCD	Quantum Chromodynamics
EW	Electroweak
LHC	Large Hadron Collider
CI	Confidence Interval

Chapter 1

Introduction

What is this dissertation about, and why should you care?

Those are both good questions. The goal of this document is to take the author's graduate research, which often addresses hyper-specific questions encased in jargon, and present it in the broader context of current physics knowledge. The research in question involves both experimental and theoretical projects in an area frequently labeled *high energy physics*, a subset of particle physics. In a sense, this branch of physics aims to answer fundamental questions involving the smallest known pieces of the universe. The essential nature of these fundamental particles becomes increasing evident at smaller distances and higher energies–which is why the theoretical scope of high energy physics is quite broadly defined. On the experimental side, the label is more commonly reserved for experiments at particle colliders, such as the Large Hadron Collider (LHC) near Geneva, Switzerland.

At its heart, all the research involved in this dissertation shares the common goal of better understanding these fundamental particles and the rules that govern them. To contextualize these studies, we will try to summarize the current state of the field and explain the lingering questions that these research projects seek to answer. We are guided by a rather bombastic question: What is the universe made of at the most fundamental level? The author hopes that anyone with an interest in this question will find something of value in this document, even if the research involved reveals only a hint of a tiny piece of an answer.

1.1 What's in This Document

We begin in Chapter 2 with an overview of modern particle physics, followed by a discussion of open problems in the field in Chapter 3. The content of these chapters will already be familiar to anyone with experience in particle physics, so they are written with a non-expert in mind as much as possible, while still covering the necessary material. We then proceed to discuss theoretical avenues of research in Chapter 4, including an overview of theoretical research completed by the author in the early years of doctoral study, rooted in the detailed mathematical study of quantum fields. Next we set our sights on experiment, giving a brief overview of experimental avenues in modern particle physics in Chapter 5. We focus in on the role of top quarks in particle experiments in Chapter 6, before introducing specifically the LHC apparatus and CMS experiment in Chapter 7. Lastly, we present a summary of the author's final doctoral research project, a measurement of the top quark Yukawa coupling, in Chapter 8.

Effort has been made to keep the majority of this thesis as accessible as possible to a wide scientific audience, to the best of the author's ability. A majority of technical details are confined to Chapters 4 and 8, though other sections may delve briefly into more precise language to offer clarity to the expert reader while offering references to the non-expert. It is the author's hope that the core concepts behind the research motivation and conclusions are presented as plainly as possible, such that it is not necessary to understand all technical details in order to appreciate the bigger picture.

Chapter 2

A Primer on Particle Physics

"From time immemorial, man has desired to comprehend the complexity of nature in terms of as few elementary concepts as possible."

— Abdus Salam

2.1 The Big Picture

It is often said that fundamental physics has two pieces which are at conflict with one another: quantum physics and general relativity. The former describes the smallest building blocks of the universe, while the latter explains its behavior at the largest scales. This thesis is concerned with the former, but it is useful to have a broad idea of how the two contrast with one another.

General relativity provides a unified framework for explaining gravitational interaction at large scales. It superseded Newton's theory of gravity, which already did a fine job at explaining most planetary and lunar orbits, along with earthly phenomena such as projectile motion and blocks sliding down frictionless ramps. Meanwhile, general relativity gave us the big bang theory, explaining much about the universe's origins and its behavior on large scales. It even predicted the possibility of extreme objects such as black holes, which have now been observed through a variety of means.

But general relativity is a theory which operates at almost incomprehensibly different scales from those of particle physics. And rather than provide a mechanism to explain gravity on small scales, particle physics ignores it completely. Our current understanding of particle physics fails to provide a small-scale origin for gravitational force, and breaks down upon attempt to forcefully include gravity. This discrepancy persists in part because it remains essentially impossible to probe the small scale nature of gravitational interaction in a lab setting. Other forces typically render gravity obsolete at small scales. Only a few extreme cases directly require a simultaneous description of quantum physics and gravity, such as the centers of black holes or the very moment of the big bang. These are of course inaccessible to direct measurement, and the energies involved are almost unfathomable.

Since these extreme scenarios are so incredibly far away from our lab capabilities, one might ask why particle physicists bother to build large machines which operate at a fraction of the energy scale. Luckily, this fundamental contradiction is not the only mystery remaining in particle physics. While General relativity provides a streamlined and unified framework for all large scale interactions in our universe, our model of particle physics is much messier and ad-hoc. It almost begs to be studied further, presenting such apparent disorganization that one immediately suspects something to be amiss. And so, within particle physics, there are a number of avenues that could point to breakthroughs in our understanding of fundamental physics. To understand these avenues of research, we present a summary of our current best understanding of particle physics, starting with its predecessors.

2.2 When Things Got Weird: Quantum Mechanics

In the early 20th century, the development of quantum mechanics radically shook humanity's understanding of the behavior and composition of matter at small scales. Evidence-based models for atoms had gained prevalence in the century prior, but patterns in their structure and behavior remained mysterious. The turn of the century brought the discovery of electrons and protons, which provided more fundamental and universal pieces of matter. The idea of all matter consisting of very small discrete pieces was not new, but the strange behavior of the pieces came as a surprise.

Before quantum mechanics, tiny pieces of matter were thought of as *particles* which behaved like tiny billiard balls, while seemingly separate phenomena like light (electromagnetic radiation) were modeled strictly as *waves*. Einstein's study of the photo-electric effect [1] showed that in fact light behaves at times like a particle, and transmitting energy not continuously but rather in discrete pieces or *quanta*. This was the first hint that light itself is made up of many small particles, now called photons. On the other hand, the famous double slit experiment (though not demonstrated until the 1961) showed that matter particles like electrons could in fact behave like waves to produce interference pattern. [2,3]

In fact, at the smallest scales, matter and light were comprised neither of particles nor waves in the usual sense. Instead, small-scale physics could only accurately be modeled by strange objects with a mixture of particle-like and wave-like properties. This was the dawn of quantum physics. The study of this duality and its implications has proceeded alongside the development of more advanced models of particle physics, and continues to this day. In terms of casual description, the term *particle* stuck, probably because it conveys their quantized nature more intuitively. And so today, we build particle colliders–though a pedantic individual could have dubbed them wave colliders instead.

In quantum mechanics, particles like electrons, protons, or photons are the fundamental objects. By including the known "shape" of the electrostatic force F, which decays over a distance r as $F \propto \frac{1}{r^2}$, this theory was able to model what happens when you take trap a collection of electrons in "orbit" around a positively charged nucleus made of protons and neutrons. Such a collection of fundamental objects stuck together by interaction is called a *bound state*. As a novel consequence of the theory, electrons orbiting a nucleus were restricted to certain discrete values of orbital energy and angular momentum. In describing these energy levels, the theory successfully predicted the structure of the periodic table, and many properties of the elements therein. For example, this could be used to explain the energies of emitted photons from many elements, though it did not explain the physical law that allowed a photon to be emitted in the first place. And this was a hint of what the theory lacked–it knew that particles could sometimes be created seemingly out of nothing, as long as energy was conserved. But why this was allowed remained somewhat mysterious.

In a sense, breaking down quantities like energy into small, finite quanta is at the heart of quantum mechanics. Of course, the full story is quite a bit more complicated, and the "true meaning" of consequences is still fiercely debated (especially by physicists who have run out of more precise research problems to solve). Quantum physics is inherently probabilistic, a fact which will carry over to modern particle physics. Second, and instrumental this probabilistic doctrine, is the importance of operators.

In quantum mechanics, observable quantities are described by mathematical *operators*. All particles exist in some quantum state, often written as $|\psi\rangle$. We don't think of of observables like position x or momentum p as being simple properties belonging to this particle; rather, they are promoted to operators denoted $\hat{x}$ and $\hat{p}$ which *act* on the particle's quantum state. Probabilistic calculations can be be performed, such as the expectation value of an observable like x or x^2 for a particle with state $|\psi\rangle$. These would be written as $\langle\psi|\,\hat{x}\,|\psi\rangle$, and $\langle\psi|\,\hat{x}^2\,|\psi\rangle$, respectively. Because the particle represented by $|\psi\rangle$ has a distribution of possible measured x-positions, note that $\langle\psi|\,\hat{x}\,|\psi\rangle^2 \neq \langle\psi|\,\hat{x}^2\,|\psi\rangle$. Rather, the two related to calculating the mean and variance of a statistical distribution. We mention these calculations to compare with field theory later on.

2.3 The Leap to Fields

Quantum field theory (QFT) is often described as the marriage of quantum mechanics with Einstein's theory of special relativity. Although many of the new ideas presented by QFT are a consequence of joining these two theories, it is easy to understand them without discussing special relativity much at all.

A particle is typically thought of as an object that exists at a specific point in space at a given time, while a field is something that has a value at each point in space simultaneously at a given time. The *values* at each point can be scalars, vectors, or more exotic objects. Classical field theory was often used to discuss gravitational and electromagnetic fields, which carry forces felt by classical objects and even by particles in the context of quantum mechanics. The simplest type of field to imagine is a *scalar field* $\phi : \mathbb{R}^d \mapsto \mathbb{R}$. The object ϕ assigns a real number to each point in d-dimensional space. Simple attempts to place a quantum scalar field into relativistic spacetime yield the Klein Gordon equation,

$$\frac{\partial^2}{\partial t^2}\phi(x) - \nabla^2\phi(x) + m^2\phi(x) = 0. \tag{2.1}$$

Where $\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$. Here we are writing in so-called natural units, where the speed of light $c = 1$, to remove excess coupling constants. Constants like mass the m remain, as well as other field coupling constants governing interactions, as we shall see.

To shorten things, in field theory we move to relativistic notation. Here x will represent the spacetime coordinate vector $x = (x^0, x^1, x^2, x^3)$, identifying x^0 with the dimension of time t. Greek letter indices such as x^μ are used to pick out the μ-th component of x. We use the shorthand $\partial_\mu = \frac{\partial}{\partial x^\mu}$. Indices can appear as both upper and lower indices, and are implicitly summed over when repeated. Lastly, the upper and lower indices are related by contraction with the Minkowski metric tensor η. For any reader not well-versed in the language of relativity, these details are not too important–we will not be performing in depth calculations, but mostly discussing the meaning of equations written with this notation.

In relativistic notation the Klein Gordon equation can be written more compactly as

$$\partial^\mu \partial_\mu \phi + m^2 \phi = 0, \tag{2.2}$$

Where we have not written the explicit dependence $\phi(x)$ as this dependence is assumed for all field objects. (2.2) follows directly from an object called a *Lagrangian density* $\mathscr{L}$, given by

$$\mathscr{L} = \partial^\mu \phi \partial_\mu \phi - m^2 \phi^2. \tag{2.3}$$

In field theory, basic properties of fields can be read off from the Lagrangian density (often just called "the Lagrangian") with minimal effort. This particular Lagrangian has two terms: first the *kinetic term* $\partial^\mu \phi \partial_\mu \phi$, also sometimes called the free field term. Such a term is present for all types of fields, and represents the field's kinetic energy. Unsurprisingly then, the remainder is called the *potential term*, and encodes the field's potential energy. In particular, a term of form $m^2\phi^2$ is called a *mass term*. The mass term, in brief, tells us that the field has some internal energy stored as a mass m. If we *quantize* this field, the meaning of the mass term becomes more interesting.

In QFT, all fields have a ground state or *vacuum state* with the lowest possible energy. The excited states of a field with a mass term are particles with mass m. This encapsulates one of the biggest revolutions brought by QFT: rather than being taken as immutable objects dropped into force-carrying fields, particles themselves are excited states of quantum fields. The number of particles is promoted to quantum operator, through a process known as *second quantization*. In fact, all fields become operator-valued upon quantization, though this is not generally designated explicitly. The details of this procedure are beyond the scope of this dissertation, but the consequence is simple and groundbreaking: particles can be created and destroyed within the theory.

2.4 Gauge Theories and Particle Interaction

One of the earliest successful models in QFT came from the Dirac equation, describing a more complicated type of field which assigns an object called a *spinor* to each point in spacetime. Spinors describe quantum objects with the attribute of *spin*, thought of as an intrinsic angular momentum. The Dirac equation describes a field ψ with spin 1/2, a type of *fermionic* field. In standard relativistic notation, the Dirac equation reads

$$(i\gamma^\mu \partial_\mu - m)\psi = 0, \tag{2.4}$$

where γ^μ are the famous Dirac matrices which make possible the relativistic invariance of a spin-1/2 field. The equation follows from the Lagrangian density

$$\mathscr{L} = i\bar{\psi}\gamma^\mu \partial_\mu \psi - m\bar{\psi}\psi, \tag{2.5}$$

where $\bar{\psi} = \psi^\dagger \gamma^0$ is called the adjoint spinor. Though it looks a bit different than the Lagrangian in (2.3), it has the same fundamental pieces: a kinetic term and a mass term. The Dirac equation was meant to describe electrons, and had a great success in predicting the existence of anti-particles called positrons. However, it is only a starting point for the gauge theories that give us our modern understanding of how particles interact to generate forces. Examining the Lagrangian, we see it has a global symmetry under the transformation

$$\psi \to e^{i\alpha}\psi. \tag{2.6}$$

In the language of group theory, this is a $U(1)$ symmetry–the same mathematical structure as a symmetry under rotation, though the rotation does not occur in physical space. If the field is altered globally throughout spacetime by a phase α, nothing changes. But in relativity, the field cannot communicate instantaneously from one point to another. So, one could ask how the field "knows" to have the same phase at all locations. Though this sounds like philosophical nitpicking, there is a proposed solution with profound consequences. We can *promote* the global symmetry to a local symmetry, one which depends on the spacetime coordinates x. The new problem is, in making α a function of spacetime coordinates $\alpha(x)$, we break the global symmetry. Separating out a constant scaling q for later use, and applying the chain rule, we now have the transformation

$$\psi \quad \to \quad e^{iq\alpha(x)}\psi \tag{2.7}$$

$$\partial_\mu\psi \quad \to \quad e^{iq\alpha(x)}\partial_\mu\psi + iqe^{iq\alpha(x)}\psi\partial_\mu\alpha(x) \tag{2.8}$$

How can we keep the Lagrangian invariant under such a transformation? We introduce a new field A_μ, called a gauge field, which transforms under a local gauge change according to the rule

$$A_\mu \to A_\mu + iq\partial_\mu\alpha(x). \tag{2.9}$$

This field can have the effect of cancelling out the unwanted extra term in (2.8), effectively enforcing local gauge invariance. We can now construct the locally gauge invariant Lagrangian,

$$\mathscr{L}_{\text{QED}} = i\bar{\psi}\gamma^{\mu}\partial_{\mu}\psi - m\bar{\psi}\psi - \bar{\psi}\gamma^{\mu}qA_{\mu}\psi - \frac{1}{4}F^{\mu\nu}F_{\mu\nu}\ , \tag{2.10}$$

Where $F_{\mu\nu} = \partial_{\mu}A_{\nu} - \partial_{\nu}A_{\mu}$ is introduced to define a gauge-invariant kinetic term for the new field. The old fermion field ψ still has its kinetic and mass terms. The new gauge field A_{μ} is a bosonic field, and has a kinetic term but no mass term. Most interestingly, there is a new term containing both fields: an interaction term, modulated by the constant q. Taken as a whole, this Lagrangian describes a massive fermionic field (whose excitations are electrons) interacting with a massless gauge-boson field. The excitations of this gauge field are in fact photons. Our construction of $F^{\mu\nu}$ matches the electromagnetic tensor from classical electrodynamics, and the free gauge term $\frac{1}{4}F^{\mu\nu}F_{\mu\nu}$ is the vacuum term from the Lagrangian formulation of Maxwell's equations.

In essence, the imposition of local gauge invariance yields a Lagrangian for a fully interacting field theory, granting our modern understanding of photons as *gauge bosons* that mediate the electromagnetic force on the quantum level. This particular gauge theory is called *Quantum Electrodynamics*, or QED for short. The Lagrangian can also be written more compactly in terms of a *gauge-covariant* derivative D_{μ}, which absorbs the gauge field interaction term to create derivative operator which itself transforms appropriately under local gauge change:

$$\mathscr{L}_{\text{QED}} = i\bar{\psi}\gamma^{\mu}D_{\mu}\psi - m\bar{\psi}\psi - \frac{1}{4}F^{\mu\nu}F_{\mu\nu}, \tag{2.11}$$

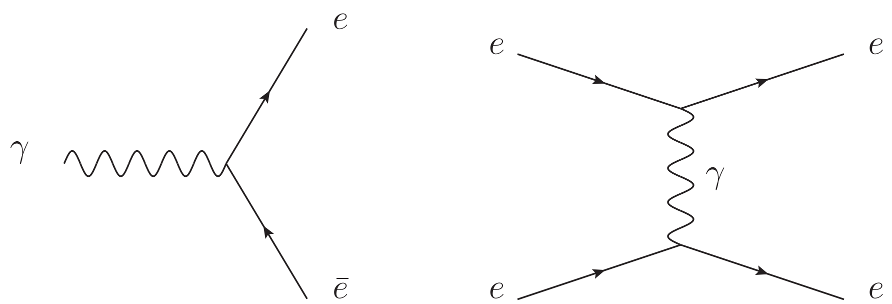

Figure 2.1: Left: Feynman rule for QED diagrams. Feynman diagrams are read with time flowing left to right, with e representing an electron and γ a photon. When the arrow attached to the e line is overall moving leftwards, it is interpreted as a positron $\bar{e}$ as indicated here. Right: a simple process of two electrons interacting via the exchange of a photon.

$$D_\mu = \partial_\mu - iqA_\mu. \tag{2.12}$$

From the QED Lagrangian, we are able not only to read off the basic properties of these fields, but also the specific interaction rules. While we won't explain here how to perform QFT calculations, it was thankfully noted by Richard Feynman that these calculations can be codified as little doodles–now called Feynman diagrams. The doodles are built from individual vertices or *Feynman rules*, and QED has just one, shown in Fig. 2.1 (left). The fundamental rule can be thought of as demonstrating the possibility for electrons to emit a photon, or for a photon to produce an electron-positron pair, or for electrons and positrons to annihilate into a photon–all depending how it is rotated. Using these rules, we are able to visualize the physical processes that this model allows, and an example is shown in Fig. 2.1 (right).

We pause here to note that the deceptive simplicity of Feynman's doodles hides some ugly truths about the theory. They are in fact *not exactly* accurate depictions of physical processes in the theory–they are terms in a perturbative expansion.

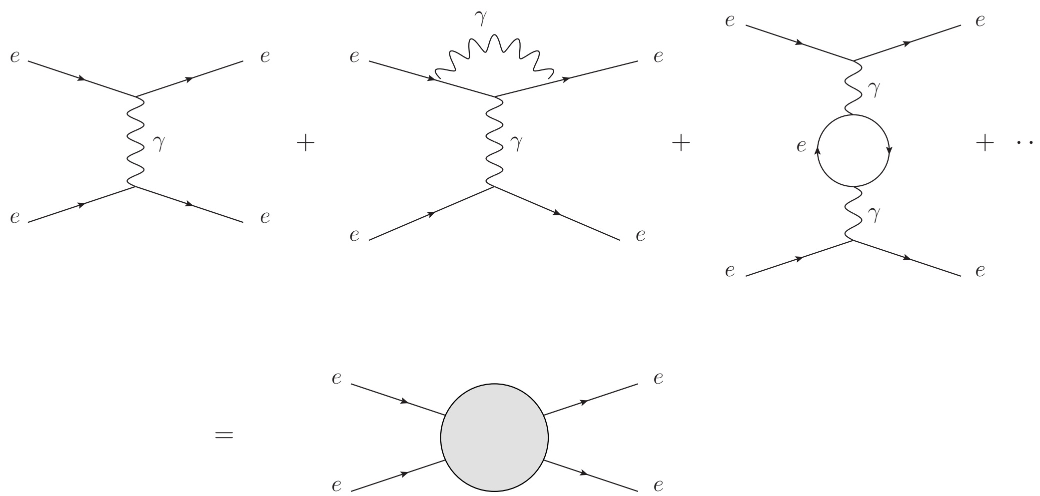

Figure 2.2: Three diagrams are shown in the perturbative expansion for electron-electron scattering. The full process has two electrons ingoing and outgoing–but we do not know exactly what happens in between, as indicated by the grey circle.

Like polynomials in a Taylor series, they each offer us only approximate glimpses of a true quantum field, translating an incomprehensible sea of interactions into individual approximations which can be understood intuitively and calculated. This means also that not all the particles in the Feynman diagrams are *real*. The real particles exist in the initial and final states of the diagrams. Anything created and destroyed in between is called a *virtual particle*. Virtual particles are not exactly real and not exactly fake–they are approximations.

For example, Fig. 2.2 shows multiple diagrams for electron-electron scattering in QED. Virtual particles are seen intersecting in closed paths, called loops. The more loops and vertices a Feynman diagram has, the harder it is to calculate. Often, diagrams with two or three loops are already at the cutting edge of evaluation. But the real process of electron-electron scattering is a combination of all these possibilities.

The accuracy of perturbation theory relies on the electromagnetic coupling strength $\alpha \propto \sqrt{q}$. Each vertex in a QED Feynman diagram carries a factor of α for the purposes of most physical calculations (discussed further in Chapter 6). Different interactions have different coupling strengths. The more vertices where particles interact that are present in a diagram, the more factors of α will be associated to it. Furthermore, in QFT, these couplings vary with the energy scale at which they are evaluated–they are called *running couplings*.

Despite the caveats, we are able to approximate many high-energy particle collisions with perturbative calculations in gauge theories where the coupling remains small, and the results agree with nature to a perhaps surprising extent. Luckily, for the electromagnetic interaction, α is generally quite small from the low energies of idle electrons ($\alpha \approx 1/137$) to all the way up to the energies obtained in particle colliders, so the series converges quickly. QED has produced some of the most accurately verified predictions ever made, such as the calculation of the electron's anomalous magnetic dipole moment, where experiment has verified the accuracy of the theory within 1 part in a trillion [4,5].

2.5 More Gauge Theories: the Nuclear Forces

Gauge theories form the basis of our understanding of particle interaction, giving rise to what are often called fundamental forces in particle physics. The strongest of these interactions is called, rather imaginatively, the *strong interaction* or *strong nuclear force*. It is a gauge theory with symmetry group $SU(3)$, describing the interaction of *quarks* through a force mediated by gauge bosons called *gluons*. Quarks are the tiny pieces which make up protons and neutrons, and the force between them is so strong that it prevents them from flying apart due to electromagnetic repulsion.

$SU(3)$ is a more complicated symmetry group than $U(1)$, with no easy analogy to visualize. In discussing some such terminology from group theory will inevitably enter our discussion of gauge theories without proper introduction. For a primer on the essential role that group theory plays in particle physics, the author recommends Ref. [6] or [7].

Rather than a conserved scalar charge q as seen in electrodynamics, quarks carry a three-dimensional analogue called *color charge*. This has led the interaction of quarks and gluons to be called quantum chromodynamics (QCD). $SU(3)$ is a non-Abelian group, meaning the generators do not commute. The Lagrangian at first looks similar to the QED Lagrangian:

$$\mathscr{L}_{\text{QCD}} = i\bar{\psi}(\gamma^\mu D_\mu)\psi - m\bar{\psi}\psi - \frac{1}{4}G_a^{\mu\nu}G^a_{\mu\nu} \,, \tag{2.13}$$

Where ψ is now the fermionic quark field. The big change comes from the new free gauge field term. This term is related to the commutator of the group elements, so it becomes more complicated in the non-abelian case, and we have had to define

$$G^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + g f^a_{bc} A^b_\mu A^c_\nu \,. \tag{2.14}$$

Here a, b, c index the 8 generators of the $SU(3)$ group that form the corresponding Lie algebra $\mathfrak{su}(3)$, and f^a_{bc} are called the structure constants of that algebra. These indices are implicitly summed over when repeated. The coupling constant for the gauge boson is now labeled g, i.e. $D_\mu = \partial_\mu - igA_\mu$. This form of Lagrangian generalizes to non-Abelian gauge theories, known collectively as *Yang-Mills theory*.

In fact, the simple case of QED is the exception and not the rule–most physical gauge theories are non-Abelian.

While this might seem very technical, in one sense the fundamental change is simple: (2.14) indicates that gluons themselves carry color charge, and its third term shows that can interact with each other. This is in contrast to photons, which mediate the electromagnetic interaction but to not self-interact. From a theory standpoint, this makes calculations difficult, and many known properties of QCD are difficult to demonstrate with rigor from the fundamental model.

QCD (along with Yang-Mills theory more generally) has a property called *asymptotic freedom*, which meant that the associated coupling constant α_s becomes small at high energies, and thus and perturbative calculations can describe collisions at particle colliders up to a point (as we will see in Chapter 6). However, the self-interaction of the gauge bosons leads to a situation where the theory is *strongly coupled* over large distances and therefore at low energies. This means that perturbation theory and its glorious Feynman diagrams break down entirely.

This is also why we don't observe quarks or gluons sitting around by themselves in nature–they quickly coalesce to form bound states. In fact, we don't observe them alone at particle colliders either–they form bound states before they even reach the detector. When a quark and anti-quark bind, it is called a meson. When three quarks bind, the resulting particle is called a hadron, a class of particles that includes protons and neutrons. Such bound states can be understood approximately via *effective field theories*, but their properties are incredibly challenging to deduce from the fundamental QCD Lagrangian. Only at high energies, like those achieved at the LHC, do quarks and gluons start to participate in the more fundamental interactions where perturbation theory holds.

The strong force can hold together protons in a nucleus against the repulsive electromagnetic force. Recall that the the standard electromagnetic force falls of with distance r as $F_{\text{EM}} \sim \frac{1}{r^2}$. However, as a consequence of non-perturbative effects and gluon self-interaction, the strength of the strong interaction has an additional exponential dampening term, $F_{\text{QCD}} \sim \frac{1}{r^2} e^{-\kappa r}$, where κ is a constant. Thus, its effects are felt only at the scale of the atomic nucleus. Such a scaling would be expected by a force with a massive mediator boson, and indeed the strong force was originally theorized to be mediated via π-mesons by Hideki Yukawa in 1935 [8]. While this is not true at the most fundamental level, the strong force effectively has a massive mediator boson due to the gluons' inability to travel unperturbed. Lone gluons and quarks will quickly form bound states such as mesons, hadrons, and hypothesized massive collections of interacting gluons called glueballs.

However, there is another nuclear force, which is mediated by massive bosons in a more fundamental sense. Understanding how this can occur was a giant leap in understanding, laying the foundations for what would become known as the standard model.

The other nuclear force is the weakest of known particle interactions (not counting gravity which is not yet understood on the particle level), so it is usually referred to simply as the weak interaction, suggesting that perhaps physicists should consider outsourcing the naming of their concepts. The weak interaction is responsible for many radioactive decays involving nucleons, electrons, and tiny neutral particles called neutrinos. Like the strong force, the interaction is confined to the nuclear scale, which might suggest massive mediator bosons. Furthermore, the famous Wu experiment [9] in 1956 demonstrated that the weak interaction violates a symmetry called *parity*[1], to the shock of many in the physics community. This essentially means that the interaction does not respect a mirror symmetry, and meaningfully distinguishes right and left directions unlike the electromagnetic and strong interactions.

Following the Wu experiment, a satisfying explanation for how a gauge theory could create the observed behavior of the weak interaction remained somewhat elusive. The full answer would require not only the unification of the weak and electromagnetic interactions, but also the introduction of a new scalar field to the model: the Higgs.

2.6 The Higgs Mechanism and Electroweak Unification

In the early 1960s, Glashow proposed a combination $SU(2) \times U(1)$ gauge theory as a way to to unify QED with the weak interaction, but reproducing the correct behavior required a forced breaking of the underlying gauge symmetry. Meanwhile, several authors (including Peter Higgs) were independently demonstrating that gauge bosons could acquire mass through a mechanism called *spontaneous symmetry breaking*. It was then Weinberg and Salam who put all of this together in the mid 1960s, independently, to build our modern understanding of electroweak theory– arguably the cornerstone of modern particle physics. The resulting interaction had three massive mediator bosons, the W^+, W^-, and the neutral Z, which obtained their mass through a natural symmetry breaking process involving a scalar field

[1]Shamefully, Chien-Shiung Wu who designed and performed this experiment was left out of the 1957 Nobel Prize, awarded to theorists Lee and Yang who proposed the test.

ϕ. This model was a triumph of theoretical particle physics; the W and Z bosons would not be observed until the 1980s, while the direct confirmation of the Higgs field as a physical entity would come only in 2012 at the LHC. We will try to give a whirlwind tour of this model.

The Higgs field is a scalar doublet $\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}$ where ϕ^+, ϕ^0 are complex scalars. The Higgs Lagrangian should have the standard form

$$L_{\text{Higgs}} = D_\mu \phi^\dagger D^\mu \phi - V(\phi). \tag{2.15}$$

The covariant derivative D_μ must respect the gauge symmetry of electroweak unification, and will turn out to be somewhat messy, so let us focus first on the potential term. In addition to the usual mass term, the Higgs potential contains a fourth order self-interaction term:

$$V(\phi) = \mu^2 \phi^\dagger \phi + \lambda(\phi^\dagger \phi)^2 \tag{2.16}$$

where μ and λ are conventional constants and $\phi^\dagger$ is the Hermitian conjugate of ϕ used to take inner products. In the standard model $\mu^2 < 0$ while $\lambda > 0$. The resulting shape of the Higgs potential is such that the ground state (or vacuum state) occurs when $| \phi | \neq 0$. The potential can most easily be visualized by restricting one component, say ϕ^+, to be zero, as shown in in Fig. 2.3. The resulting shape is often referred to as the *Mexican hat* potential by a bunch of European physicists who don't know the word *sombrero*. The Higgs field illustrates a concept called spontaneous symmetry breaking. Though the potential shown in Fig. 2.3 has a rotational symmetry, a ball placed at $\phi = 0$ is in unstable equilibrium, and will roll down in a random direction to break the symmetry. This principle is now essential

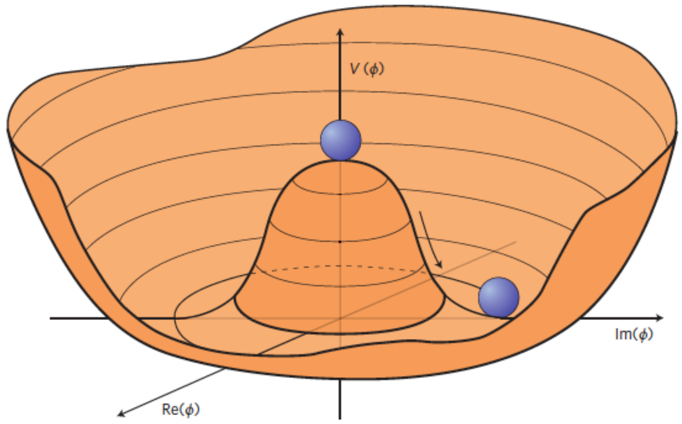

Figure 2.3: A visualization of the Higgs potential, restricted to the one component of the doublet ϕ (courtesy of John Ellis, CERN)

in our understanding of particle physics, especially in the context of the universe's evolution.

The electroweak sector of the standard model starts out as a standard gauge theory like QED and QCD, with a gauge group that is a product of two symmetries: $SU(2)_I \times U(1)_Y$, which are used to define the gauge covariant derivatives in the Eq. 2.15. The subscripts I and Y stand for the associated conserved quantities *weak isospin* and *weak hypercharge* associated with these symmetries, to distinguish them from other standard model symmetry groups such as $U(1)$ symmetry of QED. This gauge theory should have 4 massless mediators, but it is coupled to the Higgs field. As we can see, the ground state of the Higgs field spontaneously breaks this symmetry, giving the field a non-zero value.

This breaks the overall gauge symmetry of the model. Gauge bosons mediating interactions between electrons and neutrinos acquire a mass term through the Higgs field's vacuum expectation value. These include the charged W bosons, and the neutral Z boson. The Z boson in fact is a mixture of the original $SU(2)_I$ and $U(1)_Y$ mediators, as is the photon that mediates the remaining $U(1)$ symmetry. Thus, QED is what remains after the natural breakdown of a greater symmetry, the process of which separates out the weak force with its massive force carriers. The fact that these force carriers are massive dampens the effect of the force over large distances, leaving its effects confined to the scale of the atomic nuclear force–which is why it is considered one of the *nuclear forces*.

It doesn't really matter *which* way the ball rolls down the hill, as the minimum it will find is part of a rotationally symmetric ring. We can see looking at (2.16) that the minimum occurs when $|\ \phi\ |= \frac{|\mu|}{\sqrt{\lambda}} \equiv v$, called the vacuum expectation value. The excited states of this field will thus be perturbations about this vacuum. Because of the symmetry inherent in $V(\phi)$, without loss of generality we can take the ground state to fall such that the first component of ϕ is zero and the second component lies on the real axis. A small excitation of the vacuum state will then look like

$$\phi = \begin{pmatrix} 0 \\ v \end{pmatrix} \rightarrow \begin{pmatrix} 0 \\ v + H \end{pmatrix} \tag{2.17}$$

Upon quantization, this excitation H is the famed Higgs boson, discovered in 2012 at the LHC [10, 11]. The discovery of the particle itself was such a landmark because it verified the much broader model of the Higgs field and the concept of spontaneous symmetry breaking.

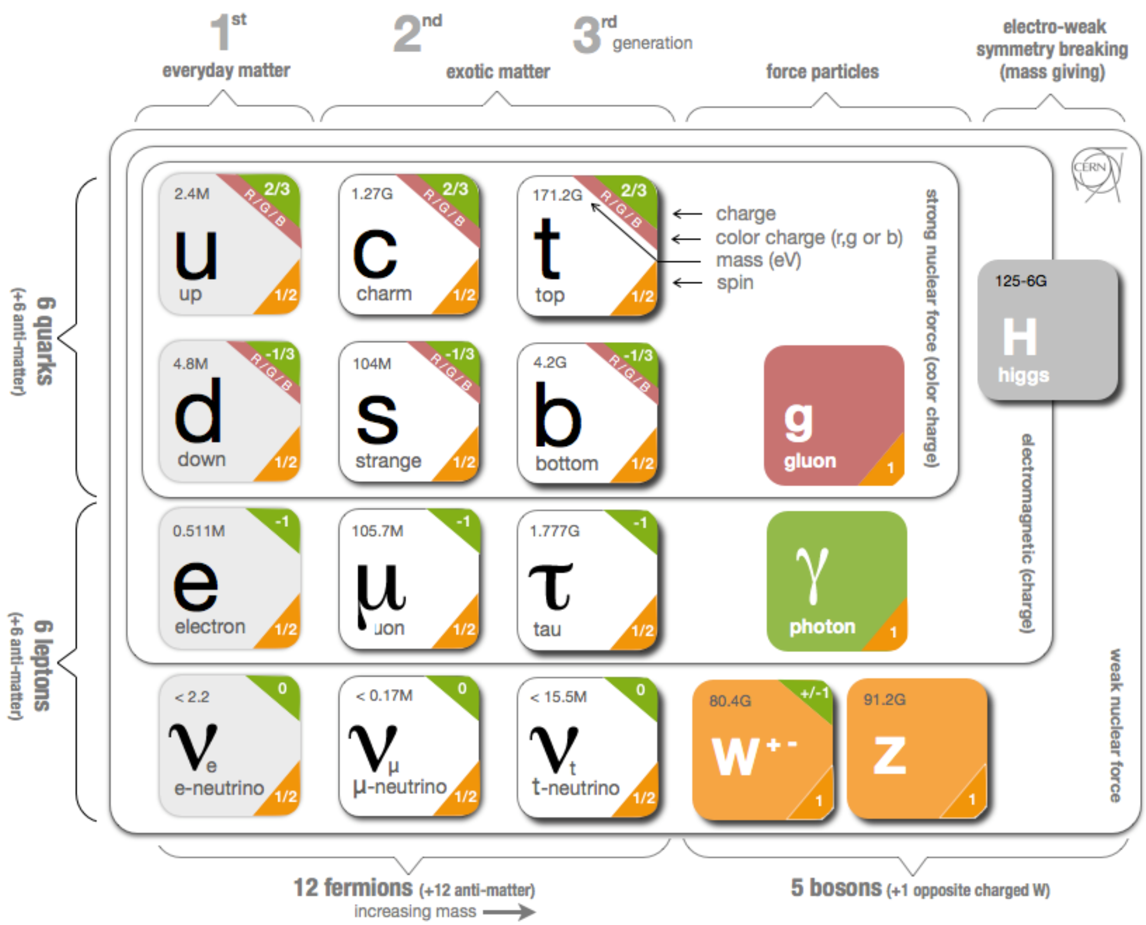

Figure 2.4: A visualization of the standard model of particle physics, courtesy of CERN

2.7 The Standard Model

The standard model is, essentially, a set of gauge field theories within the framework of QFT. It is the combination in full of the Electroweak interaction (including the Higgs field) and QCD, and is often described as one big gauge theory with symmetry group $SU(3) \times SU(2) \times U(1)$. A summary of the model, and the particles it contains, are shown in Fig. 2.4

A few things are worth noting that have not yet been mentioned. Notably, the standard model contains three generations of matter. Stable electrons have heavier counterparts in the other matter "generations," called muons (μ) and tau leptons (τ). While these particles are produced in accelerators and in the atmosphere, they are fairly short-lived and decay into lower-mass particles (though atmospheric muons survive long enough that many of them are harmlessly passing through you and your surroundings as you read this!). Similarly, protons and neutrons are made of up and down quarks, which are stable. However, two more generations of heavier quarks exist which can combine into many unstable bound states called exotic hadrons and mesons. The heaviest quark, and the heaviest particle in the standard model, is the top quark, which will be discussed later in this dissertation.

Another thing to mention is *chirality*. Fermions can come with two different types of *chirality*, separated by applying the left-handed and right-handed projection operators $P_L = \frac{1}{2}(1 - \gamma^5)$ and $P_R = \frac{1}{2}(1 + \gamma^5)$ respectively, where $\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3$ is the product over Dirac matrices. Thus, we say that all fermions can be either left-handed or right-handed. There doesn't have to be a meaningful difference in how these two behave, but in the parity-violating weak interaction, they are separated.

Under the the $SU(2)_I$ symmetry in electroweak theory, left-handed fermions transform as a doublet. Specifically, these doublets are formed by combining *up-type* and *down-type* quarks from the same generation:

$$\begin{pmatrix} u \\ d \end{pmatrix}, \begin{pmatrix} c \\ s \end{pmatrix}, \begin{pmatrix} t \\ b \end{pmatrix}; \tag{2.18}$$

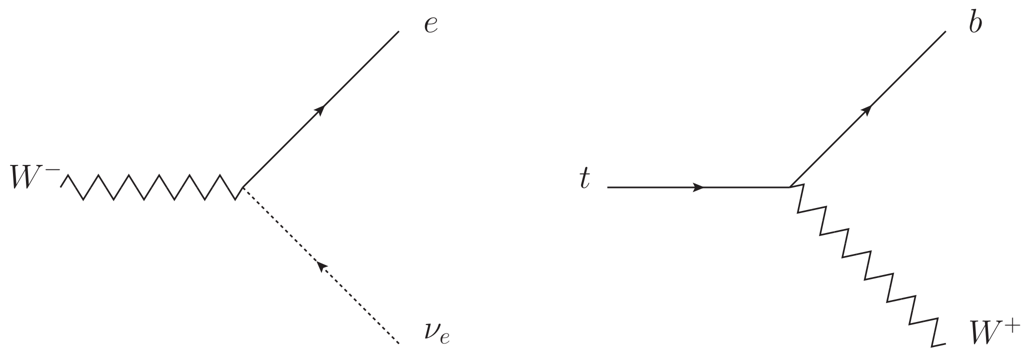

Figure 2.5: Examples of weak interactions involving particles which come together in $SU(2)$ doublets.

or a massive fermion and a neutrino from the same generation,

$$\begin{pmatrix} e \\ \nu_e \end{pmatrix}, \begin{pmatrix} \mu \\ \nu_\mu \end{pmatrix}, \begin{pmatrix} \tau \\ \nu_\tau \end{pmatrix}. \tag{2.19}$$

These same particles are naturally linked by Feynman vertices, such as those shown in 2.5. All quarks and leptons can also come with right-handed chirality, in which case they transform only as a singlet under the action of the $SU(2)_I$, and have a different value of hypercharge Y. While the two chiralities participate in most interactions equally, this is not true for interactions involving a W boson, and there is an even more glaring difference: right-handed neutrinos have not been observed in nature at all, and are not included in the standard model.

2.8 Mass in the Standard Model

On the topic of chirality, there was another profound consequence of the Electroweak unification, which was not originally intended. When writing our the full standard model Lagrangian, one obviously expects to see mass terms for the massive fermions. In the Dirac Lagrangian (2.5), such a term could be added trivially. However, in the full standard model, a mass term would have to include in equal measure left and right-handed fermions, despite their different transformation properties. A simple number m cannot fulfill this requirement as it does not obey the correct transformation symmetries. The only candidate for a mass term would be a simple linear or *Yukawa*-type coupling between the Higgs field and all massive particles. Consider, as an example, the electron. It transforms under an $SU(2)$ doublet L together with ν_e, and a right handed singlet R all by itself. The piece of the standard model Lagrangian containing a Yukawa interaction between an electron and the Higgs field ϕ would read

$$\mathscr{L}_{\text{Yuk}} = -g_e(\bar{L}\phi R + \bar{R}\phi^\dagger L) \tag{2.20}$$

After the symmetry breaking in Eq. 2.16, the neutrino field in L drops out and this becomes

$$\mathscr{L}_{\text{Yuk}} = -g_e(\bar{e}_\text{L} v\, e_\text{R} + \bar{e}_\text{R} v\, e_\text{L}) - g_e\left(\bar{e}_\text{L}\frac{H}{\sqrt{2}}e_\text{R} + \bar{e}_\text{R}\frac{H}{\sqrt{2}}e_\text{L}\right). \tag{2.21}$$

This shows that the resulting mass term comes directly associated with a Yukawa coupling to the Higgs boson H. All massive fermions interact with the Higgs boson in this way,allowing for Feynman vertices such as that shown in Fig. 2.6. Here the interaction strength for a given fermion is the Yukawa coupling constant g_f, and we obtain the mass-Yukawa relation

$$m_f = \frac{g_f\, v}{\sqrt{2}}. \tag{2.22}$$

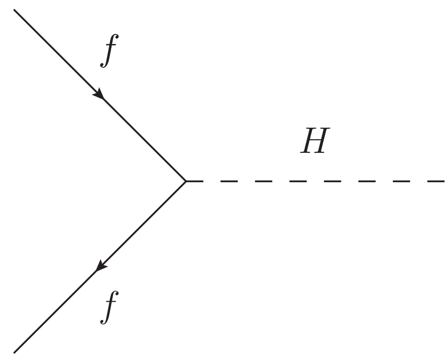

Figure 2.6: A Yukawa interaction vertex between a fermion f and the Higgs boson H

It is sometimes said that the Higgs boson thus gives mass to all particles, though some would argue that this ignores the subtlety of mass-generation in QCD bound states such as protons, not to mention the mystery of neutrino masses. Either way, the generation of lepton and quark mass through the Higgs mechanism is a core feature of the standard model. It also provides an avenue to test the validity of the standard model, as coupling strengths and mass are typically measured through very different experimental techniques.

In one sense, the idea of fermions acquiring mass from the non-zero vacuum energy density of the Higgs field after spontaneous symmetry breaking is a beautiful and thought-provoking concept. In another sense, it raises more questions than it answers. For example, the fermion masses suggest that the Yukawa coupling values vary by up to 5 orders of magnitude. The top quark, in particular, has $g_{\mathrm{t}} \approx 1$, while the electron has $g_{\mathrm{e}} \approx 3 \times 10^{-6}$.

Furthermore, we find that it is impossible for neutrinos to have mass in the standard model through this mechanism. While neutrino mass must be very small based on observation (bounded above by around 10^{-6} times the electron mass), they firmly are known to have non-zero mass based on neutrino flavor-mixing observations. There is no agreed upon way to include such a mass in the standard model, and it remains an open area of study.

2.9 A Big, Beautiful Mess of Particles

Unlike general relativity, the standard model is not a single elegant stroke of genius, but rather a mysterious and somewhat haphazard collection of objects. It does not explain why the masses of most particles are what they are, and has at least 26 free parameters with no clear explanation. Nonetheless, whatever it lacks in elegance, it makes up for in reliability. The standard model has outlived many attempts to concoct a more aesthetically pleasing set of objects, and to this date has accurately described the results of decades of particle physics experiments.

There are many ways to summarize the standard model, but it is unavoidably rather complex. It is often said to describe three of the four fundamental forces of nature, with only gravity missing. But this description really steamrolls the complexity of the electroweak interaction, and ignores the subtle role of the Higgs field, which generates arguably an entirely new particle interaction. Furthermore, particle physicists have tried their best to create images (as in Fig. 2.4) which look as minimal as possible, sometimes paving over the fact that the weak interaction has three mediators while gluons come in 8 varieties. Even in this chapter, we have glossed over some subtleties such as CKM mixing and neutrino flavor mixing, which add another layer of complexity to the story.

In many ways, the standard model of particle physics resembles the periodic table, which appeared disorganized and random at the turn of the 20th century, until the theory of quantum mechanics explained its underlying structures. Will we ever see a similar revolution, that provides a more elegant explanation for the structure of the standard model? If humanity can prevent catastrophic climate change and avoid nuclear war, perhaps such a discovery will one day be made.

Chapter 3

Unsolved Problems in Particle Physics

"I am a mystery wrapped in an enigma wrapped in a pita. Why the pita? That counts as another mystery."

— Demetri Martin

3.1 Things that are Conspicuously Missing from the Standard Model

In many ways, the tension in fundamental physics is driven by apparent incompatibilities between observations at very small scales and observations at very large scales. This goes beyond the fundamental incompatibility of quantum physics and gravity, an issue which becomes evident only when examining singularities appearing in general relativity related to black holes and the big bang. One need only to look at the large scale motion of our own galaxy to see that something is missing in current models.

The rotation curve of the Milky Way, and those of many other spiral galaxies, were found by Vera Rubin and others to differ from their expected behavior in the 1970s (see Ref. [12] for a review). The most obvious explanation is that a large quantity of non-luminous, seemingly invisible matter is floating around throughout our galaxy without interacting much with ordinary matter. Stars and dust that we can see appear to feel its pull through gravity, but do not interact with this mysterious invisible substance, called *dark matter*. If it interacted by the strong or electromagnetic forces, we would be able to detect it. Furthermore, dark matter seems to account for more mass in our universe than ordinary matter.

Since the initial arguments for the existence of dark matter, additional evidence has been found for this substance through gravitational lensing and cosmological models. Many believe that dark matter is made up of heavy particles which interact only weakly, and are missing from the standard model. They could in principle be heavier than what can be directly produced at modern colliders, and interact rarely enough to evade direct detection, being observed only by their gravitational pull. While this is the favored explanation, there could always be something unexpected at play.

On top of this, another mysterious force exists in the universe at even larger scales. While general relativity posits that the universe is expanding ever since the big bang, the rate of expansion seems to be accelerating. This was first observed by the Supernova Cosmology Project [13] and the High-Z Supernova Search [14] in 1998[1], and implies the presence of yet another mysterious substance pervading the universe, called *dark energy*. This force, unlike all others, seems to push massive bodies apart at large scales.

Curiously, a potential explanation for dark energy is already present within QFT, in the controversial concept of vacuum energy. This type of energy would exhibit exactly the right type of behavior, except for one thing: Based on the standard model, the hypothesized energy density appears too large by at least 100 orders of magnitude. While it is technically possible for as-yet-undiscovered quantum fields to yield a partial cancellation, the cancellation would have to yield a result so miraculously close to zero that many view the vacuum energy as a mathematical oddity and not a viable cosmic force. Still, it is clear that physics is missing something here.

Lastly, a more tangible problem in the standard model is that of neutrino mass. Neutrinos are observed to have a miniscule mass relative to all other particles. Yet they have no mass term at all in the standard model, and one cannot be added in quite the same way as for other fermions, because neutrinos seem to exist only as left-handed particles. One interesting possibility is the see-saw mechanism, which posits that right-handed neutrinos do in fact exist and are very heavy, making them possible candidates for dark matter. As neutrino physics is an active area of experiment today, there is hope that theorists may receive some guidance on this issue.

[1] These studies garnered a shared Nobel Prize in 2011. However, Vera Rubin was not awarded a Nobel prize for her seminal work on galactic rotation curves before her death in 2016, in another obvious injustice that suggests we should not place so much value in these awards.

3.2 Quantum Fields: Rough Around the Edges

We alluded in Chapter 2 to the fact that QFT is rather messy from a mathematical standpoint. In QFT, true quantum field interactions involve arbitrarily complex diagrams which often cannot be calculated even with today's computing resources. On top of that, most Feynman diagrams hide integrals that yield infinite results upon first attempt to calculate anything physical. They must be put through a process called *renormalization*, whereby physical quantities are redefined specifically to avoid divergent quantities. While the process of renormalization has had great success in making physical predictions, and the formulation of the renormalization group approach by Wilson in the 1970s [15] sheds some additional light on the once-controversial process, it has one fundamental issue: it is not mathematically rigorous. This is not a deal-breaker, as Newton's calculus was wildly successful for over a century before being made fully rigorous by mathematicians. Nevertheless, it sets an obvious direction for needed mathematical work.

While the process of renormalization can be quite technical, the conceptual issue is simpler: renormalization is used to answer a question in physics when the original phrasing of the question, resulting from the equations in QFT, has no apparent solution. Renormalization requires one to redefine the quantities in the problem until divergences cancel and a finite result can be obtained. This redefinition means that the original question has not been answered in a strict sense, creating an ambiguous divide between the equations of the theory and the results of the theory. For an excellent discussion on the concept of renormalization in a broad sense, see Arnab Kar's dissertation [16].

Such an issue is not so different from Newton's use of *infinitessimals* in his formulation of calculus–these objects were not defined in a mathematically rigorous way, but were used to produce many results nonetheless. However, they start to come up short when one starts to push calculus to its extreme limits, studying functions that appear bizzare and physically unintuitive. Unfortunately, most objects in QFT are bizzare and physically unintuitive, suggesting that they may benefit from more rigorous mathematical formulation.

Because there are situations where renormalization breaks down (strongly-coupled theories and gravity to name a few), it will ultimately be important to have a rigorous understanding of the procedure so we may better understand its limiting cases.

3.3 The Strongly-Coupled Regime

Despite the caveats of perturbative expansion and renormalization, those procedures take place under conditions under which QFT is still able to thrive. The theory succeeds when the interaction coupling strength is small, as the perturbative approach of Feynman diagrams and renormalization converges quickly. Then the contributions of diagrams die off as they become more and more complex. But this is not always the case, particularly for the strong interaction.

The phenomenon of running couplings in QFT mentioned in Sec. 2.4 is a direct consequence of renormalization, causing the apparent coupling strength of an interaction to vary as a function of the energy/length scale at which it is studied. In the case of the strong interaction, the coupling is small at very high energies, allowing us to predict the results of high energy particle collisions involving quarks and gluons. However, the coupling becomes large at low energies, causing quarks to join together and hadronize into bound states. QFT, for all its success, struggles to demonstrate exactly how quarks form bound states, or to explain why the proton mass is so much higher than its constituent valence quarks. This is because perturbative calculations and renormalization break down at strong coupling values. Some success has been seen in modelling these non-perturbative effects numerically through lattice field theories, for example in calculating meson masses (see Ref. [17] for a review).

Lattice QCD results, coupled with observations and effective field theories, generally are believed to provide sufficient evidence for the validity of QCD and the quark model. Nonetheless, the inability to perform calculations involving real world particles based on first principles is troubling. The strongly-coupled regime eludes formal description by the tools that we have. This leaves open the meaning of other hypothesized field theories, such as those which become strongly coupled at high energies–can such a theory exist in nature? QED was thought to be such a theory before EW unification, but the additional symmetry group removed divergences at high energy scales.

To this day, understanding the low energy behavior of QCD directly ties into one of the famous Millenium Problems [18]–so you could win $\$1,000,000$ if you figure it out! The relevant Millenium Problem is to demonstrate that Yang-Mills theories, which include QCD, have a *mass gap*. Theories with a mass gap do not have massless excitations. In QCD, this would mean that gluons are somewhat a ruse, existing only as virtual particles in perturbative diagrams at high energies. All

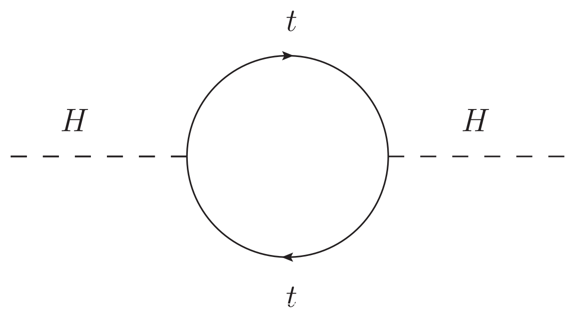

Figure 3.1: A diagram for a top quark loop correction to the Higgs propagator, which will affect the Higgs mass renormalization. Because the Higgs boson couples to all massive fermions, it is influenced by many such diagrams.

real excitations of the gluon field would have to be a bound state of many interacting constituents which acquires a mass. This is widely believed to be the case.

3.4 The Hierarchy Problem

The hierarchy problem can be stated in a few ways. First, we note that at present there is only one independent parameter in the SM which is not dimensionless in natural units: essentially, the Higgs mass. The mass of the Higgs boson sets the energy scale for the electroweak interaction, and arguably for much of the standard model, though we are ignoring a more subtle scale that arises in QCD here. Quantum Gravity should have a natural scale on the order of the Planck mass m_{pl}, 16 orders of magnitude larger than the Higgs mass. What can explain this enormous difference? In one sense, this question itself constitutes the hierarchy problem. In another sense, there is a more specific issue at play.

The hierarchy problem can also be more technically posed in terms of the Higgs mass renormalization. The trouble comes from the fact that the Higgs couples to other particles via loop diagrams, as in Fig. 3.1. Such diagrams must be included in the Higgs mass renormalization, and their contribution depends on the mass squared of the particle appearing in the loop. This means that any heavier particles we have not yet observed ought to produce large corrections to the Higgs mass. Indeed, if there exist new physics at energies much higher than the EW scale (which the Planck mass strongly suggests), it seems impossible that the Higgs mass would remain comfortably within the EW regime, as an extreme fine-tuning is needed to balance the larger and larger loop corrections. Such fine-tuning is detested by particle physicists. In general, it seems possible for nature to possess terms at different scales that bear little relation, and also for many exact cancellations. However, incredibly unlikely *approximate* cancellations are frowned upon without a good explanation.

The favored resolution to this problem has been supersymmetric (SUSY) models, in which the additional particles create a cancellation effect above the SUSY energy scale, removing the need for fine-tuning. However, with no direct evidence of SUSY, the potential energy scale for SUSY has been pushed ever higher, while the measured Higgs mass that fits comfortably with current SM data. It is not clear how new physics could eventually come into play at higher energies without greatly effecting the Higgs mass, unless these effects are wildly different than the particle interactions we understand today.

Honorable Mentions

We pause here to note a few more mysteries of the standard model that may pique the interest of some readers: There is the strong CP problem [19], in which a viable CP-violating term is conspicuously absent from the QCD Lagrangian, a fact that has been used to conjecture about hypothetical particles called axions. Another fundamental problem is the discrepancy between observed CP violation in the standard model and matter anti-matter asymmetry in the visible universe [20], which seem incompatible. Lastly, for those interested in early universe cosmology, the lack of a plausible bridge between the standard model and the model of inflationary cosmology remains an open and intriguing issue [21].

Chapter 4

Theoretical Paths Forward

"Good books tell the truth, even when they're about things that never have been and never will be. They're truthful in a different way"

— Stanisłav Lem (one could say the same of good theories)

4.1 Maybe More is Better

The standard model acquired its name only after decades of stubbornly refusing to be supplanted by a more elegant theory. In the 1970s, as the pieces of what is today known as the standard model were coming together, particle physics seemed to be progressing at breakneck pace on theoretical and experimental fronts. Theorists had a good track record solving conceptual issues, and experimentalists still had some heavy quarks and bosons left to discover. On the theoretical side, even more bold ideas came about into the 80s and 90s, with grand ambition. There was the proposal of a new type of symmetry between fermions and bosons, referred to as supersymmetry (SUSY) which would basically double the number of fundamental particles from those in the standard model. Following the success of Electroweak unification, there were attempts unify the three known forces under one higher-dimensional gauge group and create a grand unified theory (GUT). There was also the idea of string theory, which was able to propose a renormalizable theory of quantum gravity, at the cost of positing several hidden spatial dimensions existing in our universe.

In each case, adding *more* ideas to the standard model seemed to increase the elegance of particle physics as a whole, and paint a less haphazard picture of our universe. There is immense mathematical beauty in each of these ideas, which helped them to gain popularity. And yet, as of June 2020, none of them have seen any hint of experimental confirmation.

Theoretical research in the field of of high energy physics can today feel somewhat adrift. It took over 50 years for the Higgs mechanism to be verified experimentally since its theoretical inception. But in advance of the Higgs discovery, there were very convincing arguments that LHC experiments would observe either the standard model Higgs boson or some other new physics to compensate for its absence (for a simple overview of pre-LHC Higgs constraints see Ref. [22], or for more detail on precision Electroweak fits see Refs. [23, 24]).

Models which would take us beyond the standard model, in particular supersymmetry, tend to have too many free parameters to make similar predictions about how and where their effects should first be observed. As we speak, LHC experiments continue to rule out small portions of endless phase spaces in which supersymmetry could be detected. Yet, whether they will admit it today or not, many particle physicists were confident 10 years ago that the LHC would have already seen strong evidence of supersymmetry by now. Some even viewed this as the primary purpose of the LHC. The lack of such a discovery to date certainly does not disprove the model, but it has cast a shadow of skepticism on the model's strongest proponents.

Some would argue that supersymmetric models have received too much focus, dominating the conversation due to their prominent proponents in the field. Perhaps this has indeed distracted from other possibilities and open issues in the standard model and QFT more broadly. We will examine some avenues of research which proceed very differently, simplifying QFT and honing in on its individual characteristics without proposing bold new paradigms.

4.2 Mathematical Rigor in QFT

QFT has an obvious shortcoming in its lack of mathematical rigor. To get a sense of what mathematical tools are missing to define a more rigorous results, we will briefly discuss some issues with the *path integral formulation* of QFT.

The idea of a path integral predates QFT, and is rigorously defined in a select few cases. One such case is the famous process known as Brownian motion, whose formal mathematical description is often referred to as the Wiener process. This refers to a continuous random walk $W(t)$ which results as the limit of independent increments. In 1-dimension, $W(0) = 0$ and we have that W is Gaussian distributed at each point $p_W(x,t) \propto e^{-x^2/2t}$. It then has a variance

$$\mathrm{Var}[W(t_2) - W(t_1)] = \langle |W(t_1) - W(t_2)|^2 \rangle \propto |\ t_2 - t_1\ | \tag{4.1}$$

Using the mathematical tools of *stochastic calculus*, it is possible to define an integral that allows us to average a function over all such paths leading from one fixed endpoint to another, called the Wiener integral and often generalized to the Itô integral. We could write such an average of a function f, using the standard physicist stylings, as

$$\langle f \rangle = \frac{\displaystyle\int_{W(t_1)=x1}^{W(t_2)=x_2} f(W)\mathcal{D}[W]}{\displaystyle\int_{W(t_1)=x1}^{W(t_2)=x_2} \mathcal{D}[W]}\,. \tag{4.2}$$

where $\mathcal{D}[W]$ indicates that we are integrating over all paths of Brownian processes $W(t)$. This is a far more complex calculation than ordinary integration from Newtonian calculus. By integrating over all possible paths, we must integrate an infinite dimensional system. Additionally, Brownian motion has famously rough paths– though continuous, they are nowhere differentiable. A sample of two approximately generated Brownian paths with the same endpoint is shown in Fig. 4.1, which would be among uncountably infinite sample paths in a Wiener integral.

In the path integral formulation of QFT, we must perform a further generalization of this concept, known as a functional integral. Consider a scalar field ϕ. The

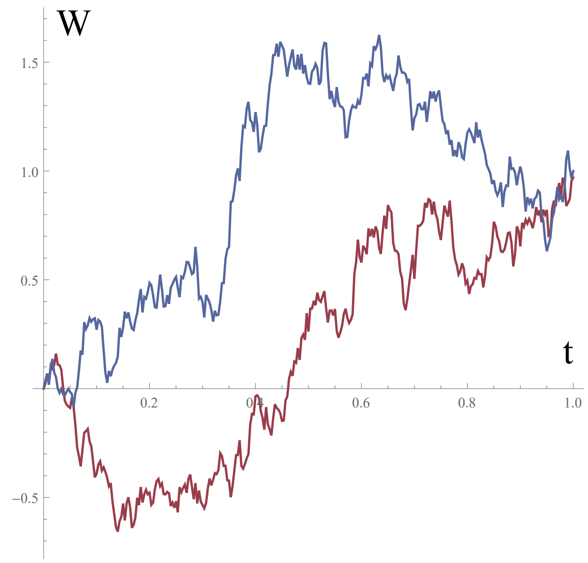

Figure 4.1: Two approximately generated Brownian paths $W(t)$ with the same endpoint, representing sample paths in a path integral.

necessary integration to perform an average of an observable f is written as

$$\langle f \rangle = \frac{\displaystyle\int f(\phi) e^{-iS(\phi)} \, \mathcal{D}[\phi]}{\displaystyle\int e^{-iS(\phi)} \, \mathcal{D}[\phi]} \tag{4.3}$$

Where the action S is the integral of the Lagrangian density, $\int \mathcal{L} \, dx^4$. The mathematical tools to rigorously evaluate such a functional integral generally do not exist. Even from a physicist's standpoint, the main such integral which can be evaluated is that of a non-interacting field, also called a *free field*, with only a kinetic term. The purpose of perturbative Feynman diagrams is, in essence, to relate interacting fields back to this one solvable case.

To get a sense what is so difficult, we consider a typical field configuration. In QFT, one of the simplest and most important calculations we can perform is called the *two-point function*, $\langle \phi(x_1)\phi(x_2) \rangle$, which is related to the correlation of the field values at these points. For a free scalar field in 4 dimensions, we arrive at

$$\langle \phi(x_1)\phi(x_2) \rangle \propto \frac{1}{\mid x_1 - x_2 \mid^2} \tag{4.4}$$

Such a field ϕ has a divergent two-point function as $x_1 \rightarrow x_2$, and is a strange object to contemplate. How can the expected product of the field value at two points vanish when the points are far away, yet diverge when they are close? The answer is they cannot–the stochastic sample configurations we are integrating over do not actually *live* in 4-dimensional space. They belong to the realm of *distributions*, mathematical objects that only make sense inside of an integral–such as the well-known Dirac δ-distribution. This is a stark difference to the traditional wavefunction $\psi(x)$ in quantum mechanics, whose physicality may be debated, but which is nonetheless mathematically well-defined in most cases.

This is obviously an area ripe for research, but it is mostly of a mathematical nature. Because the language of formal mathematical rigor is often completely different from the language used for physical calculation, research in this direction generally is classified as *mathematical physics*, an area populated mostly by mathematicians. To make any rigorous claims, one must usually work on reduced-dimensional theories with simple interactions. Thus, the work being done by mathematicians in the field of mathematical physics is often completely disparate from that being done in the realm of theoretical physics. The two worlds often diverge in terms of both interest and notation, making it a difficult divide to bridge.

However, when that divide is bridged, it can lead to wondrous results, such as Martin Hairer's work with regularity structures in Ref. [25]. Hairer, a physics PhD from the University of Geneva turned mathematician, has made substantial progress generating rigorous solutions to problems in stochastic calculus which map exactly to quantum field theories, in particular 3-dimensional $\lambda\phi^4$ theories. These are scalar field theories with a Higgs-like self-interaction term, but without the other machinery of the standard model. This work earned him a Fields medal in 2014. Perhaps if mathematicians and physicists learn one another's language and come together more often, we can expect more progress like this in the future.

4.3 Research Project: Continuity of Log-Correlated Gaussian Fields

As mentioned, Brownian motion has paths which are in a sense *barely* continuous. It is not Lipschitz continuous, meaning that for a Brownian path $W(t)$ the equality

$$\frac{|W(t_1) - W(t_2)|}{|t_1 - t_2|} < C \tag{4.5}$$

does not hold for any C as $x \to y$; the quantity is ubounded. However, it is possible to find an optimal function ω with respect to which $W(t)$ is Lipschitz continuous (see, for example, Ref. [26]). That is, there exist C such that

$$\frac{|W(t_1) - W(t_2)|}{\omega(|t_1 - t_2|)} < C \tag{4.6}$$

is true with probability 1. Such a function is called a *modulus of continuity*. One can think of it as measuring the severity of the divergence of the slope of $W(t)$. The modulus of continuity for $W(t)$, attributed to Lèvy is

$$\omega(|t_1 - t_2|) = \sqrt{|t_1 - t_2| \log \frac{1}{|t_1 - t_2|}}. \tag{4.7}$$

Brownian motion is actually related to a free quantum field in 1 dimension, which is not much of quantum field at all. All quantum fields in higher dimensional path integrals have distribution-valued sample configurations, and so are divergent. The most mildly divergent are free fields in 1+1 dimensions (1 space + 1 time). These fields have a logarithmically-divergent two point function. We can consider, for simplicity, a 1-dimensional field with the same correlations–this is of interest in the mathematical community [27], and is related to the ground state wavefunction of a 1+1 dimensional free quantum field from a physics perspective. For such an object, a physicist would write

$$\langle \phi(x_1)\phi(x_2) \rangle \propto -\log |x_1 - x_2|. \tag{4.8}$$

To be more mathematically rigorous, we must recognize that ϕ is distribution-valued, and makes sense only when integrated against a test function h where

$$\phi[h] = \int \phi(x)h(x)dx, \quad \int h(x)dx = 0. \tag{4.9}$$

Then we can say, more precisely, that

$$\langle \phi[h_1]\phi[h_2] \rangle \propto -\int \log |x_1 - x_2| h_1(x)h_2(y)\, dx\, dy. \tag{4.10}$$

Working with S.G. Rajeev, the author performed a study to compare such fields to Brownian motion by implementing simple test functions [28]. The low severity of logarithmic divergences allows us to use particularly simple test functions. For constants s and u, we can define a moving average of ϕ over an interval of width $2s$ centered at su as

$$\bar{\phi}_s(u) = \frac{1}{2}\int_{-1}^{1} \phi\left(s\left[u - w\right]\right) dw \tag{4.11}$$

There is some subtlety here as test functions must retain mean zero, but we preserve the basic properties of a moving average. This allows us to have in mind a simple picture for the output of our field ϕ against the test function–it is as if we are able to perform a sliding measurement of the field potential across a finite region. The output is easily shown to be a function, but whether it possesses continuity is not immediately clear.

Our test function allows us to define a metric ρ on the field

$$\rho(u, v) = \sqrt{\langle[\bar{\phi}_s(u) - \bar{\phi}_s(v)]^2\rangle}. \tag{4.12}$$

It is the existence of this metric which allows us to compute the modulus of continuity of our averaged field. After a short exercise outlined in [28], we find the modulus of continuity to be

$$|x - y| \log \frac{1}{|x - y|}. \tag{4.13}$$

This is similar to the standard result for $W(t)$ in Eq. 4.7, without the square root. Our averaged fields are indeed continuous, and not so different conceptually from the Brownian paths shown in Sec. 4.2.

The goal of this project was to explore the links between simple divergent quantum fields and well-studied continuous stochastic processes. We see a straightforward example of how a simple choice of test function for our quantum field creates a more familiar mathematical object with more tractable properties. This project was a fairly short exercise, and it would be interesting to see further study of the possible translation from divergent quantum fields to more ordinary functions, as well as the significance of metrics such as that defined in Eq. 4.12, the idea for which came from work done by Rajeev and Kar [29].

A

4.4 Toy Models

It is difficult to convey how challenging it is to perform complex calculations in Yang-Mills theories that describe the most interesting standard model processes. Even when perturbation theory succeeds, calculations for physical processes at 3rd order are in many cases considered cutting edge. Despite the fact that we have exponentially more computing capability afforded to us than at the time these theories were developed, we are unable to fully automate many calculations. A deep knowledge of physical properties and various renormalization techniques is required to extract meaningful results from complicated Feynman diagrams.

Rather than using mathematical techniques to search for rigorous definitions in reduced-dimensional models as in Sec. 4.3, we can take a more physics-oriented approach and study so-called *toy models*. These are simplified interacting models, often with reduced dimension $d < 4$, which can be studied via theoretical physics methods instead of pure mathematics.

Recall the unsolved conjecture of a mass gap in Yang-Mills theories such as QCD from Sec. 3.3. Another model, called the *principal chiral model*, has long been considered a toy model for Yang-Mills theory. The principle chiral model involves group-valued field g in 1+1 dimensions with a target space of $SU(2)$ (represented as matrices).

$$\mathcal{L}_{\text{PCM}} = -\frac{1}{\lambda^2}\text{Tr}(g^{-1}\partial_\mu g \cdot g^{-1}\partial^\mu g) \tag{4.14}$$

$\text{Tr}(X)$ is the trace of a matrix X, and λ a coupling constant. This model is a special case of so called *non-linear sigma models* arising in string theory, in which a scalar field theory takes values on any non-linear manifold called a target space.

The principal chiral model shares the property of asymptotic freedom with Yang-Mills theory. Most notably, it was shown to have a completely factorizable S-matrix, meaning that the perturbative expansion in QFT can in principle be evaluated at arbitrary order [30,31]. This led to calculations indicating that a dynamical mass gap is generated in the theory, despite the apparent fundamental massless excitations– exactly as is conjectured to occur in QCD. As such, this toy model provides a proof of principle for what we expect to occur theories like QCD, but the computations are unfortunately too difficult to confirm this directly in those higher dimensional theories.

4.5 Research project: Nonlinear Wave Solutions in a Strongly-Coupled Toy Model

Working with professor S.G. Rajeev, the author examined a toy model related to the principal chiral model [32]. We consider a field ϕ in 1+1 dimensions (t and x), taking values in the Lie algebra $\mathfrak{su}(2)$. This can be represented as a 3-component decomposition into the standard Pauli matrices σ_i: $\phi = \frac{1}{2i}\left[\phi_1\sigma_1 + \phi_2\sigma_2 + \phi_3\sigma_3\right]$. We can construct a model containing a third-order interaction term with Lagrangian density

$$\mathcal{L}_{\text{toy}} = \frac{1}{2}\partial_\mu\phi^a\partial_\nu\phi^a\eta^{\mu\nu} + \frac{\lambda}{6}\epsilon_{abc}\phi^a\partial_\mu\phi^b\partial_\nu\phi^c\epsilon^{\mu\nu}, \tag{4.15}$$

where ϵ is the Levi-Civita tensor. If we make the indentification

$$\partial_t\phi = \frac{1}{\lambda}g^{-1}\partial_x g, \quad \partial_x\phi = \frac{1}{\lambda}g^{-1}\partial_t g \tag{4.16}$$

we will find that the classical equations of motion (following the Euler-Lagrange equations) are equivalent between the theories. Thus, we say that they are classically dual theories. However, at the quantum level, the two diverge. The principal chiral model is known to be asymptotically free, meaning that upon renormalization the running coupling $\lambda \to 0$ at high energies. However, our toy model of interest has essentially the opposite behavior. The quantum theory features a *Landau pole*, meaning that $\lambda \to \infty$ at large but finite energies.

Such theories are not unheard of; pure QED has a Landau pole, which is avoided in the standard model by the emergence of electroweak unification at far lower energy. Additionally, $\lambda\phi^4$ theory, the theory of a solitary Higgs-like scalar field, has such behavior in 4 dimensions. It remains unclear if QFT Lagrangians which lead to a Landau pole can describe a physical theory, and thus whether a Higgs-like field could have a physical realization without its standard-model interactions.

Because the $\lambda \to \infty$ or *strong coupling* limit differs between the theories, we studied first the classical behavior of the field theories in this limit. The theory of $\mathcal{L}_{\text{toy}}$ in particular has the seemingly troubling behavior that the speed of light $c \to 0$ in the strongly coupled limit. Nevertheless we were able to find nonlinear wave-like solutions, which continue to propagate at non-uniform speeds in the $c \to 0$ limit, defying naive expectations of the field's behavior. These solutions were found as part of a general class of solutions following the ansatz

$$\phi(t,x) = e^{Kx} R(t) e^{-Kx} + mKx, \quad K = \frac{i}{2} k \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{4.17}$$

The resulting solutions are analogous (but not equivalent) to plane wave solutions. Here m and k are constants, with wavenumber k. The field ϕ is seen to have constant energy density, shifting by an internal rotation and a constant shift under translation. The formula for the solution $R(t)$ is somewhat involved, and can be found in the paper. The resulting orbit in the $R_1 - R_2$ plane, along with the solution $R_3(t)$, are plotted for a sample solution in Fig. 4.2. We see that this solution rotates much like an epicycloid in the R_1 and R_2 components, while increasing at a non-constant rate in the R_3 direction.

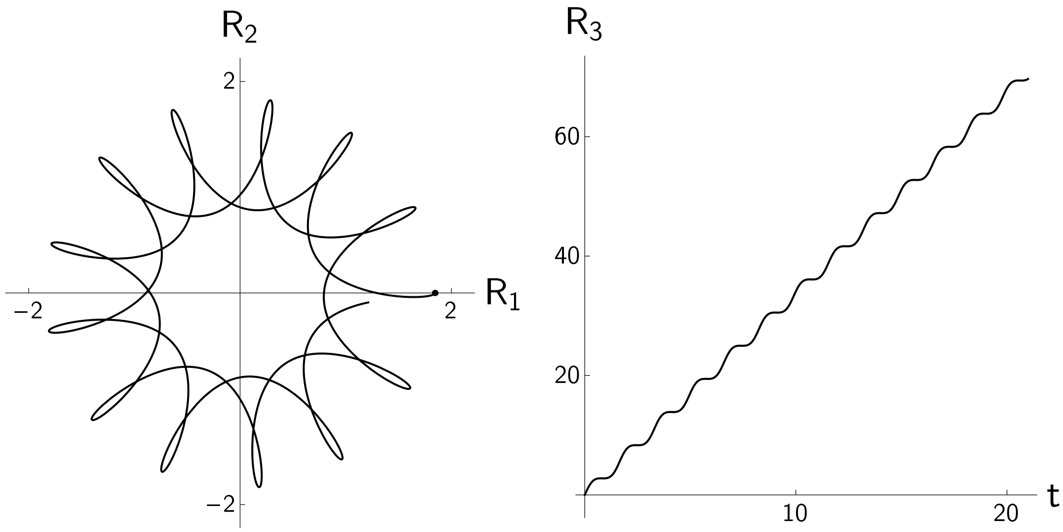

Figure 4.2: The orbit in the R_1ˇR_2 plane (left) and the evolution of R_3 over time (right). This sample solution is plotted for $0 < t < 21$ and uses parameters $k = 1$, $\lambda = 2.4$, and $m = 0.5$. The initial conditions are chosen in terms of conserved quantities of the solution set, and are detailed [32].

We can restrict our field theory to a finite-dimensional system by studying this class of wave solutions exclusively. This reduces the problem to a finite-dimensional mechanical system which can be more easily studied, and the rules of ordinary quantum mechanics can be applied to the reduced system–such an approach is called a *semi-classical* approximation.

In the semi-classical approximation, we found that the reduced system ordinarily had excitations following a dispersion relation of $E \sim |k|$, where E is the energy density and k is analogous to a wavenumber for the system. This is the typical dispersion relation for a field with massless excitations.

However, in the strongly coupled limit of the semi-classical model, we found the unusual dispersion relation $E \sim |k|^{2/3}$, which does not correspond to the spectrum of any normal particle-like excitation. We hypothesize that these nonlinear wave solutions could thus represent some exotic excitations which appear at strong coupling in the full theory.

Such a conjecture is somewhat complementary to what is already well-accepted to occur in Yang-Mills theories like QCD. In that case, easy to understand excitations in the $\lambda \to 0$ limit break down at lower energies where the coupling becomes large. The fundamental excitations of the Lagrangian cease to exist at a certain point, yielding objects that would not be expected from perturbation theory–e.g., quarks and gluons give way to hadrons and mesons. Perhaps the reverse can happen as well–perhaps theories with Landau poles, where the coupling diverges at high energies instead, have some exotic excitations which survive in this limit that physicists have yet to understand.

Following this work, the reduced mechanical system introduced in [32] was further studied by Krishnaswami and Vishnu, who found a Lax pair and a family of action-angle variables for the system in [33, 34]

4.6 Summary and Outlook

There are many fundamental problems that remain in theoretical high energy physics. A common approach has been to add new objects in order to cancel out some difficulties, rather than to examine in detail the minutia of the underlying framework of QFT. For a long time, similar strategies guided the theory to correct conclusions in advance of their experimental discovery, such as with the Higgs model. However, with no exciting experimental results to confirm any grandiose extensions of the standard model, it is becoming harder to believe that they exist as originally formulated. Subjects like supersymmetry and string theory have effectively become matters of faith.

Meanwhile, in the direction of further elucidating our existing theory, problems like the Yang-Mills mass gap conjecture have stood unsolved for many decades. Recent progress is hard to assess, as interest and funding have waned somewhat. There have been many interesting toy models and conjectures over the last 50 years. It is difficult for a newcomer to sift through them, but any one of these old ideas could have a sudden resurgence when viewed by the right perspective. Interesting results in field theory and math can sometimes come seemingly out of nowhere.

In the absence of new experimental guidance for such a long time, theory has arguably run amok. With that said, the physical utility of QFT as a framework is irrefutable, and there is interesting work being done on the mathematical side to better define its axioms and make rigorous its results. Perhaps this is a promising avenue, hindered only by the fact that mathematical physicists and theoretical physicists often inhabit different scientific worlds, separated by dense notational and conceptual barriers.

Chapter 5

Experimental Paths Forward

"Science progresses best when observations force us to alter our preconceptions"

— Vera Rubin

5.1 The Energy Frontier

One obvious path forward is to build higher and higher energy colliders and see what happens. Particle physics research which takes place at this energy frontier is often called *high energy physics*, a rather vague name which aims to set it apart from other types of particle physics experiment which do not rely on enormous colliders. The Large Hadron Collider (LHC), discussed further in Chapter 7, is currently the highest energy collider ever built, and is firmly at the epicenter of experimental high energy physics today.

It is a guarantee that new physics will emerge eventually, but the energy at which this will occur has become very unclear since the discovery of the Higgs boson in 2012. This is a fairly new situation–prior to 2012, there were widely accepted arguments that if the Higgs boson was not found in a certain mass range upon turning on the LHC, some other new physics would have to be found around the same energy energy scale. But in June 2020, there are currently no easily accessible energy ranges which have an obvious need for new physics. Models for extra particles and dark matter candidates often have many free parameters that cause them to become relevant at arbitrary energy scales.

Although discovering new particles is perhaps the most exciting and glamorous way to advance experimental high energy physics, it is not the only sort of result the field has to offer, and seems increasingly unlikely in the near future. The lack of a clear direction for high energy experimental physics pushes us towards an era of precision measurement, in which every aspect of the existing model and data is pored over and vetted extensively.

Previously, precision measurements have pointed the way to new physics well in advance of direct discoveries. For example, the mass of the top quark was predicted in advance of its discovery by precision fitting in the electroweak sector of the standard model [23]. However, particle physics at the energy frontier is today a bit more directionless than in the 1990s, as discussed in Sec. 4.1–we no longer expect to find new heavy quarks or a long-missing scalar boson. We know that something is missing from the existing model, but we can't agree what. Thus, a more broad approach to precision measurements of standard model parameters is a promising path forward, as it allows us to scour the data and see if anything in particular is amiss in our models.

There are many precision measurements being done by experiments at the LHC, but one obvious target is verifying the expected properties of the Higgs boson. This direction, specifically with regards to its interaction with the top quark, will be the final research project presented in this dissertation.

5.2 Other Approaches

Before proceeding into our experimental research, we note briefly for the reader other exciting avenues of physics research which can crack away at the fundamental questions addressed in this document.

Neutrino experiments are making slow but steady progress in illuminating the properties of these often invisible particles. For example, shortly before the writing of this dissertation, the T2K experiment in Japan published a result regarding the constraint of CP violation terms involved in neutrino oscillation [35]. Arguably, neutrinos are the only particles currently seen to disobey the standard model, because they have mass (albeit very little). In the usual formulation of the standard model, they are massless. While adding in neutrino mass does not exactly shatter the standard model into pieces, there is some ambiguity on what mechanism gives rise to it, and whether they may have heavy unobserved counterparts.

Neutrinos are difficult to detect, and generally we must have trillions of neutrinos pass through large liquid-filled detectors to observe even a few. Otherwise, they simply pass right through everything they encounter. There are many interesting experiments which are either collecting data now or will begin in the near future, such as the T2K experiment in Japan and the DUNE collaboration in North America.

These have the potential to revise our understanding of the standard model in the coming decades. In a completely different approach to study neutrinos, there are also experimental searches for neutrino-less double beta decay, which could determine if neutrinos are Majorana fermions and shed light on the problem of neutrino mass. Overall, this sector of research tends to generate results at a reliable, steady pace.

Dark matter searches are experiments which try to detect signs of dark matter. Direct detection experiments can be large tanks of liquid underground, similar to neutrino detectors, hoping to detect the signature of a new particle passing through. There are also indirect detection experiments–for example, those that search for decay products of dark matter annihilation in outer space. Relative to the slow and steady trickle of information from neutrino experiments, this sector is a bit of a lottery. At any time, one of them could detect the signature of a dark matter candidate, completely upending the world of particle physics and surely meriting a Nobel Prize.

Cosmological experiments are closely linked to particle physics as well–the standard model governs the evolution of the early. Thus, measurements of the cosmic microwave background, and perhaps someday the cosmological gravitational wave background, can have profound implications for dark matter models and particle physics in general.

These are just a few to highlight the breadth of experimental approaches that could shed light on the fundamental mysteries of particle physics. There are many others, from searches for magnetic monopoles, to cosmic ray detection, to the generation of stable antimatter bound states, to precision studies of radioactive decay... As the rate of new discoveries from the energy frontier slows, perhaps the next big breakthrough in particle physics will come from a new type of experiment which generates little excitement today.

Chapter 6

Top Quarks and Collider Physics

"The top quark was discovered in 1995 and since then the Higgs has become our obsession, because the standard model was incomplete without it."

— Fabiola Gianotti, Director General at CERN

6.1 Why are Top Quarks Interesting?

The top quark is the heaviest known fundamental particle at the time of this dissertation's writing. It currently has a mass of $172.4 \pm 0.7\,\mathrm{GeV}$, a bit higher than the Higgs boson mass of $125.1 \pm 0.14\,\mathrm{GeV}$ (the latest values taken from combination measurements in Ref. [36]). This makes its mass 40 times higher than that of the next heaviest quark, and over 300,000 times heavier than that of an electron. Its heavy mass means the top quark is highly unstable, making it the only quark to decay

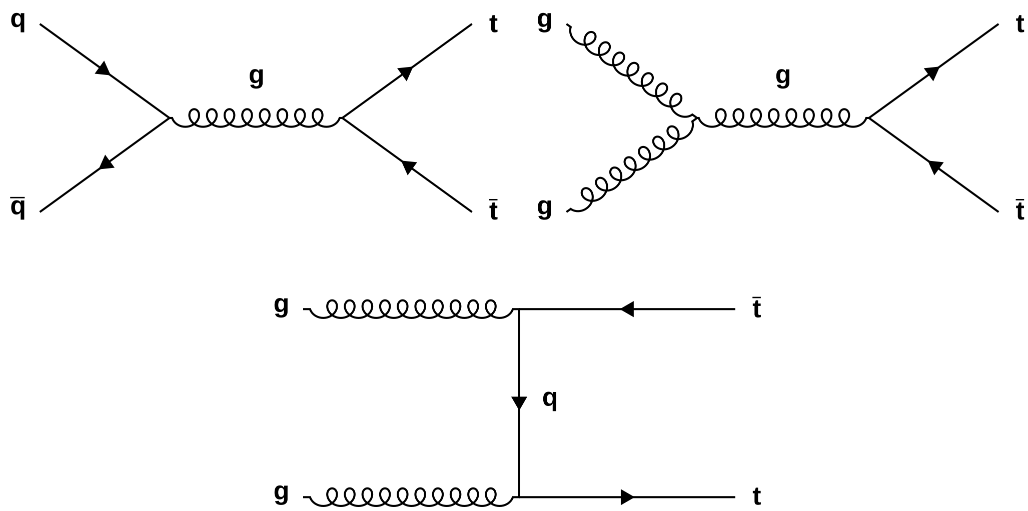

Figure 6.1: Example diagrams for leading order $t\bar{t}$ production via qq and gg processes.

prior to hadronization. This opens up an experimental window into fundamental physics interactions at the highest energies currently attainable, making the top quark one of the most useful tools to advance our understanding of the standard model.

6.2 Top Quark Production at Hadron Colliders

For the purposes of this document, we will focus on top quarks produced as quark anti-quark pairs. Top quark pair production, or $t\bar{t}$ production for short, occurs when a top and anti-top quark are produced together in an inelastic collision. At hadron colliders, this happens mostly due to strong interactions between quarks and gluons, as shown in Fig. 6.1. Pair production is generally the easiest way to observe top quarks, and served as the initial means of discovery for the top quark at the D∅ and CDF experiments in 1995 [37,38].

At colliders, the starting point for QFT calculations is a scattering *cross section* σ. As the name might suggest, this has dimensions of area, and is often expressed in terms of the unit Barns $\mathrm{b} = 10^{-28}\,\mathrm{m}^2$. In classical theory, a scattering cross section is the size of the area perpendicular to two particle's motion within which they must pass to interact and scatter. In quantum theory, it is related to the probability that two initial state particles interact and product a given final state. For example, from 2016–2018 the CMS experiment recorded an integrated luminosity of about $L_{\text{tot}} = 137\,\text{fb}^{-1}$ of collision data. The total cross section for $\mathrm{t\bar{t}}$ production from proton-proton collision at the appropriate energy is calculated to be about $\sigma_{\mathrm{pp}\to\mathrm{t\bar{t}}} = 840\,\mathrm{pb}$ [39]. Multiplying the two, we can estimate that the number of top quark pairs produced and recorded at the CMS detector during this period was about $L_{\text{tot}} \times \sigma_{\mathrm{pp}\to\mathrm{t\bar{t}}} = N_{\mathrm{t\bar{t}}} \approx 10^8$ events.

The basic relationship between Feynman diagrams and cross sections will be important. While we won't calculate any in this dissertation, Feynman diagrams stand in for integrals used to calculate elements in a perturbative expansion of the scattering matrix. Each vertex of a Feynman diagram carries a factor relating to the coupling constant governing the relevant interaction, such as α_{s} for QCD diagrams in Fig. 6.1. The associated *matrix element* $\mathcal{M}_{i\to f}$, for initial state i and final state f, carries a factor of $\sqrt{\alpha_s}$ for each vertex in the diagram. The cross section follows the relationship

$$d\sigma \propto |\ \mathcal{M}_{i\to f}\ |^2, \tag{6.1}$$

and so cross section calculations themselves effectively carry one factor of the coupling α_s for each vertex (though things can get more subtle when summing over initial and final states).

Thus, the diagrams in Fig. 6.1 are said to be of order $\mathcal{O}(\alpha_s^2)$, which is the leading order (LO) $\mathrm{t\bar{t}}$ production. Diagrams of order $\mathcal{O}(\alpha_s^3)$ are then next-to-leading order (NLO) QCD diagrams, and so on. These higher order diagrams can contain either loops of virtual particles, or emissions of additional real particles in the final state.

One can also calculate *differential cross-sections*, to look at the relative probability of final states as a function of some kinematic variable. This will tell you how many events to expect as a function of the desired variable. For example, we will be looking at the differential $\mathrm{t\bar{t}}$ cross section with respect to the invariant mass of the $\mathrm{t\bar{t}}$ system, $d\sigma_{\mathrm{t\bar{t}}}/dM_{\mathrm{t\bar{t}}}$. We show a theoretical result for this differential cross section in Fig. 6.2. We see that pair production begins at threshold mass of $2m_{\mathrm{t}} \approx 345\,\mathrm{GeV}$ and peaks slightly after, trailing off at higher values of $M_{\mathrm{t\bar{t}}}$.

6.3 Top quarks: Decay and Detection

Top pair production has a relatively high cross section and leaves unique byproducts which are easy to distinguish from other processes. When they are produced, top quarks have an incredibly short lifetime of about 10^{-25} seconds. Because of this short lifetime, the top quark is the only quark which decays via a fundamental Feynman vertex modeled by perturbative calculations. When an unstable particle decays, each possible decay process has a probability called a *branching fraction*. To simplify things further, top quarks in the standard model decay almost exclusively to a bottom (b) quark and a W-boson, with a branching fraction of over 99%. The vertex that allows this decay is shown back in Fig. 2.5. We briefly discuss each byproduct of this decay.

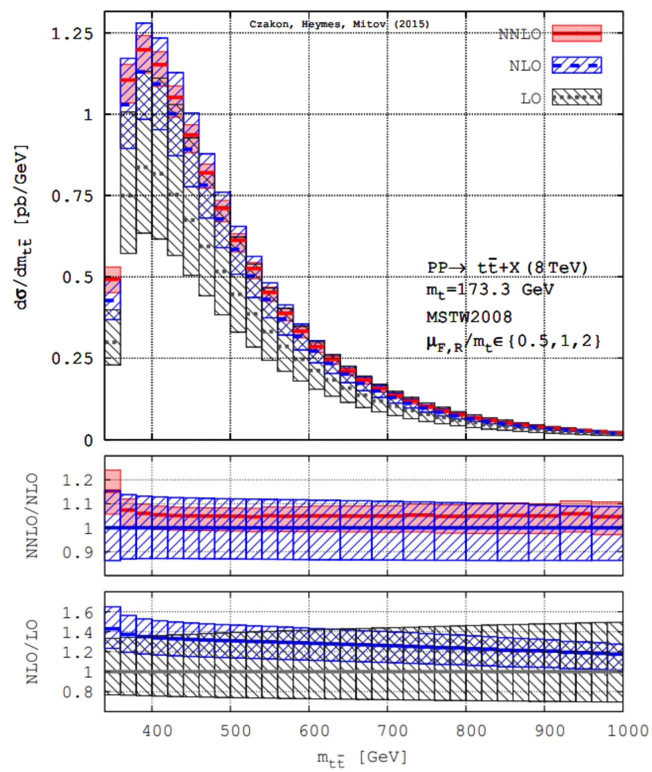

Figure 6.2: Fixed-order theory calculations of $d\sigma_{\mathrm{t\bar{t}}}/dM_{\mathrm{t\bar{t}}}$ are shown at LO, NLO, and NNLO in QCD for proton-proton collisions at $\sqrt{s} = 8\,\mathrm{TeV}$ and a top mass of $m_{\mathrm{t}} = 173.3\,\mathrm{GeV}$, which was a popular choice for simulation until recently. This figure is taken from Ref. [40] by Czakon, Mitov, and Heymes. Uncertainty bands are based on varying the renormalization and factorization scales, μ_{r} and μ_{f}, respectively.

The **b quark** survives long enough to hadronize, generally forming a B hadron. This hadron then typically travels a distance on the order of several millimeters before decaying at a secondary vertex. Like other quarks and unstable particles, the b quark is ultimately detected as a shower of particles called a jet. Luckily, information about particles which appear to come from secondary vertices, together with additional jet kinematics, can be used to distinguish b jets from other jets with reasonable success [41]. Different algorithms and working points are available to identify b jets, which have varying rates of sensitivity (the rate of correct identification fraction of b jets) and specificity (the rate of mis-identification of other jets). $t\bar{t}$ decays benefit from having two associated b-jets, which helps greatly in distinguishing them from other physics processes.

The **W boson** has two options for decay. First, it can decay *hadronically* into a quark anti-quark pair, which both hadronize and are in practice detected as one or two jets. Secondly, it can decay *leptonically* into a lepton and neutrino pair. In this case, the W boson decays into an electron (e), muon (μ), or tau lepton (τ), and the corresponding neutrino, with equal probability (see the vertex from Fig. 2.5). If the lepton is an electron or muon, it can be detected and reconstructed with high quality. Tau leptons decay quickly and can sometimes be reconstructed by their decay products, though many measurements try to avoid events with taus due to the added challenge. Meanwhile, the neutrino (ν) escapes undetected, leading to an imbalance in transverse momentum (p_{T}) with respect to the beam axis which contributes to $\boldsymbol{p}_{\mathrm{T}}^{\mathrm{miss}}$. While this presents a challenge in reconstructing the neutrino momentum, it also offers another way of distinguishing $t\bar{t}$ events, as they will generally have a large value for $p_{\mathrm{T}}^{\mathrm{miss}}$.

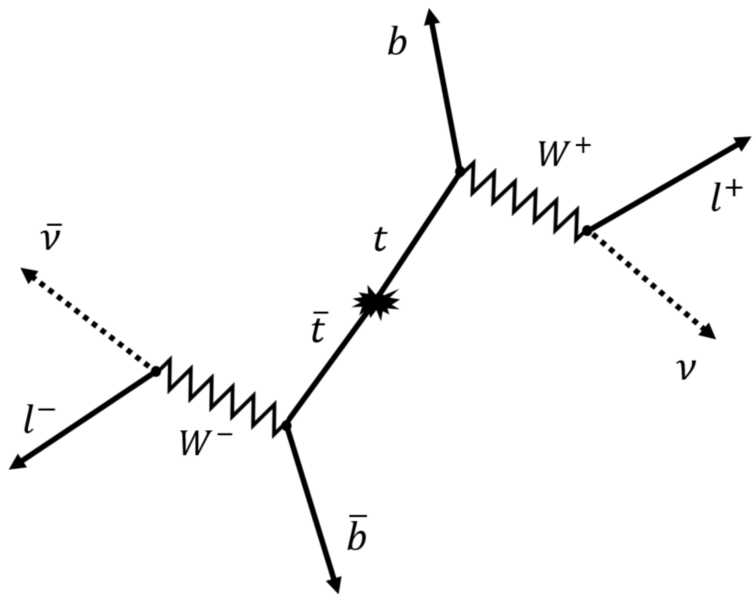

Figure 6.3: A $t\bar{t}$ pair decaying into a dilepton final state

The process of identifying $t\bar{t}$ decays is helped by the fact that these decay products will generally have high p_T relative to decays of many other physics processes. The channel with the lowest background rate is the dilepton final state, where both W bosons decay leptonically. The presence of two b jets, two high p_T leptons, and large p_T^{miss} makes it easy to distinguish these events and obtain samples of high purity. However, it does have lower statistics than the other decay channels, with a branching fraction of only 11% in $t\bar{t}$ decays.

Because of the high purity and relatively high cross section, $t\bar{t}$ events offer a unique ability to measure differential cross sections with high precision. We will see in chapter 8 how the shape of such measurements can be used to study the interaction of top quarks with other particles.

6.4 Kinematics of dilepton event reconstruction

The process of two top quarks ($t\bar{t}$) each decaying leptonically is visualized in Fig. 6.3. The two undetected final state neutrinos create a challenge in reconstructing the top quark kinematics. However, if one forces the decays to be strictly on-shell, meaning that 4-momentum is exactly conserved at each vertex and the mass width of t and W are ignored, the neutrino momenta can be in theory determined exactly. This approximation is not exactly valid–recall that intermediate particles in $t\bar{t}$ decays are technically virtual. The precise physicality of such particles is a rather philosophical debate, but what is clear is that they obey conservation of momentum and energy only in approximation.

If we make this approximation for each top quark decay, conservation of energy E and momentum $\vec{p}$ at each vertex yield

$$E_{\mathrm{t}} = E_{\mathrm{b}} + E_{\mathrm{W}} = E_{\mathrm{b}} + E_{\ell} + E_{\nu}, \tag{6.2}$$

$$\boldsymbol{p}_{\mathrm{t}} = \boldsymbol{p}_{\mathrm{b}} + \boldsymbol{p}_{\mathrm{W}} = \boldsymbol{p}_{\mathrm{b}} + \boldsymbol{p}_{\ell} + \boldsymbol{p}_{\nu}. \tag{6.3}$$

The momenta $\boldsymbol{p}_{\mathrm{b}}$ and $\boldsymbol{p}_{\ell}$ and the corresponding energies are known from measurement. The masses of the top quark and W boson are taken as fixed, entering through the quadratic relationship

$$E^2 = |\ \boldsymbol{p}\ |^2 + m^2. \tag{6.4}$$

With the neutrino's mass taken to be zero, each vertex of the decay generates a set of quadratic equations which constrain $\boldsymbol{p}_{\nu}$ to an ellipsoid in 3-dimensional momentum space (6.4 left). The neutrino momentum is thus constrained to lie on

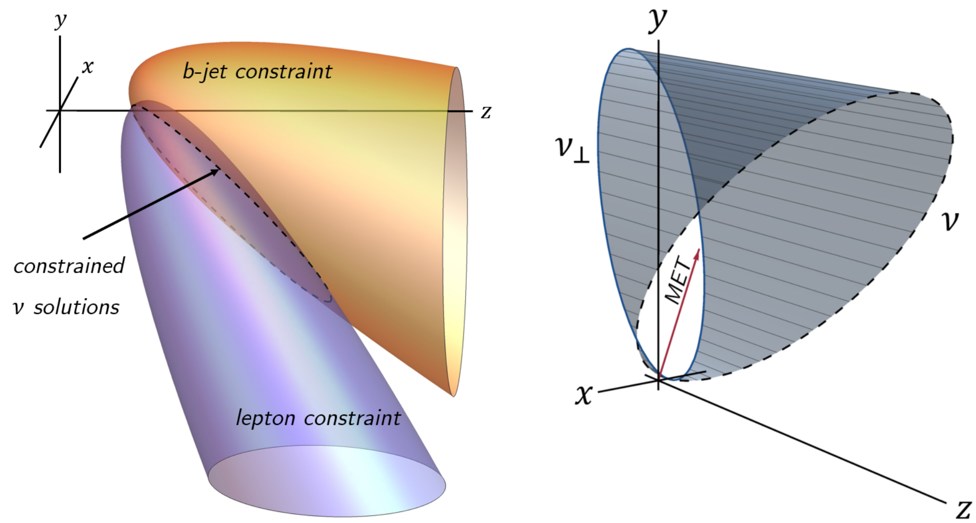

Figure 6.4: Left: ellipsoidal constraints on neutrino momentum, whose intersection defines an ellipse. Right: The elliptic constraint projected into the transverse plane.

the intersection of these two surfaces, which traces out an ellipse. We call this the *mass constraint*, and it is not always possible to satisfy, as the two ellipsoids will not generically intersect. Occasionally, this is due to small deviations from the on-shell assumption (recalling that intermediate particles are virtual), or from measurement resolution. More commonly in practice, it occurs in the case of an incorrect event reconstruction where a jet has been mis-identified or placed at the wrong vertex.

If the mass constraint can be formed, this ellipse can then be projected into the transverse plane to define a constraining curve $\boldsymbol{\nu}_\perp$, which can be related to the measured quantity $\boldsymbol{p}_\mathrm{T}^\mathrm{miss}$ (also known as the MET vector). In the case of semi-leptonic decays, with only one undetected neutrino, we can simply introduce a parametrization $\boldsymbol{\nu}_\perp(\theta) : [0, 2\pi] \mapsto \mathbb{R}^2$ and find the point $\boldsymbol{\nu}_\perp(\theta)$ which minimizes the distance to $\boldsymbol{p}_\mathrm{T}^\mathrm{miss}$.

In dilepton decays, with two final state neutrinos carrying away undetected momentum, we must repeat this argument twice. This results in two separate

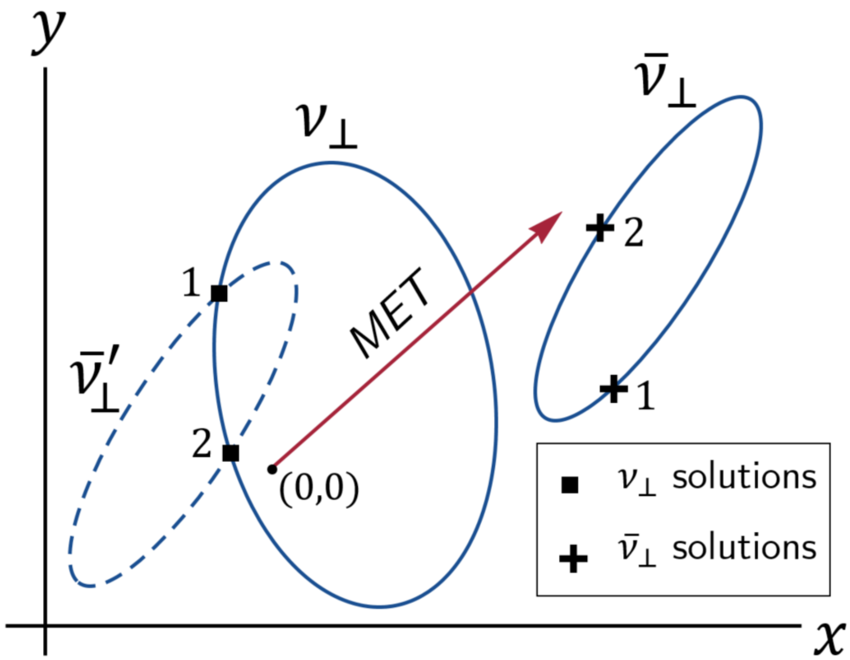

Figure 6.5: combination of 2 ellipses in the dilepton case

transverse momentum constraints, $\boldsymbol{\nu}_\perp$ and $\bar{\boldsymbol{\nu}}_\perp$ (Fig. 6.5). We wish to find points on these two ellipses, specified by parameters θ_1 and θ_2, such that

$$\boldsymbol{\nu}_\perp(\theta_1) + \bar{\boldsymbol{\nu}}_\perp(\theta_2) = \boldsymbol{p}_{\mathrm{T}}^{\mathrm{miss}} \tag{6.5}$$

This can be rewritten as

$$\boldsymbol{\nu}_\perp(\theta_1) = -\bar{\boldsymbol{\nu}}'_\perp(\theta_2) \tag{6.6}$$

where $\bar{\boldsymbol{\nu}}'_\perp(\theta_2) \equiv \boldsymbol{p}_{\mathrm{T}}^{\mathrm{miss}} - \bar{\boldsymbol{\nu}}_\perp(\theta_2)$.

This new ellipse maintains the overall shape of $\bar{\boldsymbol{\nu}}_\perp$, but is shifted and rotated 180° about its center. The problem then reduces to the intersection of two ellipses. We call this the *MET constraint*. The MET constraint can be satisfied if the two

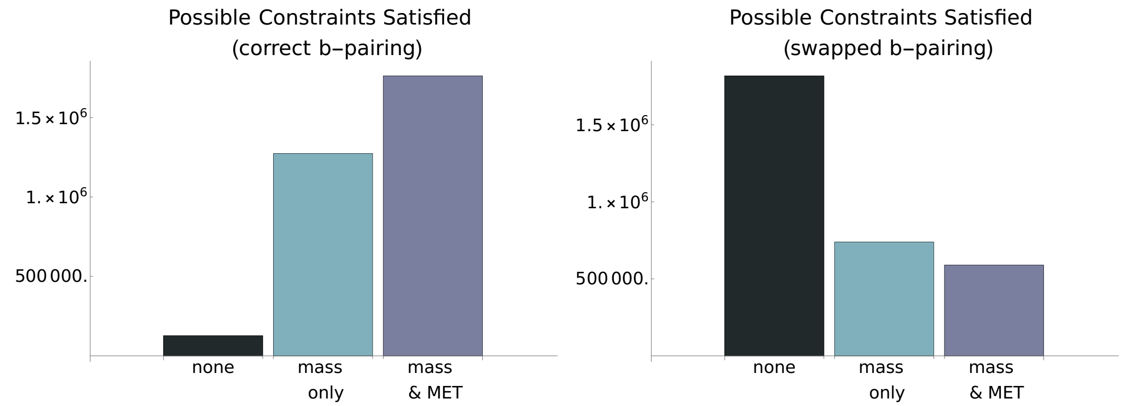

Figure 6.6: We show the frequency with which the different kinematic constraints can be satisfied, comparing the case where the two b jets are paired with the correct leptons against the case where they are not.

ellipses do intersect, which will yield a discrete set of either 2 or 4 possible momenta for the neutrinos. Because the MET is a difficult quantity to measure in practice, it is not infrequent that the MET constraint cannot be exactly satisfied even if the other objects in the event were correctly identified and well reconstructed. In this case, an approximate solution can be taken from the points of closest approach of the two ellipses.

We show in Fig. 6.6 the results of examining the two kinematic constraints when comparing correct and incorrect matching of b jets to leptons in simulated events. This emphasizes the role of these kinematic constraints as a tool to aid with proper selection and assignment of b-jets. The reconstruction of $t\bar{t}$ events from detected particles can be a non-trivial procedure, particularly in the dilepton channel. While $t\bar{t}$ events are easy to identify in this final state which has very low background rates, it remains difficult to accurately reconstruct the kinematics of the top quarks themselves due to the undetected neutrinos. The reconstruction technique outlined here is very sensitive to detector resolution, and deviations from on-shell behavior of the decays. The kinematic variables used for the measurement presented in Chapter 8 ultimately do not rely on a full kinematic reconstruction of the top quarks, but the kinematic constraints outlined here are still used indirectly to aid in determining which b-jets and leptons are likely to originate from the same top quark.

Chapter 7

The CMS Experiment

7.1 The Large Hadron Collider (LHC) apparatus

The LHC [42] is the largest and highest-energy particle collider ever constructed, spanning across the border of France and Switzerland near the city of Geneva. The main ring, which uses the tunnel from the Large Electron Positron collider (active 1989-2000), has a diameter of 8.6 km and ranges from 50-175m underground along its circumference. Its primary function is to store and collide circulating beams of protons at energies of up to 14 TeV, though it is also used to collide heavy nuclei during limited data-taking periods. Particle energies are ramped up with a series of smaller linear and circular accelerators before entering the main ring. In this document, we will discuss the machine from the proton-proton collision perspective; for more on the physics resulting from heavy ion collisions at the LHC, see for example the review [43].

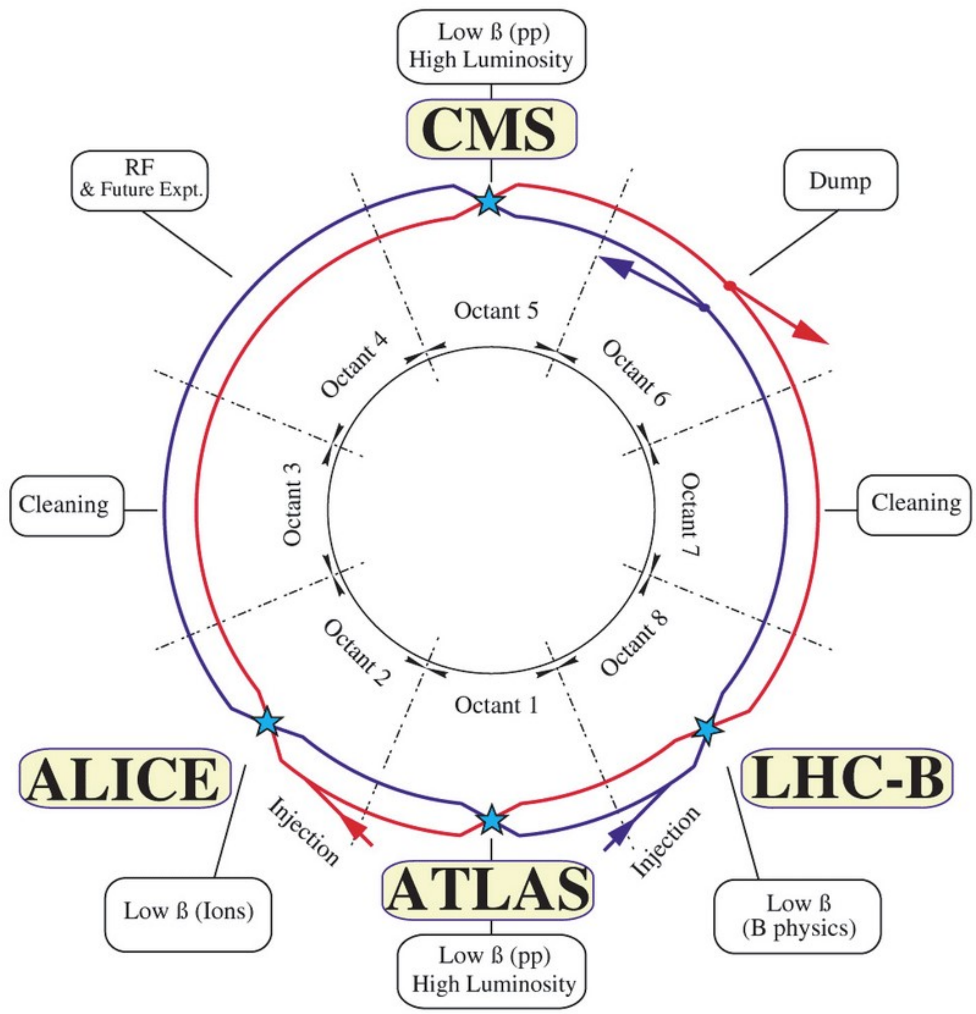

Figure 7.1: A map showing the layout of LHC. The primary experiment at each of the 4 bunch crossings is labeled, but there are also smaller experiments associated with these points.

Protons circulating in the LHC ring are divided into bunches prior to injection. Each bunch contains around 10^{11} protons and is about 30cm long (from the laboratory reference frame). The main ring has over 2000 bunches circulating simultaneously when running at its current beam capacity.

There are several experiments designed to detect results of collisions at 4 beam crossing points, shown in Fig. 7.1. Of these detector experiments, the two largest are the CMS and ATLAS experiments, both of which are general purpose detectors aimed at measuring and reconstructing a broad range of collision byproducts and outcomes. CMS is an acronym for *Compact Muon Solenoid*, while ATLAS supposedly results somehow from the words *A Toroidal LHC Apparatus* (though many scientists believe that this shortening violates the fundamental rules of an acronym).

The LHC was first turned on in 2008. The first data-taking run for experiments took place from 2010-2013, where collisions took place at an center of mass energy of 7 to 8 TeV. After a long shutdown for upgrades and maintenance, the beam energy was increased to 13 TeV for the second run of data taking, which took place from 2015-2018. In particle physics, the center of mass energy (total energy of the colliding protons in the rest frame) is denoted by $\sqrt{s}$, due to its relationship to the traditional Mandelstam variable s. The data-taking period from 2016-2018 was the second major data-taking effort at the LHC, and this period is referred to collectively as Run 2. Run 3 is tentatively scheduled to begin in 2021, after which the machine will begin upgrading to the High-Luminosity LHC (HL-LHC), planned to begin operating around 2030. This machine, while operating at the same energies, will produce vastly greater luminosities.

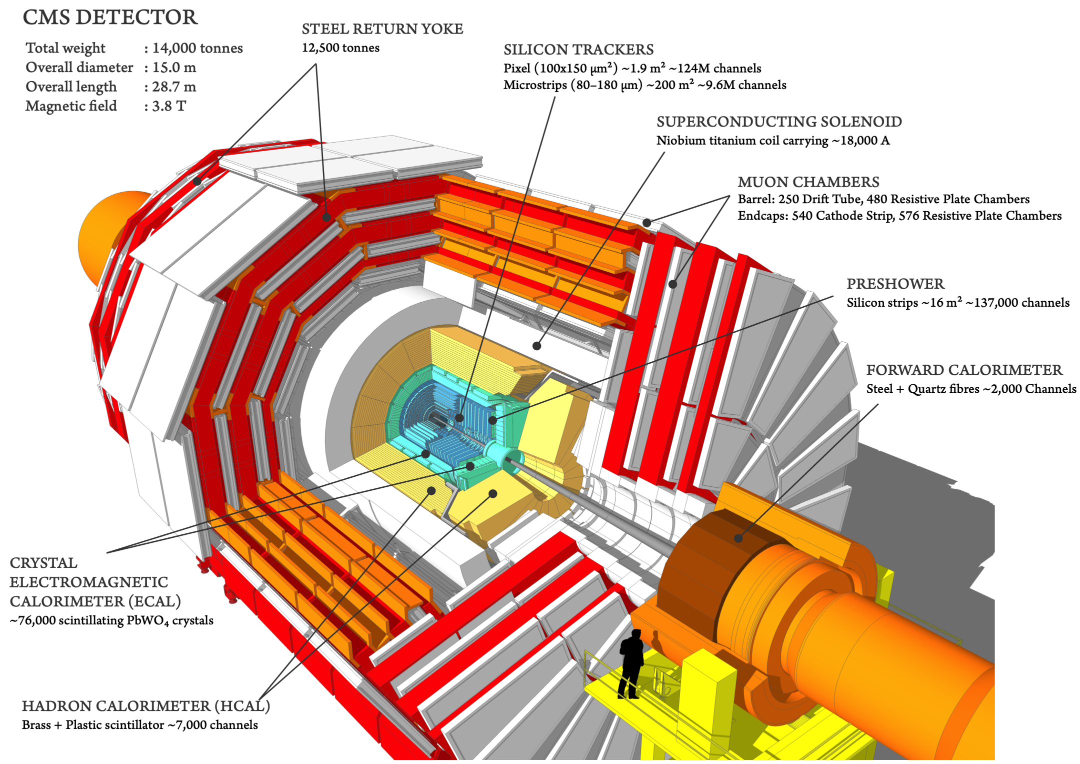

Figure 7.2: a 3-dimensional rendering of the CMS detector is shown, indicating the scale of the detector and the overall layout of the internal components with respect to the beam pipe.

7.2 The CMS detector

The detector has 5 main layers, consisting of 4 types of detectors and the solenoid magnet which is the experiment's namesake, as shown in Figs. 7.2-7.3. We briefly outline the detector components here, from innermost to outermost.

1. The **tracker** is comprised of an inner portion with high-granularity silicon pixels and an outer level of silicon strips. A particle passing through the silicon material will create loose electron/hole pairs. A reverse bias voltage generates an electric field which prevents these pairs from recombining, and

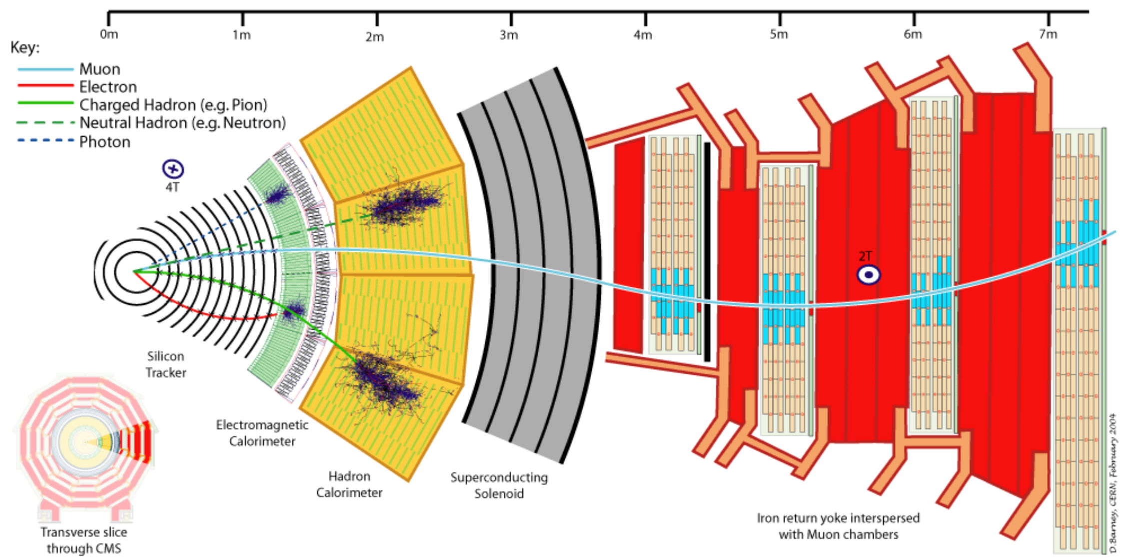

Figure 7.3: A cross sectional slice of the CMS detector is shown with example particle tracks demonstrating the function of each detector layer

they drift through the bulk of the material to readout sensors where the effect can be detected.

With the many layers of silicon pixels and strips, we can reconstruct precisely particle trajectories. Due to the presence of a strong magnetic field, these trajectories will be curved, and the charge-to-mass ratio of particles can be estimated. The silicon tracker is essential for estimating the location of the vertex at which particles are produced, which may be displaced from the primary vertex of a collision. This is essential in b-jet identification, for example.

2. The **electromagnetic calorimeter** (ECAL) consists of a dense material which detects electrons and photons, causing them to decay into a well-defined burst/shower of particles. The energy deposited by this burst can be measured and used to estimate the energy of detected particles. As the electromagnetic interaction scales like $1/m^2$ for particles with mass m, heavier particles

like hadrons/mesons and muons can pass through this layer with minimal disturbance.

3. The **hadronic calorimeter** (HCAL) is another calorimeter designed to detect hadrons/mesons. It can detect charged and neutral particles of this type. Among standard model particles, only muons and neutrinos can easily pass through this layer undisturbed.

4. The **superconducting solenoid** (magnet) is the largest magnet of its kind ever constructed, producing magnetic field strengths of up to 3.8 T. This strong magnetic field causes the paths of charged particles to curve based on their mass and momentum, aiding greatly in particle identification and reconstruction. This is arguably the central design element of the detector and its namesake.

5. The **muon detector** detects only muons, and also contains several layers of iron return yoke to ensure that any particles other than muons are blocked before entering this region (excluding neutrinos). While the muons pass easily through all solid components of CMS, the muon detector primarily uses small gas chambers to detect them. The muons will knock electrons off of particles in the gas, creating a detectable signal. Though the spatial resolution of this type of detector is worse than others, the muons are detected at several locations and are the most easily identified particle, making them the cleanest data points in most collisions. The magnet allows their velocity to be easily computed by the observed radius of curvature of their flight path.

The particle-flow (PF) algorithm [44] aims to reconstruct and identify each individual particle in an event, with an optimized combination of information from the various elements of the CMS detector. The energy of photons is obtained from the ECAL measurement. The energy of electrons is determined from a combination of the electron momentum at the primary interaction vertex as determined by the tracker, the energy of the corresponding ECAL cluster, and the energy sum of all bremsstrahlung photons spatially compatible with originating from the electron track. The energy of muons is obtained from the curvature of the corresponding track. The energy of charged hadrons is determined from a combination of their momentum measured in the tracker and the matching ECAL and HCAL energy deposits, corrected for zero-suppression effects and for the response function of the calorimeters to hadronic showers. Finally, the energy of neutral hadrons is obtained from the corresponding corrected ECAL and HCAL energies. A more detailed description of the CMS detector can be found in [45].

The standard coordinate frame uses the z-axis as the beam axis, while the x-axis points inwards towards the center of the ring and the y-axis points vertically upward. The $x - y$ plane is referred to as the transverse plane, and projections of momenta p into this plane are labeled p_{T}. Conservation of momentum dictates that a collision should produce no net momentum in this plane.

The pseudo-rapidity η is frequently used in discussing particle kinematics, as it is closely related to the polar angle θ measured from the beam axis by the formula $\eta = -\ln(\tan(\theta/2))$. Notably, the tracker covers the angular range $\mid \eta \mid < 2.5$ (Fig. 7.4), while the muon detector covers $\mid \eta \mid < 2.4$ (Fig. 7.5). Though some specialized forward calorimeters allow some detection of particles up to $\mid \eta \mid < 5$, high-quality reconstruction is generally restricted to objects within the tracker/muon detector coverage.

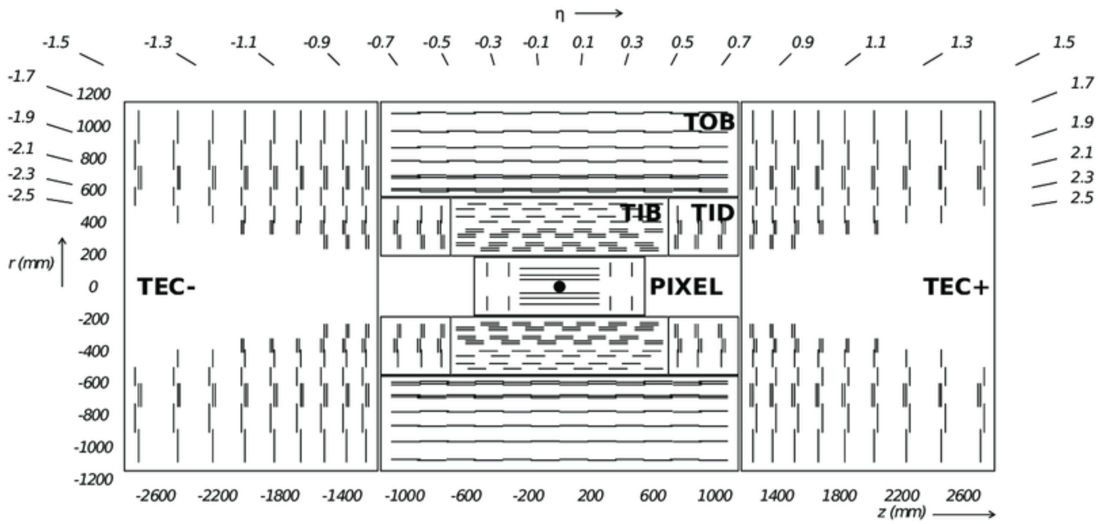

Figure 7.4: A longitudinal slice of the silicon tracker showing the coverage in η. TIB, TID, TOB, and TEC are sections of the strip tracker.

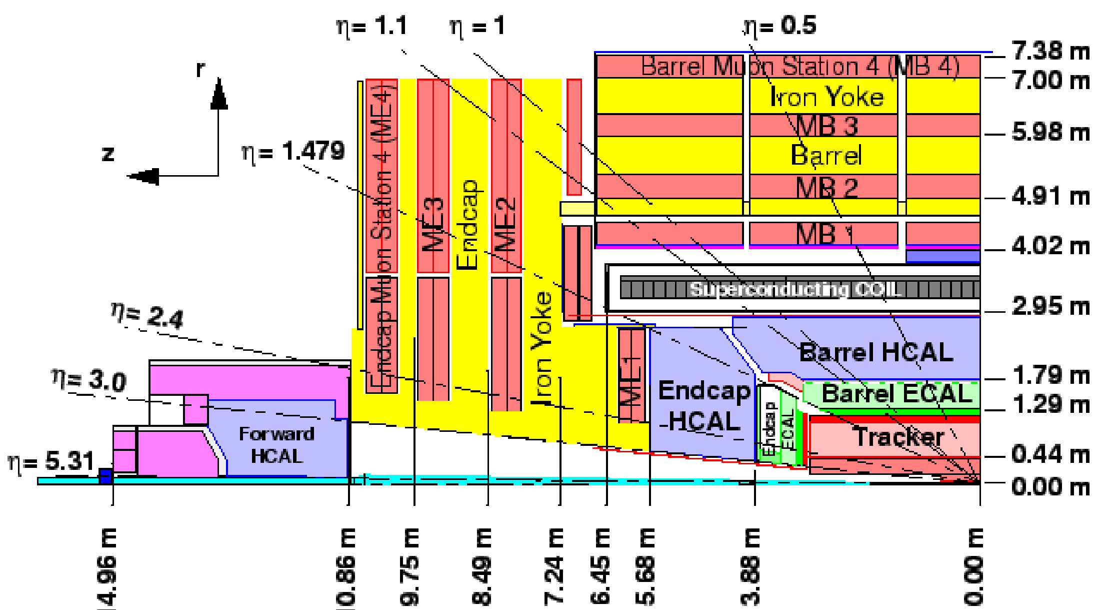

Figure 7.5: A cross sectional slice of the full detector showing the coverage in η of the muon chambers as well as other components.

7.3 Collision basics

Though the LHC generates proton-proton collisions, protons are not the most fundamental participants in the inelastic collisions that we aim to study. The collisions of most interest are those in which pieces of protons break apart and form a variety of heavy and unstable particles. Recall that protons are a bound state of 3 quarks, called the *valence quarks* quarks. Due to the strongly coupled nature of QCD interactions, this bound state is less of an inanimate collection of 3 stable ingredients, and more of a living organism whose innards are constantly interacting and rearranging. In addition to the three valence quarks, these protons contain what is often called a "sea" of non-valence quarks, anti-quarks, and gluons. When two protons collide at a certain energy, we do not know a priori which fundamental particles start the interaction, or what fraction of a proton's energy those fundamental constituents contain. Instead, this can only be known probabilistically, meaning a large amount of data or *luminosity* is needed for most precise observations.

The collision of two proton bunches is referred to as a bunch crossing, and each of these results in dozens of interactions along the beam axis within the CMS detector. Due in part to the wide range of possible energies of the fundamental participants, the vast majority of interactions which occur in any bunch crossing are physically uninteresting. Some are elastic collisions which do not break apart the protons, and others are simply inelastic collisions that do not aid in making new physics measurements. In order to filter out these events from the enormous amount of data that is produced during each second of data collection, a multi-tier trigger system selects interesting events on the fly, preventing a wildly impractical amount of data from being stored permanently. For the purpose of this thesis, we will only refer explicitly to the high-level triggers, which are used to compile different datasets used for physics analysis as in Chapter 8. The potential for overlap between individual interactions taking place very close to each other creates an issue called pileup. Generally speaking, this will involve a small number of physically uninteresting events whose byproducts sneak into the data, an effect that is accounted for in simulation.

When an event is reconstructed, the negative vector sum of all transverse momenta is designated as $p_{\mathrm{T}}^{\mathrm{miss}}$ and is for historical reasons often referred to as the MET or *missing transverse energy*. This vector, in principle, corresponds to momentum that was carried off from an interaction without being detected. While there will always be some small transverse momentum imbalance due to detector effects, this quantity can be quite large if undetectable particles such as neutrinos are produced in the event.

7.4 Simulation of CMS events

Probabilities of different proton constituents interacting, and the energy they carry, are determined from *parton distribution functions*. These have been derived and tuned from many experiments over time, and multiple sets are available.

Starting from the fundamental constituents of quarks and gluons, individual collision interactions or *events* are simulated in a series of stages. First, a fixed order perturbative QFT calculation is performed to evaluate the relevant matrix element. While this process models the fundamental interaction taking place at short scales, a variety of QCD effects beyond perturbative hard scattering must be taken into account in order to make the simulation more realistic.

The first of these is emission of lower energy particles, such as gluons and photons. Because of the large value of the strong coupling constant α_s, there is a strong tendency for the high-energy participants in an interaction to radiate away some energy as a shower of lower-energy particles. Luckily, there are some mathematical methods in QCD that allow one to calculate series of such emissions to partially model this effect. The second is the non-perturbative effect of hadronization which affects quarks and gluons. These particles can travel only very short distances before they form more stable bound-state particles, like exotic hadrons and mesons, which can go on to interact with the detector. Both showering and hadronization effects are modeled by *parton shower* algorithms, which take a hard interaction involving a small number of fundamental particles and turn it into myriad of more stable objects capable of reaching the detector. The matching of shower simulations to matrix element computations is nontrivial, and multiple approaches are available.

Lastly, the detector response to all simulated events is modeled with the GEANT4 software toolkit [46], which simulates the interaction of particles with a variety of materials. While the previous steps can be run by individuals with access to reasonable computing resources, this step is very computationally expensive. For this reason, only essential samples are generated including full detector response, and we sometimes have to make due with having lower statistics in some samples than what is desirable.

Event simulation is typically done individually for specific processes and decay channels. However, the potential effects of generic physics events with overlapping primary vertices are later added into the simulation by including extra pileup events.

Chapter 8

A New Measurement of the Top Quark Yukawa Coupling

8.1 Basic Motivation

The top quark mass is well measured at $172.4 \pm 0.7\,\text{GeV}$. Recalling the mass-Yukawa relationship from Sec. 2.8, we see that the top-Higgs Yukawa coupling, g_t, is within a few percent of $g_t = 1$. Given that the value is completely arbitrary in the standard model, one might consider it a bit suspicious–especially in relation to other fermion Yukawa coupling strengths, which are orders of magnitude lower. It is up to experimentalists to determine that the standard model mass-Yukawa relation resulting from coupling to the Higgs field holds as expected. To wuantify any deviation from the standard model, we define the dimensionless *relative* Yukawa parameter $Y_\mathrm{t} = g_\mathrm{t}^{\mathrm{obs}}/g_\mathrm{t}^{\mathrm{SM}}$, where $g_{\mathrm{t,\,obs}}$ is the observed Yukawa coupling and $g_{\mathrm{t,\,SM}}$ is the standard model value. Some extensions of the standard model, such as two-

Higgs-doublet and composite Higgs models, can induce altered couplings between the top quark and the Higgs boson where $Y_t \neq 1$ [47,48]. This makes the top–Higgs interaction in particular an interesting quantity to measure at the LHC, especially because it is experimentally accessible through multiple avenues.

The must obvious way to measure this is through processes containing a top-Higgs vertex which produce a real Higgs boson, as was recently done in Refs. [49,50]. Indeed, the most precise measurement of the top quark Yukawa coupling comes from a combination of such measurements in Ref. [51], which yields $Y_t = 0.98 \pm 0.14$. However, this is not the full story. These measurements detect the Higgs via its decay products, and so they inevitably fold in additional assumptions about its branching fractions in the standard model, which depend on other Yukawa couplings. This is a typical problem; the standard model is so rich and interconnected, it is often difficult to isolate a single parameter for measurement. Thus, it is valuable to measure the same quantity using different techniques. In this section, we outline a measurement that places limits on Y_t using only information from $\mathrm{t\bar{t}}$ decays, where the Higgs boson can be exchanged virtually and does not decay. While this measurement has its own drawbacks, it has the advantage that it does not require us to model Higgs decay, avoiding dependence on other coupling constants.

8.2 Electroweak Corrections and $\mathrm{t\bar{t}}$ Production

Top pair production at CMS is simulated based on matrix element calculations performed at order $\mathcal{O}(\alpha_s^3)$ (as discussed further in Sec. 8.3). This includes only effects from QCD interactions, which are by far the dominant means of production and generally much stronger than other interactions at high energy scales. Electroweak

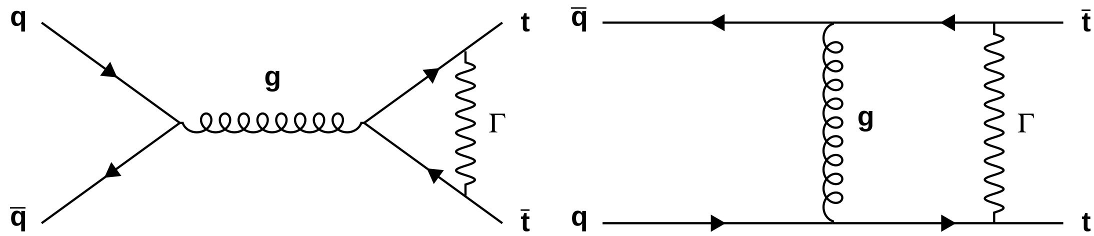

Figure 8.1: $t\bar{t}$ production from qq processes with virtual exchange of a vector or scalar boson Γ, such as the Higgs

(EW) contributions begin to have a noticeable effect at order $\mathcal{O}(\alpha_s^2\alpha)$. Here we note that α stands in for the suite of EW interactions, as the weak and electromagnetic interaction couplings are related and of the same magnitude. Additionally, here we treat equally at this order diagrams with EW interactions that carry a factor of the Yukawa coupling Y_t rather than an explicit factor of α to be of this order. This allows us to consider one-loop diagrams such as those in Fig.8.1 which include Higgs exchange. This process of virtual exchange allows us to isolate the Yukawa coupling g_t from other Higgs couplings.

The leading correction resulting from these diagrams is obtained by including them only as interference terms, hence the lone factor of α–the one-loop term diagram interferes with the LO diagram, while its square is ignored in the matrix-element calculation.

We refer to the diagrams in Fig.8.1 as *EW corrections* to $t\bar{t}$ production. Their effect on the total cross section is small, so they are typically ignored. However, if the Yukawa coupling entering these corrections via Higgs exchange is anomalously large, these diagrams can have a measurable effect on differential cross sections. For example, doubling the value of g_t can alter the $M_{t\bar{t}}$ distribution by about 10% near the production threshold as described in [52] and shown in Fig. 8.2 (left). Another

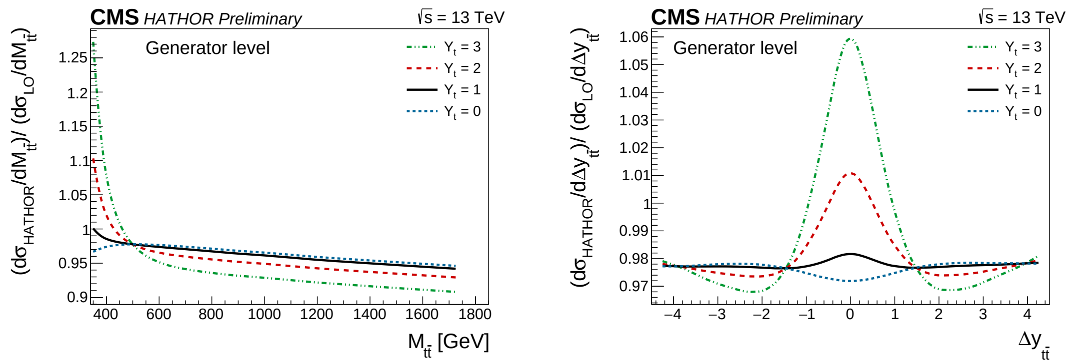

Figure 8.2: Effect of the weak corrections on $\mathrm{t\bar{t}}$ differential cross sections for different values of Y_{t}, after reweighting of generator-level events from POWHEG +PYTHIA8.

variable sensitive to the value of g_{t} is the difference in rapidity $y = \frac{1}{2}\ln\frac{E+p_z}{E-p_z}$ between top and anti top quarks. Defining $\Delta y_{\mathrm{t\bar{t}}} = y(\mathrm{t}) - y(\bar{\mathrm{t}})$, the effect of the EW corrections is shown in Fig. 8.2 (right). Together, these two variables serve as proxies for the traditional Mandelstam variables s and t, and can be used to include the EW corrections on arbitrary samples via simulated event reweighting without generating new MC samples.

This is achieved by applying weight corrections to the current standard $\mathrm{t\bar{t}}$ MC samples used by the CMS TOP group. Contributions to the top quark pair production due to the interference of QCD+EW diagrams, shown in Fig. 8.1, with the leading order (LO) QCD diagrams are evaluated using HATHOR (v2.1) package [53]. The ratio of these contributions to the LO QCD production is applied as parton-level weight corrections to events simulated using POWHEG+PYTHIA8, which incorporates fully NLO QCD calculations and parton shower effects.

8.3 Simulation of $t\bar{t}$ Events at CMS

Here we provide some technical details on $t\bar{t}$ Monte Carlo samples used by CMS during Run 2 data-taking. Simulation of top quark pair production at CMS begins with a NLO QCD matrix-element computation, using the POWHEG generator [54–57]. This calculation is performed with the renormalization μ_r and factorization scale μ_f set to the transverse top mass, $m_T = \sqrt{m_t^2 + {p_T}^2}$. A top quark mass of $m_t = 172.5\,\text{GeV}$ is used in all simulations. The matrix element calculations obtained from POWHEG are then matched with parton shower simulation from PYTHIA8 [58–60]. Pythia uses the underlying event tune M2T4 [61] in 2016 and the CP5 tune in 2017 and 2018 [62]. The parton distribution functions from NNPDF3.0 at NLO [63] are used in 2016 and updated to NNPDF3.1 [64] at next-to-NLO (NNLO) in 2017 and 2018. In the near future, samples will be available which update the simulation in 2016 to match the other years, but these are not available at the time of this document's writing.

We note here a few fundamental sources of uncertainty in the modeling. First, the scales μ_r and μ_f are varied up and down by a factor of 2 to obtain a rough model of uncertainty coming from higher order QCD terms in the matrix element calculation. This is a standard practice in assessing uncertainty on perturbative QFT calculations. The choice of scales μ_r and μ_f is required by the renormalization procedure for perturbative calculations, but the dependence of the result on these scales should vanish as we go to higher and higher orders in perturbation theory. Thus, differences in simulation at different scale choices give a sense of the sensitivity of the matrix-element calculation to higher-order effects. The variation by factor of two is rather arbitrary, but frequently performs well and has become a rule of thumb in the field.

The scale is also varied in the parton shower algorithm, performed separately to assess both the effects of scale choice on initial state radiation (ISR) and final state radiation (FSR). Finally, the makers of the NNPDF sets provide variations which model the uncertainty on the underlying parton distribution functions.

8.4 Formalism of Weak Corrections

The standard jargon of perturbation theory can sometimes be confusing, when applied to the EW corrections, since perturbative expansions are done in orders multiple separate coupling constants, principally α_s and α. Then it is not immediately clear what is meant when someone says "NLO EW diagrams," for example. In the author's experience, this designation always includes terms of order $\mathcal{O}(\alpha_s^2\alpha)$, frequently includes terms of order $\mathcal{O}(\alpha_s^3\alpha)$, and sporadically includes the diagrams for Higgs exchange which are technically $\mathcal{O}(\alpha_s^2 g_{\mathrm{t}}^2)$. We obviously want to include the third type, so we treat top-Higgs vertices as contributing an order α for notational simplicity, as they are part of the full electroweak interaction and the magnitude of their effect is comparable to diagrams involving W and Z bosons. To clarify further, we will utilize notation similar to [65], beginning with an observable such as $\sigma^{t\bar{t}}$ is expanded in a standard series as

$$\sigma^{t\bar{t}}(\alpha_s, \alpha) = \sum_{m+n\geq 2} \alpha_s^m \alpha^n \sigma_{m,n}. \tag{8.1}$$

In practice, contributions due to terms with $m < 2$ or $n > 2$ are small, and as such often neglected in the calculation of the differential cross sections. In this analysis the following contributions are of interest:

$$\sigma_{\text{LO QCD}} = \alpha_s^2 \sigma_{2,0} \tag{8.2}$$

$$\sigma_{\text{NLO QCD}} = \alpha_s^3 \sigma_{3,0} \tag{8.3}$$

$$\sigma_{\text{NLO EW}} = \alpha_s^2 \alpha \sigma_{2,1} \tag{8.4}$$

$$\sigma_{\text{extra EW}} = \alpha_s^3 \alpha \sigma_{3,1} \tag{8.5}$$

In using the Hathor package to compute the differential cross section, as a function of Y_{t}, we include the terms

$$\sigma_{\text{HATHOR}}(Y_{\text{t}}) = \sigma_{\text{LO QCD}} + \sigma_{\text{NLO EW}}(Y_{\text{t}}). \tag{8.6}$$

The EW corrections K-factor can be defined at the LO QCD level:

$$K_{\text{EW}}^{\text{NLO}} = \frac{\sigma_{\text{LO QCD}} + \sigma_{\text{NLO EW}}}{\sigma_{\text{LO QCD}}} \tag{8.7}$$

Meanwhile, the POWHEG +PYTHIA8 samples include terms

$$\sigma_{\text{POWHEG}} = (\sigma_{\text{LO QCD}} + \sigma_{\text{NLO QCD}})_{+\text{PS}} \tag{8.8}$$

here the subscript +PS indicates that parton shower simulation has been included as well. Potential variation of these parton shower effects should be covered by the standard uncertainties discussed in Sec. 8.8.

Since $\alpha \ll \alpha_s$, it often makes some sense to work at fixed order $m \leq N$, $n \leq 1$, rather than at $m + n \leq N$ (as mentioned in [66]). Thus, the EW corrections to the NLO QCD calculations are frequently included using the K-factor, defined at the

LO QCD level:

$$\sigma_{\text{NLO QCD}\times\text{EW}} \equiv K_{\text{EW}}^{\text{NLO}} \times \left(\sigma_{\text{LO QCD}} + \sigma_{\text{NLO QCD}}\right) \tag{8.9}$$

We refer to this as the "multiplicative approximation", which is similar to the "multiplicative approach" referenced in [65], aside from the missing NNLO QCD terms, which in our case should be covered by existing modeling uncertainties. This is the best approximation for EW corrections realistically available to us, and in some kinematic regimes these corrections do indeed factorize in this way from their QCD counterparts. Thus, this approach allows us to roughly approximate the terms in $\sigma_{\text{extra EW}}$ in addition to those directly generated in HATHOR. Specific to this analysis is the dependence of Higgs-based contributions on Y_{t} and the use of POWHEG, so we can write out our formula for observables more explicitly as

$$\sigma^{Y_{\text{t}}}_{\text{NLO QCD}\times\text{EW}} \equiv K_{\text{EW}}^{\text{NLO}}(Y_{\text{t}}) \times \sigma_{\text{POWHEG}}. \tag{8.10}$$

In order to estimate the uncertainty on this approach to including EW corrections, we investigate also the more naive additive approach,

$$\sigma^{Y_{\text{t}}}_{\text{NLO QCD}+\text{EW}} \equiv \left(K_{\text{EW}}^{\text{NLO}} \times \sigma_{\text{LO QCD}}\right) + \sigma_{\text{NLO QCD}}, \tag{8.11}$$

which neglects the term $\sigma_{\text{extra EW}}$ entirely. We take the difference between the two approaches as an uncertainty of the method:

$$\begin{aligned}
\text{EW unc.} &= \sigma^{Y_{\text{t}}}_{\text{NLO QCD}\times\text{EW}} - \sigma^{Y_{\text{t}}}_{\text{NLO QCD}+\text{EW}} && (8.12)\\
&= \left(\sigma_{\text{POWHEG}} - \sigma_{\text{LO QCD}}\right)\left(K_{\text{EW}}^{\text{NLO}}(Y_{\text{t}}) - 1\right) && (8.13)\\
&\equiv \sigma_{\text{POWHEG}} \times \delta_{\text{QCD}} \times \delta_{\text{EW}}, && (8.14)
\end{aligned}$$

where we have introduced the shorthand

$$\delta_{\mathrm{QCD}} = \frac{\sigma_{\mathrm{POWHEG}} - \sigma_{\mathrm{LO\,QCD}}}{\sigma_{\mathrm{POWHEG}}}, \quad \delta_{\mathrm{EW}} = (K^{\mathrm{NLO}}_{\mathrm{EW}}(Y_{\mathrm{t}}) - 1), \tag{8.15}$$

to illustrate that this uncertainty is effectively a cross term which arises from the difference in additive and multiplicative approaches. Note that δ_{EW} represents the marginal effect in terms of which EW corrections are often understood. This uncertainty is not a perfect description, but it is important to include near the threshold where two-loop diagrams of order $\mathcal{O}(\alpha_s^3\alpha)$ can make significant contributions. At high values of $M_{\mathrm{t\bar{t}}}$, large negative EW corrections result from Sudakov logarithms associated with W/Z exchange [65]. Such diagrams facctorize well, making the multiplicative approach valid to good approximation, and we thus do not include this uncertainty in that kinematic regime.

The implementation of this uncertainty is detailed further in Sec. 8.7.

8.5 Object & Event Selection

Backgrounds for $\mathrm{t\bar{t}}$ events in the dilepton channel are low due to the presence of 2 high-p_{T} leptons and 2 b-jets in the final state, which together are an ample signature to identify the process with high confidence. Backgrounds can thus be sufficiently modeled with simulated events, as opposed to data-driven estimation. The main backgrounds are single top production and Drell-Yan decays, examples of which are shown in Fig. 8.3. Single top production simulated at NLO with POWHEG in combination with PYTHIA8, while rare diboson events are simulated in PYTHIA8 with LO precision. Drell–Yan production is simulated at LO using the MADGRAPH generator [67] interfaced to PYTHIA8 with the MLM matching algorithm [68]. Overall, after event selection, the expected non-$\mathrm{t\bar{t}}$ background rate in this channel is about 5%.

Events are selected first using single electron or single muon triggers, designed to pick out events with at least one high-p_{T} lepton from the data. In the case of muons, we select events with a trigger p_{T} threshold of 24 GeV except in 2017 data, where the threshold is raised to 27 GeV. For electrons, we select events with a trigger p_{T} threshold of 27 GeV with the exception of 2018 when a threshold of 32 GeV is used due to availability. While dilepton triggers are also available with lower p_{T} thresholds, we found that they have no significant benefit in our case, as most events which interest us have a leading lepton p_{T} over 30 GeV.

To ensure that all electrons and muons are within the silicon tracker coverage, we require they have pseudo-rapidity $|\ \eta\ | < 2.4$. To operate well above the online trigger threshold, we then require at least one isolated electron or muon reconstructed with $p_{\mathrm{T}} > 30$ GeV, except in 2018, where offline selection of leading electrons is raised to $p_{\mathrm{T}} > 34$ GeV to match available triggers. A small buffer of a few GeV

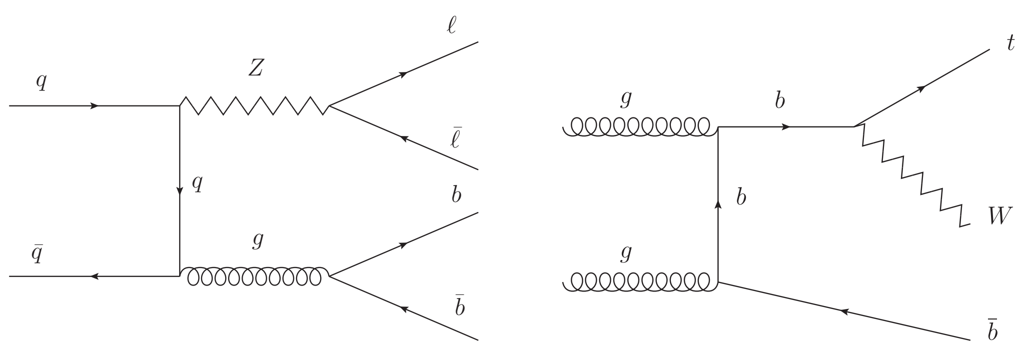

Figure 8.3: Examples of background processes for $\mathrm{t\bar{t}}$ dilepton decays are shown. The left diagram shows a Drell-Yan process involving a Z boson decaying to a lepton + anti-lepton pair, while the right diagram shows an example of single-top production.

between the online trigger selection and the offline p_T selection helps prevent potential disagreement between simulation and data near the trigger threshold. A second isolated electron or muon with $p_T > 20\,\text{GeV}$ is required. Any multi-lepton events with additional isolated leptons having $p_T > 15\,\text{GeV}$ are discarded, helping to lower diboson background contributions and avoiding events where additional work would be required due to ambiguity with lepton assignment.

Jets are clustered from particle-flow objects through the anti-$\text{k}_\text{T}$algorithm [69,70] using a distance parameter of 0.4. The jet momentum is calculated by taking the vector sum of momenta over all constituents identified as part of the jet. Important corrections to the jet energy are derived as a function of jet p_T and η, using both simulation and measurements of energy balance in data [71]. We select jets with $\mid \eta \mid < 2.4$ and $p_T > 30\,\text{GeV}$. We identify jets originating from b quarks using the DeepCSV algorithm [41], requiring all events to have at least two b-tagged jets according to the "loose" working point, which has the highest sensitivity. Events with more than two b jets are considered only if exactly two of the jets pass a more stringent selection. In this case these two jets passing the more stringent requirement are assumed to be those originating from top quark decays.

We find that in $\text{t}\bar{\text{t}}$ events, the two b jets are among the candidates passing the looser b-tagging requirement in roughly 93% of the events.

Drell-Yan background events (as in Fig. 8.3 left), mostly involving the decay of a Z boson, make a substantial contribution to the ee and $\mu\mu$ channels. To reduce their prevalence, we reject events with an invariant mass of the two leptons below 50 GeV or within 10 GeV around the Z boson mass of 91.2 GeV.

Source	2016 (X fb^{-1})	2017 (Y fb^{-1})	2018 (Z fb^{-1})	All (X+Y+Z fb^{-1})	% total MC
$\mathrm{t\bar{t}}$ MC	140 831±134	170 554±95	259 620±146	571 005±220	96.1%
V + jets MC	1944±54	3178±75	4958±133	10 080±162	1.7%
Single t MC	3022±25	3521±24	5827±32	12 370±47	2.1%
Diboson MC	142±9	151±10	250±15	543±20	0.1%
Total MC	145 940±148	177 404±124	270 655±201	593 999±279	100%
Data	144 817	178 088	264 791	587 696	98.9%

Table 8.1: MC and data event yields for all three years and combined. Statistical uncertainty on simulated event counts is given. For an estimate of systematic uncertainty, see Figs. 8.5-8.7.

The breakdown of expected signal and background yields for each year is shown in Tab. 7.1. The Drell–Yan and Single top quark production processes each account for roughly 2% of the estimated sample composition. Diboson decays, as well as other multi-jet events typically referred to as QCD background, make negligible contributions.

Additionally, there are multiple small discrepancies in efficiency when one applies the same selection criteria to simulation and data. This applies to the single lepton triggers, lepton identification, and b-jet identification. These disagreements are addressed via scale factor corrections to the event weights, which can be derived by comparing data and simulation in special cases–for example, the tag and probe method on Z decays is used for lepton-associated scale factors [72,73]. Uncertainty on these corrective scale factors is addressed in Sec. 8.8.

8.6 Event Reconstruction

While it is possible to completely reconstruct the neutrino momenta as described in Sec. 6.4, such reconstruction is highly sensitive to the MET ($p_{\mathrm{T}}^{\mathrm{miss}}$). Because the MET computation for each event relies on measurements taken throughout the detector and involving many different particles, it is vulnerable to large resolution effects and also systematic uncertainty resulting from the energy scales of various particles. For this reason, it was decided to forgo the full top quark reconstruction, instead defining proxy variables for $M_{\mathrm{t\bar{t}}}$ and $\Delta y_{\mathrm{t\bar{t}}}$. We found that a more sensitive measurement results from using the variables

$$M_{\mathrm{b}\ell} = M(b+\bar{b}+\ell+\bar{\ell}), \tag{8.16}$$

$$|\Delta y|_{\mathrm{b}\ell} = |y(b+\bar{\ell}) - y(\bar{b}+\ell)|, \tag{8.17}$$

which were found to have good sensitivity to EW corrections while being less vulnerable to resolution effects and systematic uncertainties. $M_{\mathrm{b}\ell}$ has the advantage of not dependening on the pairing of b jets with leptons at W decay vertices, meaning that it is reconstructed quite accurately. We do however utilize the kinematic constraints discussed in Sec. 6.4 to help with this pairing of b jets and leptons in order to reconstruct $\Delta y_{\mathrm{b}\ell}$. Including this variable was shown to add additional sensitivity, despite some resolution effects.

The full procedure for b-jet pairing is as follows:

1. The mass constraint is checked for both possible pairings. If only one pairing is found to satisfy the mass constraint, that pairing is used. If both pairings fail to satisfy the mass constraint, the event is discarded. If both pairings satisfy the mass constraint, we check the MET constraint.

2. If only one pairing allows for the MET constraint while the other does not, the pairing yielding an exact solution to the MET constraint is used.

3. If the neutrino kinematics do not suggest a clear pairing, the b-jets, b_1 and b_2, are paired with the leptons (l, $\bar{l}$) by minimizing the quantity

$$\Sigma_{1(2)} = \Delta R(b_{1(2)}, l) + \Delta R(b_{2(1)}, \bar{l})$$

among the two possible pairings, where $\Delta R(,\ell) = \sqrt{(\eta_{\mathrm{b}} - \eta_\ell)^2 + (\varphi_{\mathrm{b}} - \varphi_\ell)^2}$.

In the end the correct b-jet pairing is selected for approximately 82% of events in which both b-jets have been correctly identified.

The sensitivity of our chosen kinematic variables to Y_{t} is shown before and after reconstruction in Fig. 8.4 . The effect of Y_{t} remains substantial enough, and unique enough in shape, that a measurement can be extracted. When looking back at Fig. 8.2, we see that most sensitivity in $M_{\mathrm{t\bar{t}}}$ is more narrowly confined to the threshold region, which is particularly difficult to reconstruct accurately in the dilepton channel. By comparison, the sensitivity of Y_{t} to $M_{\mathrm{b}\ell}$ is distributed across a wider region.

We also see that sensitivity to our parameter of interest Y_{t} is not greatly reduced by the reconstruction process. This stands in contrast to typical attempts to reconstruct $M_{\mathrm{t\bar{t}}}$, where the sensitivity is confined to a narrow peak at the threshold and much information is lost to resolution effects.

The comparison between data and Monte Carlo is shown in Figs. 8.5-8.7. The agreement is generally seen to be well within known uncertainty. We do see some slopes in the ratio of data to MC prediction in the distributions of lepton and b-jet p_{T}, to varying extent for each year. These are likely related to a well known disagreement in top quark p_{T} distribution between POWHEG samples and data (See Ref. [74], for example).

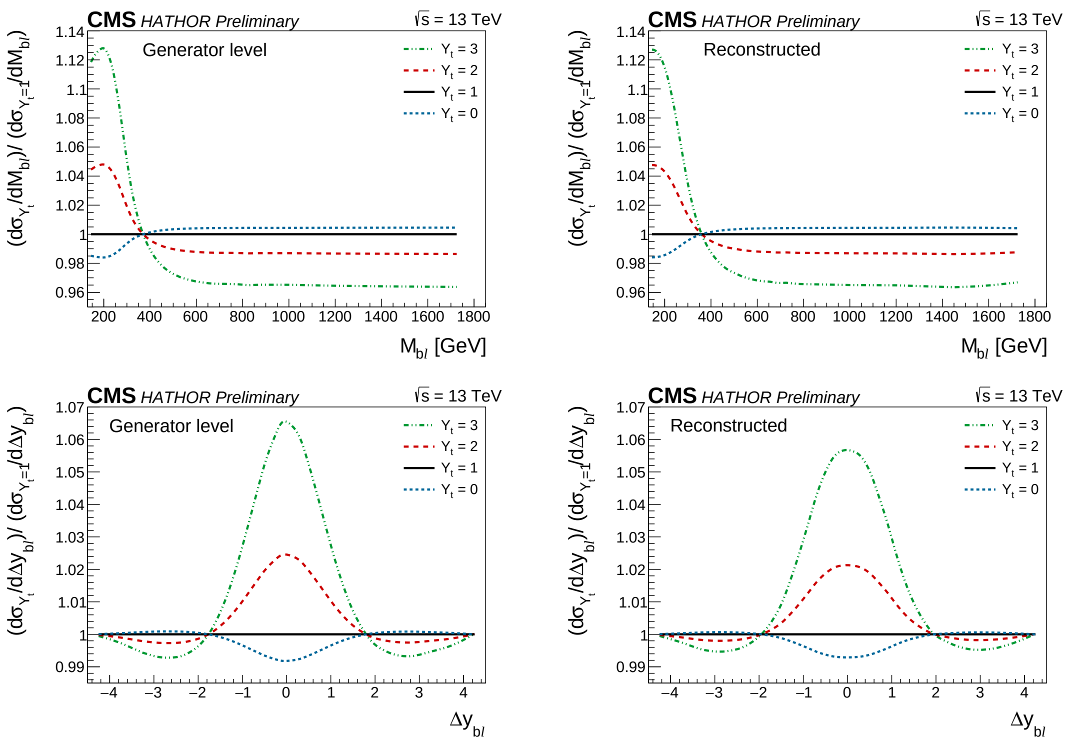

Figure 8.4: Effect of the EW corrections on $M_{b\ell}$ and $\Delta y_{b\ell}$ kinematic distributions are shown at generator level (left plots) and reco level (right plots). The effect is shown relative to the Standard Model EW corrections, as this is the best way to assess our sensitivity to different Yukawa coupling values, rather than comparing to a sample with no EW corrections.

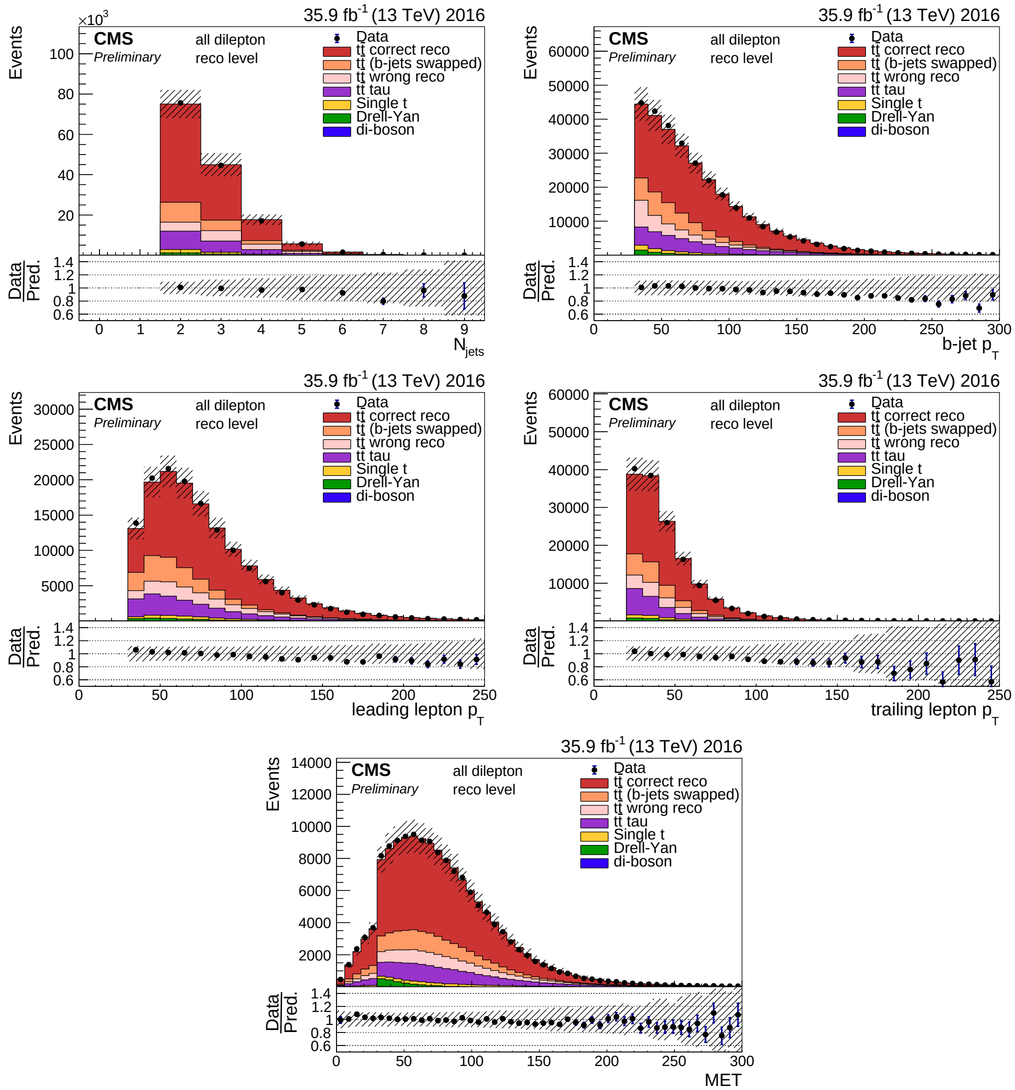

Figure 8.5: 2016 Data to MC comparison, with the shaded region indicating known uncertainty.

After reconstruction, events in each year are separated into two $|\Delta y|_{b\ell}$ bins and binned more finely in the $M_{b\ell}$ variable. The predominant sensitivity to Y_t comes from shape distortions to the $M_{b\ell}$ distribution, but the additional separation by $\Delta y_{b\ell}$ helps to distinguish these shape distortions from those associated with systematic uncertainties. The precise binning is chosen to obtain optimal information about shape variation while maintaining low statistical uncertainty by requiring $\gtrsim 10,000$

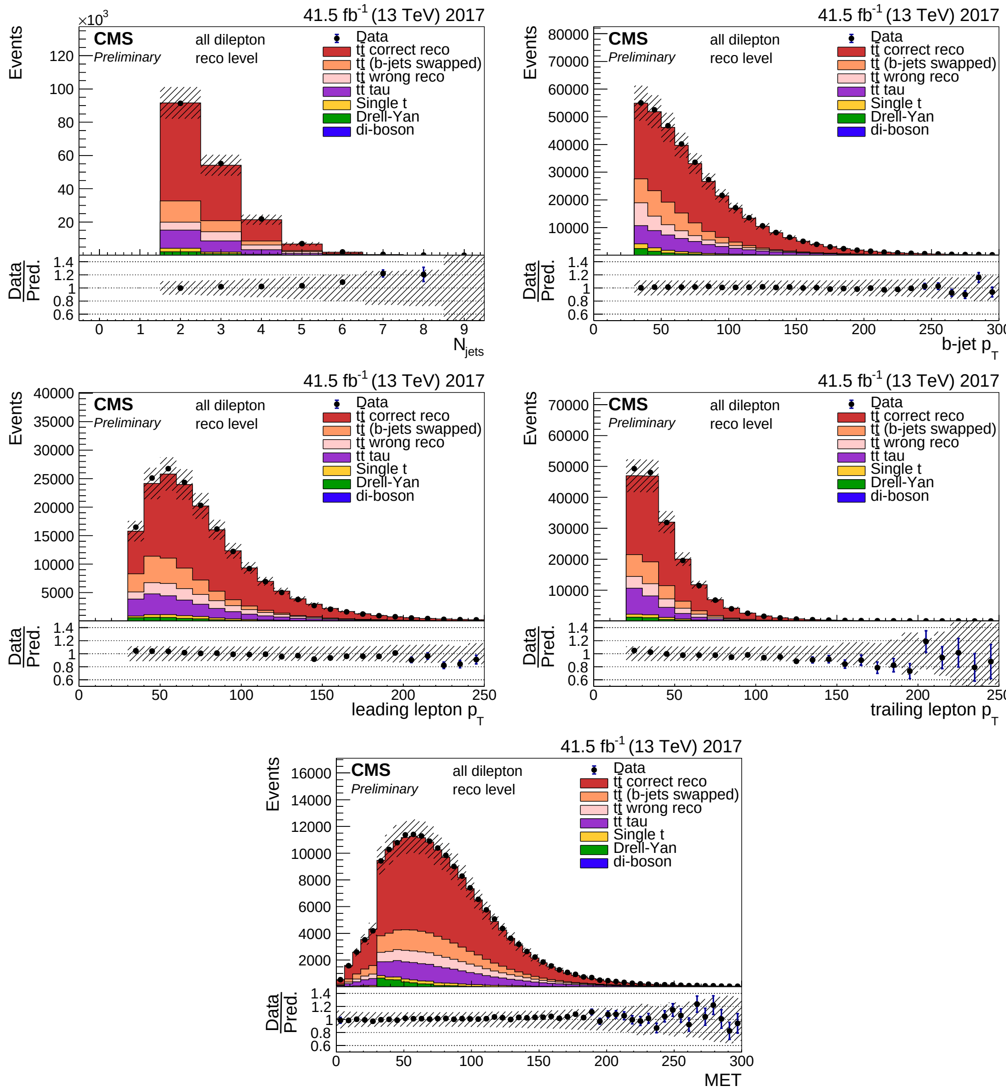

Figure 8.6: 2017 Data to MC comparison, with the shaded region indicating known uncertainty.

events present in each year in each bin. The data and simulation is shown with the final binning in Fig. 8.8.

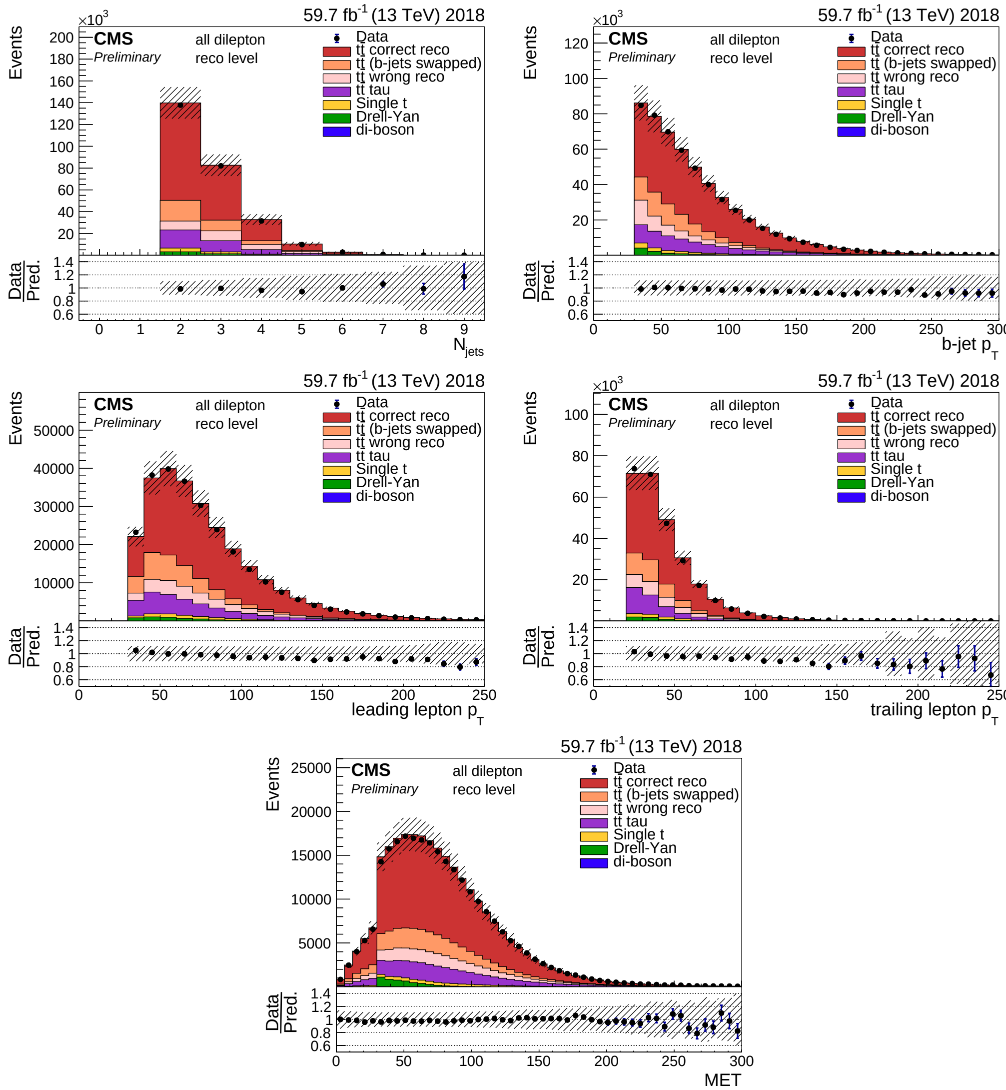

Figure 8.7: 2018 Data to MC comparison, with the shaded region indicating known uncertainty.

8.7 Statistical Methods

We define a likelihood function $\mathcal{L}$, which itself is defined as a product over separate likelihood functions evaluated at each bin,

$$\mathcal{L} = \prod_{\mathrm{bin} \in (M_{\mathrm{b}\ell}, |\Delta y|_{bl})} \mathcal{L}_{\mathrm{bin}} \times p(\phi) \times \prod_i p(\theta_i), \tag{8.18}$$

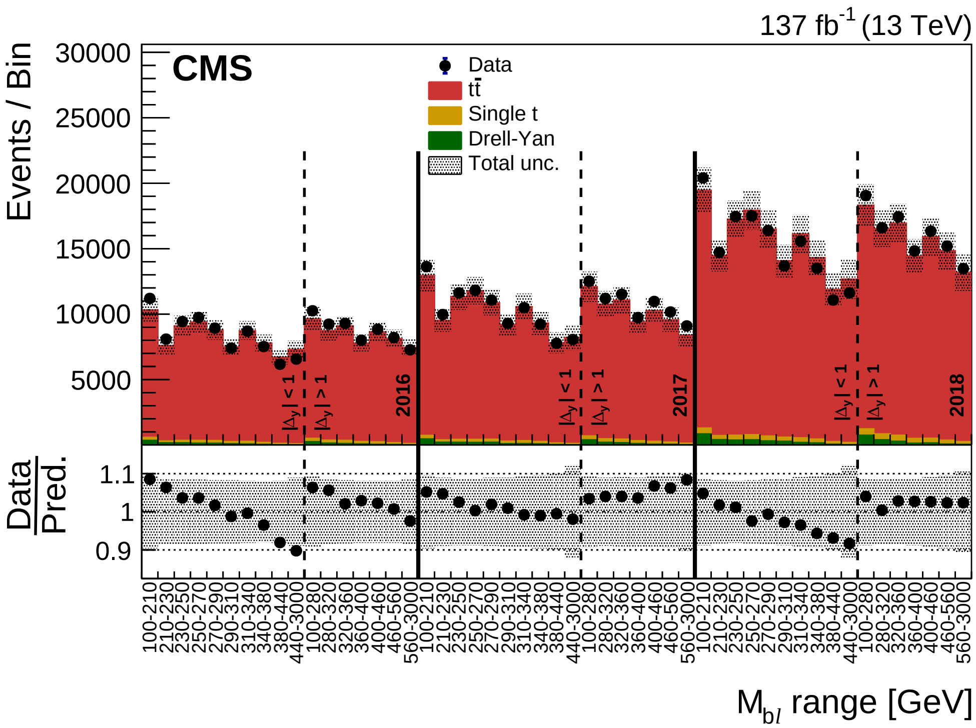

Figure 8.8: Binned data and simulated events prior to performing the statistical fit. The dashed line divides the two $|\Delta y|_{b\ell}$ bins, while the horizontal axis shows the bins' $M_{b\ell}$ ranges.

where in each bin we have a likelihood based on a Poisson distribution which describing the statistical fluctuation of the event count,

$$\mathcal{L}_{\text{bin}} = \text{Poisson}\left[n_{\text{obs}}^{\text{bin}} \middle| s^{\text{bin}}(\{\theta_i\}) \times R_{\text{EW}}^{\text{bin}}(Y_{\text{t}}, \phi) + b^{\text{bin}}(\{\theta_i\})\}\right]. \tag{8.19}$$

Systematic sources of uncertainty are represented by nuisance parameters θ_i and ϕ, which have penalty terms $p(\theta_i)$ and $p(\phi)$ accounting for for each nuisance parameter's deviation from its estimated value. In each bin, $n_{\text{obs}}^{\text{bin}}$ is the total observed bin count, with the expected bin count being the sum of the predicted signal s^{bin} and background b^{bin}. The number of expected signal events is modified by the additional rate R_{EW}, which encodes the EW corrections depends on the Yukawa coupling ratio Y_{t} and the special nuisance ϕ. The full expression for the rate $R_{\text{EW}}^{\text{bin}}$, including the modifying factor from this uncertainty, is

$$R_{\text{EW}}^{\text{bin}}(Y_{\text{t}}, \phi) = (1 + \delta_{\text{EW}}^{\text{bin}}(Y_{\text{t}})) \times (1 + \delta_{\text{QCD}}^{\text{bin}}\delta_{\text{EW}}^{\text{bin}}(Y_{\text{t}}))^{\phi}, \tag{8.20}$$

where $\delta_{\text{QCD}}^{\text{bin}}\delta_{\text{EW}}^{\text{bin}}$ represents the cross term arising from the difference in multiplicative and additive approaches of applying EW corrections as described in Sec. 8.2. This uncertainty is applied only to bins where EW corrections are increasing as a function of Y_{t}. The normally distributed nuisance parameter ϕ modulates this uncertainty.

Each of the other nuisance parameters, θ_i, represents an unknown parameter that can alter the expected bin counts. These are described by a standard normal Gaussian probability distribution function of $p(\theta_i)$, often called a penalty term, that affects the likelihood when this parameter deviates from our best-estimate value.

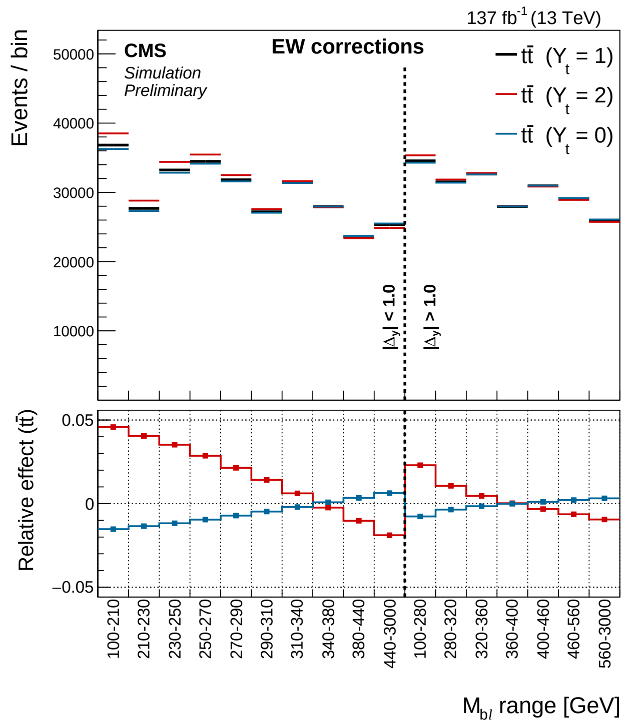

Figure 8.9: The effect of the yukawa parameter Y_t on binned reconstructed events is shown for two discrete choices of Y_t. Data from the full Run 2 period is combined in the kinematic binning.

The effects resulting from nuisance parameters $\{\theta_i\}$ are modeled by generating up and down variations of the bin content s^{bin} for each θ_i in each. These variations result from changing the underlying physical/technical sources, which are outlined in Sec. 8.8, usually by an amount of one standard deviation (σ) based on the uncertainty in our best estimates.

Collectively, the up or down variation across all bins generates a shape template associated with the nuisance parameter. These shape templates are then enforced as the bin modifiers associated with $\theta_i = \pm 1$, while $\theta_i = 0$ corresponds to the nominal estimate at which simulated data is produced.

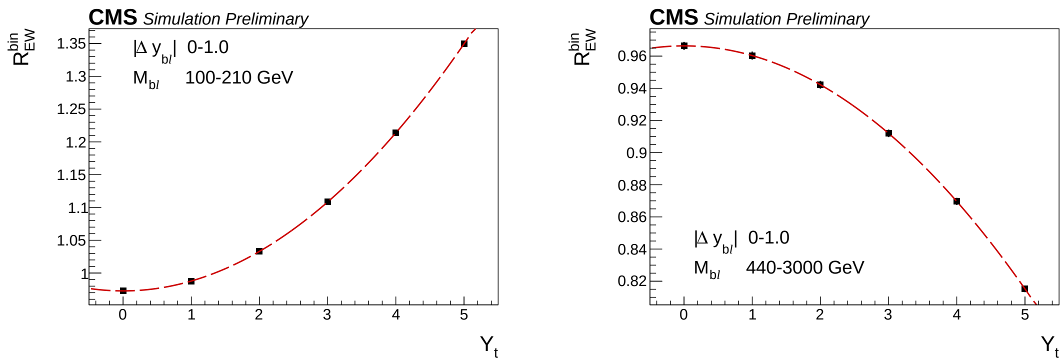

Figure 8.10: The weak correction rate modifier $R_{\mathrm{W}}^{\mathrm{bin}}$ in two separate $[M_{\mathrm{b}\ell}, \Delta y_{\mathrm{b}\ell}]$ bins from 2017 data, demonstrating the quadratic dependence on Y_{t}. All bins will have an increasing or decreasing quadratic yield function, with the steepest dependence on Y_{t} found at lower values of $M_{\mathrm{b}\ell}$.

An example of a shape template for one year of data, compared to the effect of the parameter of interest (POI), are shown in Fig. 8.9. The Quadratic dependence of the electroweak corrections on individual bins are shown in Fig. 8.10. Further examples of shape templates are shown in Sec. 8.9 We impose these shape templates through vertical template morphing as a function of the underlying nuisance θ_i, such that each bin the modifier is interpolated as a cubic spline for values of $\theta_i \in [-1, 1]$ and linearly outside of that region. This is the simplest procedure available to ensure that $s^{\mathrm{bin}}(\theta_i)$ is a continuous and differentiable function.

After constructing the full likelihood function, the best-fit value of Y_{t} is obtained by maximizing the likelihood. This maximization is then repeated with Y_{t} fixed over a fine array of fixed values of the POI and comparing the maximum log likelihood, we build confidence intervals (CI) at the 68% and 95% level.

8.8 Sources of Uncertainty

8.8.1 Experimental sources of uncertainty

The dominant experimental uncertainty in this analysis comes from the detector's jet energy response. Because jets are the result of clustering many individual particles, accurate measurement of the jet energy is challenging. Corrections are derived as a function of jet p_{T} and η based on simulation, which are then improved using data. We follow the standard approach outlined in [71] to consider 26 separate uncertainties which are involved in determining these calibrations. In this approach, uncertainty on the resolution of the jet reconstruction is considered as well as a separate uncertainty. When deriving shape templates for the associated uncertainties, the altered jet energy is propagated to the calculation of $p_{\mathrm{T}}^{\mathrm{miss}}$.

Other experimental sources of uncertainty are comparatively minor. The uncertainty on the integrated luminosity has an overall normalization effect on signal and background processes in each data-taking year, ranging from 2.3% to 2.5% depending on the year.

Unwanted pileup events are modeled and injected into simulation to include their effects on event reconstruction. Uncertainty on the number of pileup events included in this way is assessed by varying the inelastic cross section, 69.2 mb, by its current uncertainty of 4.6% [75].

Efficiencies in b jet identification and mis-identification are corrected, as a function of jet p_{T} and η, to better match data [41]. These uncertainties have a normalization effect of a few percent and a smaller shape component, and do not have much effect on the fit.

Similarly, scale factors are applied as functions of p_{T} and η to correct simulated efficiencies of lepton identification, isolation requirements, and triggers to match data. These are derived for muons and electrons using the tag-and-probe method on Z decays [72,73]. The tag-and-probe fit accounts for uncertainty from statistics as well as difference in jet multiplicity. Overall, the effect is seen to be below 2%, and approximately uniform.

8.8.2 Theoretical sources of uncertainty

As a standard technique to estimate potential contributions of higher-order QCD terms at the matrix element level, the QCD renormalization scale μ_{r} and factorization scales μ_{f} are varied up and down in the POWHEG simulation by a factor of 2. The scales are each varied individually to form templates, and then a same-sign variation of the two scales together is included as an additional uncertainty. These three shape templates together cover various higher-order effects, as higher order perturbative calculations depend less and less on the scale, ultimately in principle converging to a scale-independent limit. In generating the scale variation templates, the overall normalization effect is removed, as the NNLO $\mathrm{t\bar{t}}$ cross-section is already used to improve MC normalization. This allows us to consider mainly shape effects, which we are far more sensitive to.

A 5% normalization uncertainty is included on the $\mathrm{t\bar{t}}$ cross section, which more than covers expected contribution from terms beyond NNLO QCD. The backgrounds in this analysis are small enough (each making up $\approx$2% or less of the expected sample composition) that we do not generate shape templates for their response to individual systematic uncertainties. We do, however, include substantial normalization uncertainties meant to broadly cover uncertainty on background modelling, which will have an effect on reconstructed distribution shapes. A 15% sample is included on single top quark production, which was seen to cover the scale variation and jet energy correction uncertainties in magnitude. A 30% normalization uncertainty is included on Drell-Yan and diboson event simulation, because these MC samples are generated with LO calculations and have larger uncertainties associated with scale variation.

We include an uncertainty on the weak corrections, based on our methods for generating and applying these additional terms, as outlined in Sec. 8.4 and 8.7. Like the scale variations, this uncertainty stands in for higher-order effects at the ME level, predominantly those coming from diagrams of order $\alpha_s^3\alpha$. This effect is included only for the region near the threshold where the EW corrections have a positive effect on the $t\bar{t}$ cross section. The uncertainty manifests as an envelope around the EW corrections in this regime, as shown in Fig. 8.11. This is a key uncertainty in the fit, as it encodes our limited knowledge of the EW interactions from which we derive our sensitivity.

To model uncertainty in parton distribution functions, NNPDF set [63] provides 100 individual variations (replicas) as uncertainties. As in [76], variations with similar shape templates are combined to reduce the number of variations to a more manageable 8 in total. These uncertainties are treated as fully correlated between 2017/2018 datasets, but not in the 2016 data as the simulation uses a different version of NNPDF with incompatible variations. Variation of the strong coupling constant α_s used by NNPDF is included as well. Though the these uncertainties can have significant shape effects, they are generally small, on the order of $\sim$1% or less.

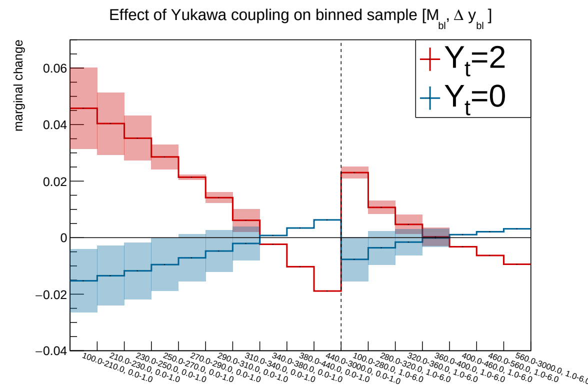

Figure 8.11: The marginal effect of the EW corrections is shown at two values of top Yukawa coupling. The colored lines show the nominal value of the effect, while the shaded envelope indicates the included uncertainty on the EW corrections themselves.

Dedicated MC samples are generated with the top mass varied up and down by 1 GeV to model uncertainty on the measured top mass (though at present this value is an overestimate of that uncertainty). Due to statistical fluctuations in the dedicated samples, simulations from all three years are combined to generate one shape template, and the uncertainty is taken as correlated between years. It should be noted that, although the top mass and Yukawa coupling are generally treated independently in this measurement strategy, varying the mass will slightly modify the definition of $Y_t = 1$. However, the effect is below 1%, making it much smaller than the sensitivity of the measurement and safe to ignore.

The h_{damp} parameter, which controls the matrix element parton shower matching in POWHEG + PYTHIA8, is varied in order to estimate the uncertainty in the matching algorithm. Dedicated MC samples are also generated with variations of the PYTHIA8 tune in order to estimate uncertainty in the hadronization model.

Uncertainty due to initial and final state radiation in the parton shower algorithms is assessed by varying the energy scale in the initial and final states by a factor of 2.

Lastly, there are two more modelling uncertainties related to b-jet modeling. The branching ratio of semi-leptonic B decays affects the b-jet response, so this quantity is varied within its measured uncertainty [36] to model its effects. Additionally, The momentum transfer from b quarks to B hadrons is dependent on the fragmentatiom rate $x_b = \frac{p_{\mathrm{T}}(B)}{p_{\mathrm{T}}(b-jet)}$. To include the uncertainty of this ratio, the transfer function is varied up and down within uncertainty of the Bowler-Lund parameter [77] in PYTHIA8 and the resulting effect is applied via event weights.

8.8.3 Processing of shape templates

Shape templates are generated for all uncertainties which are not explicitly normalization-based uncertainties. These templates are generate separately for each year, due to minor differences between data taking periods. In many cases the underlying nuisance parameter is fully or partially correlated between years, which is enforced in the likelihood fit. Some shape templates are also smoothed to eliminate noise, or made to be exactly symmetric, both of which help to ensure stability and performance of the fit. Templates which do not exhibit an effect of $> 0.1\%$ after smoothing are converted to an overall normalization uncertainty, as shape distortions are not meaningful at this level. Example shape templates are shown for the 3 dominant uncertainties in Fig. 8.12, where the effects of smoothing and symmetrizing can be seen.

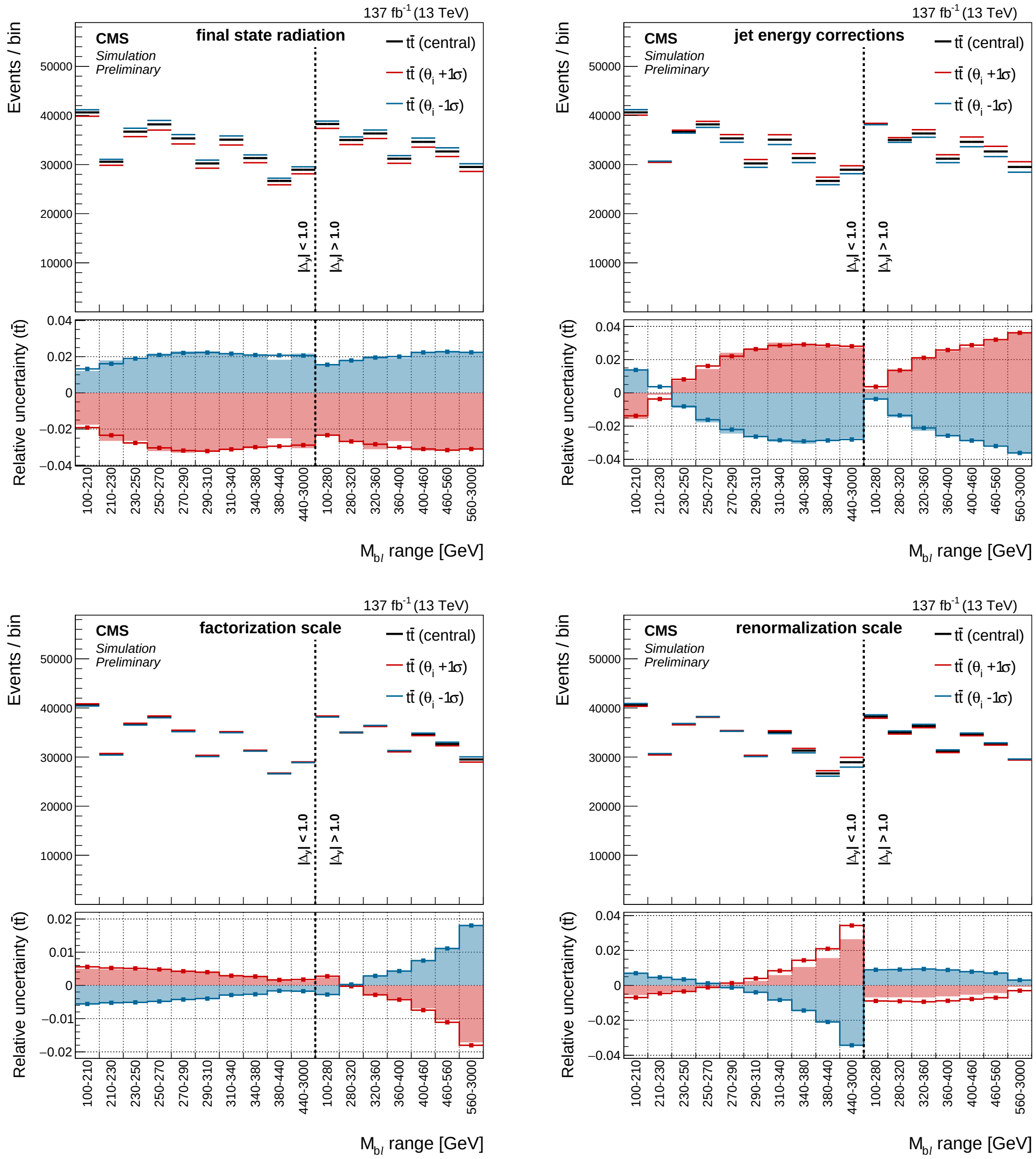

Figure 8.12: Shape templates are shown for dominant uncertainties in the fit, associated with the final-state radiation in PYTHIA8, the jet energy corrections, the factorization scale, and the renormalization scale. The shaded bars represent the raw shape template information, while the lines show the shapes after smoothing and symmetrization have been applied. In the fit, the jet energy corrections are split into 26 different components, but for brevity only the total uncertainty is shown here. Variation between data-taking years is minimal for each of these uncertainties, so their effect is shown on the full dataset in the kinematic binning. They are treated separately by year in the actual fit, to ensure accurate modelling.

8.9 Results

The maximum likelihood fit yields a best fit value of $Y_t = 1.16^{+0.24}_{-0.35}$, with $Y_t < 1.54$ at a 95% confidence level, with impacts of significant uncertainty sources shown in Fig. 8.13 Agreement of data and simulation in the final analysis binning is shown for the final configuration after performing the maximum likelihood fit in Fig. 8.14. We see that the fit is able to obtain good agreement and remove sloping seen in the pre-fit data (Fig. 8.8). A likelihood scan visualizing the construction of the confidence intervals is shown in Fig. 8.15, performed for both data and simulation. The result is in agreement with the analogous measurement obtained previously in Ref. [76]. While it does not mark a major improvement in sensitivity, this measurement serves as a valuable complement, particularly due to the use of a different decay channel with different backgrounds.

Though we achieve lower sensitivity compared to production-based constraints in ttH [49] and Higgs combination measurements [51], it should be noted that these measurements depend as well on other Yukawa coupling values through additional branching assumptions. This is part of the challenge of measurements in the standard model: the theory is so complex and contains so many free parameters, it can be challenging to isolate just one. Thus, it is valuable to have measurements with different approaches which fold in different assumptions. Our result serves as a compelling orthogonal measurement to other techniques that involve Higgs production and decay. Another recent measurement which possesses a similar independence to other branching ratios comes from searches for the production of four top quarks [78], which yields a 95% CI upper limit of $Y_t < 1.7$ [78], compared to the 95% upper limit of $Y_t < 1.54$ seen here.

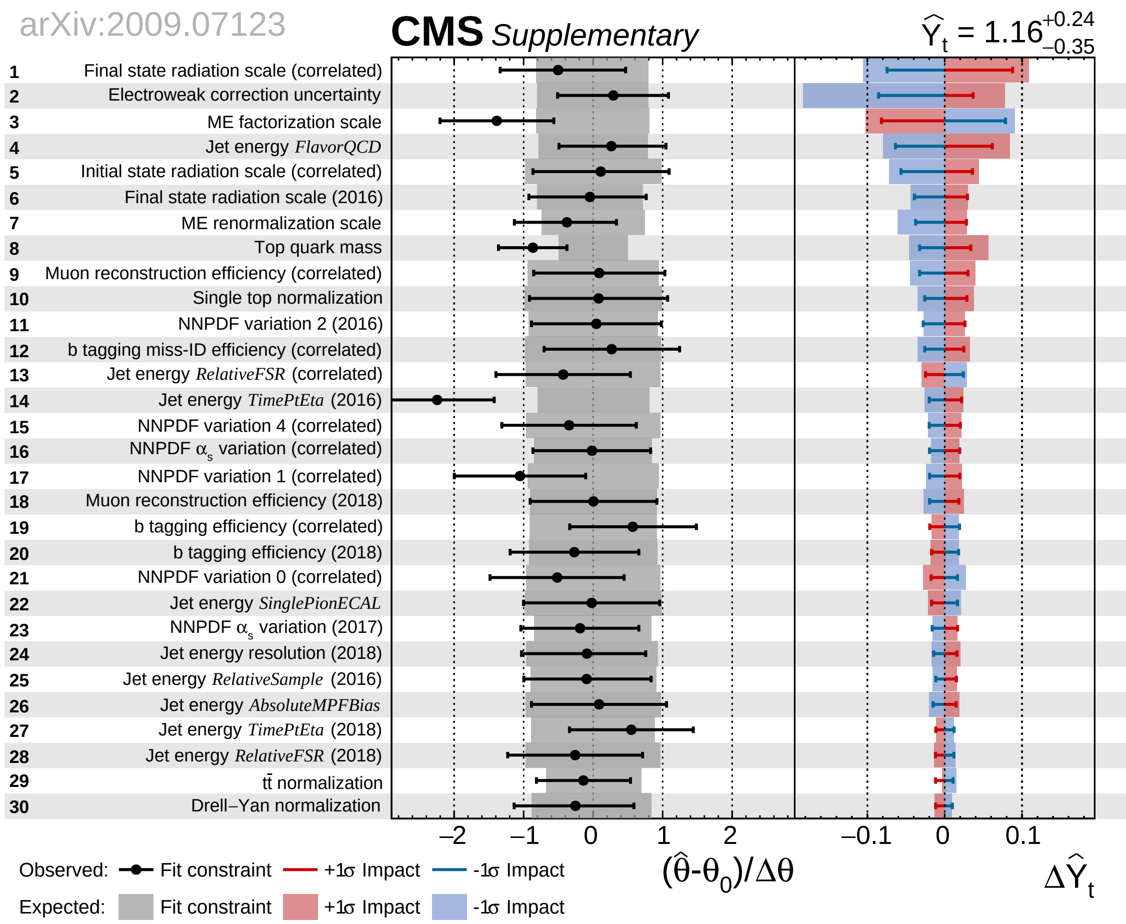

Figure 8.13: Nuisance parameter post-fit constraints and impacts are shown for the 30 uncertainties in the fit with the largest impact on the best fit value of Y_t. This information is displayed for an Asimov dataset generated with $Y_t = 1$ (expected) and for the final fit result on data (observed). Impacts are calculated by repeating the fit with individual nuisance parameters varied up and down by one standard deviation according to the their post-fit uncertainty and held fixed. Some uncertainty sources are split into a correlated nuisance parameter and an uncorrelated parameter unique to each data-taking year. Uncertainty on corrections to jet energy is broken down into several sources, whose names are indicated in italics. Technical details on the meaning of each such component can be found in Ref. [71]

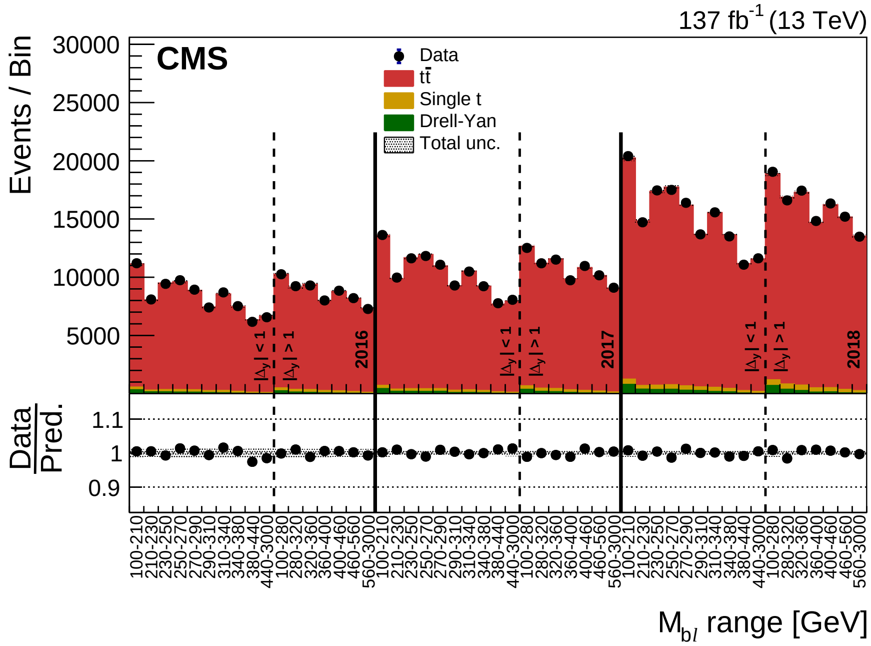

Figure 8.14: The agreement between data and MC simulation at the best fit value of $Y_t = 1.16$ after performing the likelihood maximization.

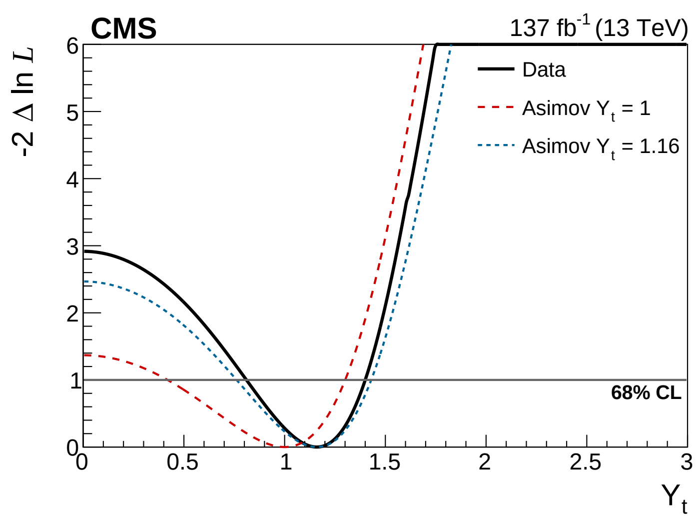

Figure 8.15: The result of a likelihood scan, performed by fixing the value of Y_t at values over the interval [0, 3.0] and minimizing the negative log likelihood ($-\ln \mathcal{L}$). Expected curves from fits on simulated data are shown produced at the SM value $Y_t = 1.0$ (red, dashed) and at the final best-fit value of $Y_t = 1.16$ (blue, dashed).

Chapter 9

Concluding Remarks

"The goddess of learning is fabled to have sprung full-grown from the brain of Zeus, but it is seldom that a scientific conception is born in its final form, or owns a single parent. More often it is a product of a series of minds, each in turn modifying the ideas of those that came before, and providing material for those that come after."

— J.J. Thomson

The standard model of particle physics has been wildly successful at explaining the behavior of matter at the smallest scales and highest energies which can be probed in a lab setting. However, for as many successful predictions as it makes, it raises an equal number of questions. The most famous of these questions arise when the standard model is compared to large scale behavior of the observed universe. But, as we have seen, the standard model leaves many open paths worthy of further pursuit–from neutrino mass generation, to dark matter candidates, to precision coupling measurement. Even QFT, which forms the basic framework for the standard model, has mysterious mathematical underpinnings and a host of unsolved conjectures.

In this document, we have examined two very different approaches to research in fundamental particle physics pursued by the author, while providing broader context to each. First, in Chapter 4, we discussed ways of examining reduced-dimensional quantum fields as a way to study the framework that governs particle interactions. There is much left to be understood here, and there are unsolved problems that have stood for over half a century. The path forward remains completely unknown. Some of the same problems can be approached from either the field of mathematics or the field of theoretical physics, two groups which see the issues from surprisingly different perspectives. In attempting to learn pieces of both perspectives, the author performed detailed studies of reduced dimensional systems, focusing on both mathematical [28] and physical [32] properties of quantum fields. While this work hardly answers the most fundamental questions involving QFT, both projects demonstrate that there are remaining avenues of study that can produce interesting, unintuitive results.

Secondly, in Chapters 7-8, we looked at the CMS experiment and a novel measurement of the top-Higgs interaction strength from Ref. [79]. Though we see a measured value consistent with the standard model measurement, we also see that further study with greater precision is needed. Because the top quark and Higgs boson are both the newest particles to be discovered and the heaviest known to exist, it makes sense to study their properties and interactions in detail. This interaction has only recently begun to be measured, and it is important to continue performing such measurement with a variety of orthogonal techniques. Historically, precise measurements at a given energy scale have often provided clues about what will be seen at higher energies, and even small deviations from what is expected of the Higgs properties could have huge ramifications.The LHC will continue to collect data for years to come, and for decades as the upgraded HL-LHC. We will have to wait and see what the program yields in the long term.

It is impossible to know which direction of research will yield groundbreaking results. And it is sometimes said that the state of physics in the early 21st century bears a striking resemblance the early 20th century. Back then, the story goes that many believed all fundamental rules of matter and energy had been laid out by Newton's gravity, atomic theory, thermodynamics, and Maxwell's electromagnetism. Some say that these models were so successful, it created a sense that it would be difficult to progress much further. And yet, physicists at the time unknowingly sat at the edge of a revolution, in which general relativity and quantum mechanics would upend their understanding of light, matter, and gravity.

Of course, groundbreaking physics results have continued in fields not mentioned here. But particle physics of late seems to have slowed its pace. What is to become of our fundamental understanding of particles, the tiny pieces that make up the universe and ourselves? Could we be sitting on the edge of a new revolution, that will upend our understanding in the coming century? If so, where might it come from? Or has the standard model indeed brought us close to a limit of sorts, such that particle experiments in the foreseeable future will only serve confirm the model we already have? Only time will tell, but it's best to keep an open mind.

Appendix A

Continuity of Scalar Fields With Logarithmic Correlations

Partial reprinting of Ref. [28] with permission, co-authored with S.G. Rajeev.

A.1 Introduction

In traditional geometry, the distance between two points is the length of the shortest curve that joins them. This fits well with classical physics, as this shortest path is the one followed by a free particle. But in quantum physics, the shortest one is only the most likely of many paths that the particle can take. Moreover, no particle can follow a path connecting two spacelike separated points. Taking these facts into account, we should hesitate to associate distance with the length of one particular curve. Instead, we can average over all paths connecting two points, yielding the Green's function of a quantum field (also called the two point function, correlation function or propagator.) A metric does emerge out of the correlation, but turns out to be non-Euclidean [29].

The idea of defining a metric from the correlation of a random process is a staple of modern stochastic analysis [80, 80–82]. This can be illustrated with Brownian motion. The Brownian paths are continuous, but not differentiable with respect to the usual time parameter. A particle executing Brownian motion is knocked around by other particles in the medium. As the time between collisions tends to zero, the velocity at any instant is no longer a physical quantity. Furthermore, even the speed cannot be bounded; as $x \to y$ the probability of $\frac{|B(x)-B(y)|}{|x-y|}$ being bounded is zero (To make comparison with a quantum field easier, we call the time parameter of the Brownian process x rather than t. Since the diffusion constant has dimension $(\text{length})^2/\text{time}$, dimensional analysis suggests that

$$\frac{|B(x) - B(y)|}{\sqrt{|x-y|}} \tag{A.1}$$

would be a better quantity to measure the speed of a Brownian particle. But it turns out that even this is unbounded with probability 1 (more commonly stated as "almost surely" or a.s.) as $x \to y$. The proper way to quantify the time that has elapsed between two measurements is not $|x-y|$ or even $\sqrt{|x-y|}$. We seek a metric ω with respect to which the sample paths are locally Lipschitz continuous, meaning $\frac{|B(x)-B(y)|}{\omega(x,y)}$ is almost surely bounded as $x \to y$. The correct such "modulus of continuity," attributed to Lévy, is

$$\omega(x,y) \propto \sqrt{|x-y| \log \frac{1}{|x-y|}} \tag{A.2}$$

for small $|x-y|$. This quantifies the roughness of Brownian paths (One can bound the variations precisely with a proportionality constant $\sqrt{2}$, but we will generally ignore multiplicative constants in discussing continuity/roughness here).

We look at the spatial metric in the simplest relativistic theory, a massless scalar quantum field in $1+1$ dimensions. Such logarithmically correlated fields have generated interest in purely mathematical contexts, and have potential applications in areas ranging from finance to cosmology (see [27]). It is enough to understand the continuity of sample fields in the ground state; those in any state of finite energy will exhibit identical behavior over small distances (see the appendices in the original publication of Ref. [28] for supplementary discussion of the ground state wavefunction).

A complication is that scalar quantum fields are random distributions rather than functions: $\phi(x)$ at some point in space is not a meaningful quantity. But we will show that a mild smoothing procedure (averaging over an interval) is enough to get around this difficulty, yielding a continuous but not differentiable function. This average can be viewed as a model for the potential measured by a device: such a measurement will always take place over some finite width. A peculiar property of the logarithmically correlated field is that the probability law of the average is independent of the size of the interval (size of the measuring device). That is, the field does not appear any rougher if we average over smaller intervals.

We then obtain a result analogous to that of Levy: a metric in space with respect to which the scalar field is a.s. Lipschitz (we will use this term exclusively in the sense of local continuity). Our result is a particular case of the much deeper mathematical theory of regularity of random processes [80, 80–82] . The idea of using a moving average (instead of an inner product with smooth test functions) seems to be new, and yields simple explicit results.

We then apply this moving average technique to other random fields of physical interest, noting that a new procedure is sometimes needed if the field considered has more severe divergences. Supplemental material can be found in the original publication of Ref. [28] discussing the intimate relationship of this work to the resistance metric on a lattice, connecting to an earlier paper [29], and describing connections to a functional analytic approach to regularity of random processes.

Although we work with Gaussian fields in this paper, the short distance behavior is the same for asymptotically free interacting fields (up to sub-leading logarithmic corrections). The regularity of renormalizable but not asymptotically-free theories (such as QED or the Higgs model) can be quite different. The strength of interactions grow as distances shrink, possibly leading to a singularity (Landau pole). In the case of QED, we know that this is not physically significant, due to unification with weak interactions into a non-Abelian gauge theory.

But the question of what happens to the self-interaction of a scalar quantum field (Higgs boson) at short distances is still open. In the absence of evidence at the LHC for supersymmetry or compositeness of the scalar, we have to consider the possibility that the Higgs model is truly the fundamental theory. The short distance behavior is dominated by interactions, necessitating new mathematical methods beyond perturbative renormalization. The extensive mathematical literature [80, 80–82] on continuity of non-Gaussian processes ought to contain useful tools for physics. In order to apply this work to a full interacting theory, we must first know what happens in the simpler case of a free theory. This is part of the physical motivation for this paper.

A.2 Continuity

A.2.1 Continuity of Random Processes

A random process $r(x)$ assigns a random variable to each value of x in some space X. The quantity

$$d(x,y) = \sqrt{\langle [r(x) - r(y)]^2 \rangle} \tag{A.3}$$

satisfies the triangle inequality and so defines a metric on X (provided we identify any originally distinct points x, y for which $d(x,y) = 0$).

This metric need not be Euclidean or even Riemannian. A standard example is Brownian motion, where $d(x,y) = \sqrt{|x-y|}$, which is neither. Supplementary material on another simple example can be found in the appendices of the full publication, Ref. [28].

One commonly successful approach to the study of continuity is to leave behind the intuitive structure associated with the space X, and begin instead by looking at structures related to the process of interest (such as the metric d above). At first, one might expect that the sample paths for a random process $r(x)$ will be necessarily continuous with respect to d. Although true for Brownian motion, almost sure continuity with respect to d does not hold in general. For Gaussian processes, a sufficient condition for continuity is the convergence of the Dudley integral [80, 80–82]

$$J(\delta) = \int_0^\delta \sqrt{\log N(D,\epsilon)}\, d\epsilon, \quad \delta < D \tag{A.4}$$

where $N(D,\epsilon)$ be the minimum number of balls of radius ϵ it takes to cover a ball of radius D in (X,d) (We suppress the D dependence of J for simplicity of notation). The possible divergence comes from the lower limit of the integral $\epsilon \to 0$.

Can we go beyond continuity? To speak of differentiable functions, a metric is not enough: we would need a differentiable structure on X which we do not have intrinsically. The closest analogue to differentiable functions on a metric space (X,d) are Lipschitz functions, for which $\frac{|f(x)-f(y)|}{d(x,y)}$ is bounded. For comparison, differentiable functions on the real line are Lipschitz, but not all Lipschitz functions are differentiable. Of course, all Lipschitz functions are continuous.

Even in cases where $J(\delta)$ converges, indicating that the sample paths are continuous, they may still not be Lipschitz with respect to the metric d above. Again, the Dudley integral comes to the rescue: using it we can define a more refined metric

$$\omega(x,y) = J(d(x,y)). \tag{A.5}$$

Since $\sqrt{\log N(D,\epsilon)}$ is a decreasing function of ϵ, $J(\delta)$ is a convex function. Thus $J(d(x,y))$ satisfies the triangle inequality as well.

The sample paths of a Gaussian process for which $J(\delta)$ converges are [83] a.s. Lipschitz in this refined metric ω. Thus ω, rather than d, is the metric ("modulus of continuity") we must associate to a Gaussian random process.

What would one do if the Dudley integral does not converge? There is a more general theory [81] which gives necessary *and sufficient* conditions for continuity: a "majorizing measure" must exist on X . We will not use this theory in this paper, but hope to return to it, as it can deal with more general cases than Gaussian processes (e.g., interacting quantum fields).

A.2.2 Quantum Fields

In this paper, we consider a quantum scalar field ϕ in the continuum limit. In the trivial case where ϕ is massless with 1 spatial dimension and no time dimension, the correlations of Brownian motion are reproduced and ϕ remains a continuous function. However, in any fully relativistic field theory, ϕ lives on a space of distributions, not functions. To get a sensible random variable, we must then take the inner product with respect to some test function h with zero average.

$$\phi[h] = \int \phi(x)h(x)dx, \quad \int h(x)dx = 0. \tag{A.6}$$

We study the case where ϕ is a distribution with the weakest possible singularities; one might say we want a field that is "close" to being a function. The obvious candidate is the case of logarithmic correlations (For a recent review, see [27])

$$\langle \phi[h]\phi[h'] \rangle = -\int \log|x-y|\; h(x)h'(y)dxdy. \tag{A.7}$$

The corresponding Gaussian measure can be thought of as the (square of the) ground state wavefunction of a massless scalar field in $1+1$ dimensions. (More precisely, the continuum limit of the resistance metric of a row on an infinite square lattice, as discussed further in the supplementary material of Ref. [28]).

The condition $\int h(x)dx = 0$ ensures that the covariance is unchanged if $\log|x - y|$ is replaced by $\log \lambda|x-y|$, meaning ϕ is scale invariant. Since ϕ has the physical meaning of a potential, observables such as $\phi[h]$ must be unchanged under a shift $\phi(x) \mapsto \phi(x) + a$, which equivalently suggests the requirement $\int h(x)dx = 0$.

A.2.3 Moving Average of a Quantum Field

Quantum fields which are only mildly singular can act on test functions which are not smooth or even continuous. It is not necessary to consider the whole space of test functions as in [82]; in this paper our test functions will be piecewise constant with compact support and zero mean.

We define a *moving average* of ϕ:

$$\bar{\phi}_s(u) \;=\; \int_{-\frac{1}{2}}^{\frac{1}{2}} \left\{\phi\left(s\left[u-w\right]\right) - \phi\left(s\left[0-w\right]\right)\right\} dw, \quad s > 0. \tag{A.8}$$

This is the inner product of ϕ with a discontinuous test function h that has support on two intervals of width s based at su and at 0; the sign is chosen so that $\int h(x)dx = 0$. The probability law of $\bar{\phi}_s$ is not translation invariant: the second term ensures the boundary condition

$$\bar{\phi}_s(0) = 0. \tag{A.9}$$

We will mostly work with the quantity

$$\bar{\phi}_s(u) - \bar{\phi}_s(v) \;=\; \int_{-\frac{1}{2}}^{\frac{1}{2}} \{\phi\left(s\left[u-w\right]\right) - \phi\left(s\left[v-w\right]\right)\}\, dw \tag{A.10}$$

which has a translation invariant law. It is convenient to rescale the coordinate of the midpoint by the width (as we have already done), so that the variable u is dimensionless. Then the quantity

$$\rho(u,v) \equiv \sqrt{\langle[\bar{\phi}_s(u) - \bar{\phi}_s(v)]^2\rangle}, \tag{A.11}$$

which is just a special case of (A.3), is finite and defines a metric. Moreover, it is independent of s in the logarithmically correlated case. This means the process $\bar{\phi}_s(u) - \bar{\phi}_s(v)$ has a probability law that is independent of s: a consequence of scale invariance, which is specific to logarithmic correlations. As an interesting aside, we note that $s\bar{\phi}_s(u)$ produces a solution to the wave equation in u and s.

The moving average does not depart from the essence of the standard idea of averaging over a test function. It is simply that a piecewise constant test function is especially convenient for a mildly singular quantum field as opposed to a smoother function. For more singular fields (e.g. scalar field in four dimensions) we would have to revert to more regular test functions.

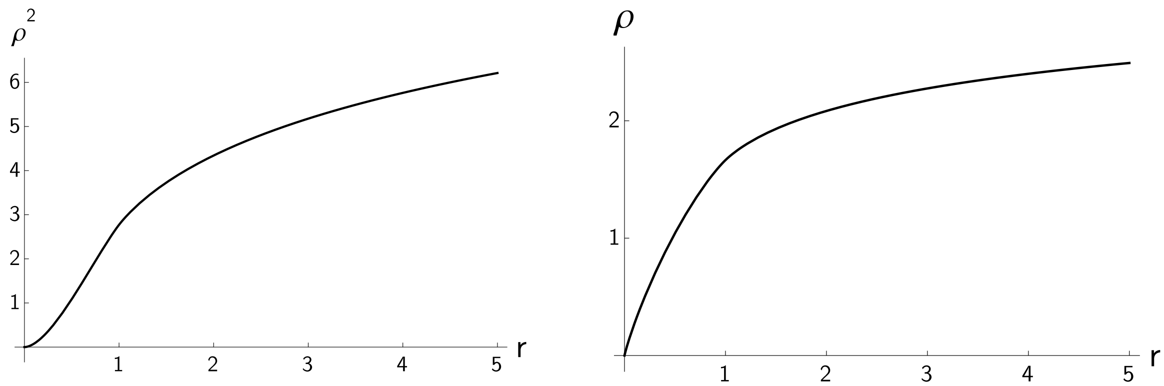

Figure A.1: The behavior of $\rho^2(r)$ and $\rho(r)$, setting $s = 1$.

A.3 Logarithmically correlated scalar field in 1 dimension

In the logarithmically correlated case, we obtain explicit formula

$$\rho(u, v) = \rho(|u - v|) \tag{A.12}$$

$$\rho(r) = \sqrt{L(r+1) + L(r-1) - 2L(r)}, \tag{A.13}$$

Where

$$L(r) = \frac{1}{2} r^2 \log r^2. \tag{A.14}$$

Being a convex function of $r = |u - v|$, this $\rho(u, v)$ will satisfy the triangle inequality (not true of ρ^2, as seen in Fig. A.1). Thus, ρ defines a translationally invariant metric.

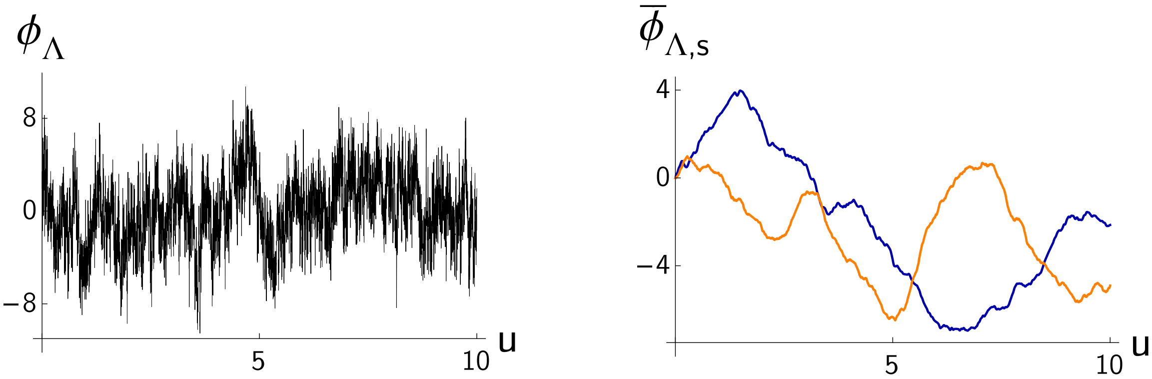

Figure A.2: We show an approximate ϕ and two averages $\bar{\phi}_{\Lambda,s}$ where $\Lambda = 4000$, $L = 5$ and we average over width $s = 0.05$ (blue/dark) and $s = 0.2$ (orange/light). We see that despite the factor of 4 difference in averaging windows, the two appear interchangeable, demonstrating the scale invariance even for this approximate representation. The two appear to follow the same law, and the continuity is seen to be greatly improved.

Simple calculations (see Sec. A.4) show that the Dudley integral J converges, so that $\bar{\phi}_s$ is a.s. continuous in ρ. Moreover we can construct a refinement

$$J\left(\rho(u,v)\right) \equiv \omega(u,v) \approx |u-v| \log \frac{1}{|u-v|} \tag{A.15}$$

with respect to which $\bar{\phi}_s$ is a.s. Lipschitz. This is a "modulus of continuity" for the moving average of a quantum field, analogous to that of Lévy for Brownian motion. (Note that there is no square root, however.)

We can obtain a crude picture of the moving average process by generating noise which has the same power spectrum as a log-correlated field, but with some high frequency cutoff. This is given by the Fourier series

$$\phi_\Lambda(x) = \sum_{m=1}^{\Lambda} \frac{1}{\sqrt{m}} \left[X_m \cos\left(\frac{\pi m x}{L}\right) + Y_m \sin\left(\frac{\pi m x}{L}\right)\right] \tag{A.16}$$

where X_m, Y_m are independent standard Gaussian variables. For large Λ (ultraviolet cutoff) and L (the infrared cutoff, $-L < x < L$) this creates an intuitive "approximation" to the divergent field. Such a technique is often used to visualize white noise. While one must be careful claiming to "approximate" a distribution with a truncated series, ϕ_Λ gives us some sensible object on which to numerically test the properties of our moving average. This is carried out in Fig. A.2. We see that $\bar{\phi}_s$ has the desired properties without requiring the full distribution.

While we focus on the log-correlated field for its mathematical simplicity, it is worth noting that such objects are not necessarily confined to the realm of mathematical fantasy. A free scalar field in one dimension can in principle be a good approximation for a real physical system, with one possible example being the electromagnetic field of certain optical fibers. If the refractive index of the fiber is chosen appropriately, only a finite number of transmission modes will be allowed. We can think of the wave equation as analogous to the Schrödinger equation, with the variable refractive index providing an effective potential. This potential can be chosen to allow only a finite number of bound states. Single-mode fibers have only one such state, leading to a system with one effective spatial dimension.

Even in the absence of light in the fiber (ground state of the electromagnetic field), there will be quantum fluctuations in the potential. In the absence of severe nonlinearities, these fluctuations can be modeled as two noninteracting scalar fields, one for each polarization mode. If the wire is transparent over a sufficient frequency (maintaining its single-mode property and minimal dispersion for propagating waves), then the potential difference between two points will be a Gaussian random variable whose variance is approximately logarithmic with distance. The measurement of the potential would require a probe of finite size, so the averaging process employed in this paper provides a convincing model for the potential as seen by a measuring apparatus at a given instant. The considerations of this paper can be viewed as a model of the spatial regularity of the electromagnetic potential in such an optical fiber. This model could also, in principle, describe the ground state fluctuations of a quantum system confined to a very narrow region in 2 spatial dimensions, sometimes called a quantum wire.

Perhaps an experimental test of the sample field behavior in Fig. A.2 is indeed possible. However, the details of realizing such a system and carrying out such measurements is highly nontrivial and not suited to the themes of this paper; we include this discussion mainly as a reminder that lower dimensional systems are often not so unphysical as they seem.

A.4 Explicit calculations and further examples

A.4.1 Variance of $\bar{\phi}_s$ for the log-correlated field

The calculations that justify the above assertions are straightforward, but worth outlining as they help illuminate the properties discussed above. Because of the divergences, we cannot use the standard approach directly to the quantum field, but only to its moving average. Begin with the observation that

$$F(a,b,c,d) \equiv -\int_a^b dx \int_c^d dy \log|x-y| \tag{A.17}$$

$$= \frac{3}{2}(a-b)(c-d) + \frac{1}{2}\left[L(c-a) - L(d-a) - L(c-b) + L(d-b)\right], \tag{A.18}$$

where L is defined in (A.14). Note this quantity is not quite scale invariant: there is an "anomaly" proportional to $\log \lambda$.

$$L(\lambda r) = \lambda^2 L(r) + r^2 \log \lambda \tag{A.19}$$

$$F(\lambda a, \lambda b, \lambda c, \lambda d) = \lambda^2 F(a,b,c,d) - (b-a)(d-c) \log \lambda \tag{A.20}$$

Then

$$\langle [\bar{\phi}_s(u) - \bar{\phi}_s(v)]^2 \rangle = \frac{F(a,b,a,b)}{(b-a)^2} + \frac{F(c,d,c,d)}{(d-c)^2} - 2\frac{F(a,b,c,d)}{(b-a)(d-c)}, \tag{A.21}$$

where

$$\frac{F(a,b,a,b)}{(b-a)^2} = \frac{3}{2} - \frac{1}{2}\log(b-a)^2 \tag{A.22}$$

which only depends on the width of the interval $[a,b]$. We can then consider two intervals of equal width s, centered at su and sv, yielding

$$\langle [\bar{\phi}_s(u) - \bar{\phi}_s(v)]^2 \rangle = 3 - 2\log(s) - \frac{2}{s^2} F\left(su - \frac{s}{2}, su + \frac{s}{2}, sv - \frac{s}{2}, sv + \frac{s}{2}\right) \tag{A.23}$$

From the scale transformation property above of F we can see that this quantity is independent of s: the "scale anomaly" of F cancels against $2\log s$. So we can simplify by putting $s = 1$ and expressing F in terms of L:

$$\begin{aligned} \langle [\bar{\phi}_s(u) - \bar{\phi}_s(v)]^2 \rangle &\equiv \rho^2(u,v) \\ &= L(u-v-1) + L(u-v+1) - 2L(u-v) \end{aligned} \tag{A.24}$$

as was claimed.

A.4.2 Continuity of Brownian Paths

In using the Dudley integral, it is useful to begin with a well-known example. The most familiar example of a Gaussian process is Wiener's model of Brownian motion, for which $d(x,y) = \sqrt{|x-y|}$. If an interval $[0,1]$ is divided into N equal parts, each part is contained in a $d-$ball of radius $\epsilon = \sqrt{\frac{1}{2N}}$. Thus $N(\epsilon) = 1 + \text{Floor}\left(\frac{1}{2\epsilon^2}\right)$ and for small δ, [where $\text{Floor}(a)$ is the integer part of the real number a]

$$J(\delta) \approx \delta\sqrt{-2\log\delta}. \tag{A.25}$$

Thus Brownian sample paths B are almost surely continuous. More quantitatively, we may construct

$$\omega(r) = J(d(r)) = \sqrt{r\log(1/r)}. \tag{A.26}$$

to obtain the result of Lévy that, with probability one,

$$\frac{|B(x) - B(y)|}{\sqrt{|x-y|\log\frac{1}{|x-y||}}} < C \tag{A.27}$$

as $x \to y$ for some constant C.

A.4.3 Continuity of $\bar{\phi}_s$ for logarithmically correlated fields

We can now show that the sample paths $\bar{\phi}_s(u)$ are continuous with probability one. Again, if $[0,1]$ is divided into N intervals, each will have radius $\epsilon = \rho\left(\frac{1}{N}\right)$. To get small ϵ, we must choose a large N; using the asymptotic behavior

$$\rho(r) \approx r\sqrt{-\log r} \tag{A.28}$$

for small r,

$$\epsilon \approx \frac{1}{N}\sqrt{-\log\left[\frac{1}{N}\right]} \implies N(\epsilon) \approx \frac{1}{\epsilon}\sqrt{-\log \epsilon}. \tag{A.29}$$

The Dudley integral converges:

$$J(\delta) \approx \delta\sqrt{\log\frac{1}{\delta}}, \quad \delta \to 0 \tag{A.30}$$

$$J\left(\rho(r)\right) \equiv \omega(r) \approx r\log(1/r) \tag{A.31}$$

which yields the claimed modulus of continuity.

A.4.4 Additional Examples for Comparison

A.4.4.1 Moving Average of Brownian Paths

It is informative to apply the moving average procedure to the Brownian case, where the paths $B(x)$ which we average over are continuous functions to begin with. Proceeding analogously, consider two intervals with width s with centers su and sv respectively. Then we can define, analogous to (A.17) but with some added foresight,

$$F\left(su-\frac{s}{2},su+\frac{s}{2},sv-\frac{s}{2},sv+\frac{s}{2}\right) \equiv -\int_{su-s/2}^{su+s/2} dx \int_{sv-s/2}^{sv+s/2} dy|x-y| \tag{A.32}$$

$$= \begin{cases} \frac{1}{3}\left(r^3\left(-3s^2+3s-1\right)+3r^2s^3+s^3\right) & 0<r<s \\ s^3 r & r>s, \end{cases} \tag{A.33}$$

where $r=|u-v|$. We then have

$$\begin{aligned} \langle[\bar{B}_s(u)-\bar{B}_s(v)]^2\rangle &\equiv \rho^2(r) && \text{(A.34)}\\ &= \frac{2s}{3}-\frac{2}{s^2}F\left(su-\frac{s}{2},su+\frac{s}{2},sv-\frac{s}{2},sv+\frac{s}{2}\right). && \text{(A.35)} \end{aligned}$$

It is easily seen that $\rho^2(\lambda r)=\lambda\rho^2(r)$, breaking scale invariance. Still for comparison purposes, we consider averaging over intervals of width $s=1$, noting that the scaling behavior will only change $\omega(r)$ by a constant factor. As before, ρ^2 does not define a metric, but its square root ρ does. In the large-r limit we have

$$\rho(r)\sim\sqrt{r}, \tag{A.36}$$

while for small r,

$$\rho(r)\sim r. \tag{A.37}$$

This short distance behavior suggests by dimensional analysis that $\bar{B}_s(x)$ might be Lipschitz in the usual metric $|u-v|$, but the Dudley integral yields a weaker limit

$$\omega(r) \approx r \log\left(\frac{1}{r}\right) \tag{A.38}$$

$$\frac{|\bar{B}_s(u) - \bar{B}_s(v)|}{r \log(\frac{1}{r})} < C. \tag{A.39}$$

Thus $\bar{B}_s$ is just shy of being Lipschitz in the usual metric, but is a.s. Lipschitz with respect to the metric $\omega(u,v) \approx |u-v| \log \frac{1}{|u-v|}$. Interestingly, this is the same ω we obtained in (A.31) for the log-correlated case, even though the short distance behavior of ρ is not quite the same (the difference in the Dudley integral vanishes for small δ). However ω for the Brownian sample paths prior to averaging (A.26) contains a square root not present here.

A.4.4.2 Power Law Correlations in 1D

We can use the same method as with Brownian motion to consider the moving average of a more general power-law correlated field such that

$$\langle \phi[h]\phi[h'] \rangle = \mathrm{sign}(\alpha) \int |x-y|^{\alpha} h(x)h'(y)dxdy, \qquad \int h(x)dx = 0. \tag{A.40}$$

When $\alpha > 0$ this is related to fractional Brownian motion [84]. When $\alpha = -1$ it is the restriction to one dimension of a massless scalar quantum field in $2+1$ dimensions. The moving average is no longer independent of the width of the intervals. Still, for purposes of comparison, we consider the average on intervals of fixed width $s = 1$.

It is not difficult to evaluate the integrals to find that, in the small r limit,

$$\rho^2(r) \sim \begin{cases} r^2 & \alpha > 0 \\ r^{\alpha+2} & -2 < \alpha < 0,\ \alpha \neq -1 \\ r \log r & \alpha = -1. \end{cases} \tag{A.41}$$

The moving average is Lipschitz with respect to the modulus

$$\omega(r) = \begin{cases} r \log(1/r) & \alpha > 0 \\ r^{\frac{\alpha}{2}+1} \log(1/r) & -2 < \alpha < 0. \end{cases} \tag{A.42}$$

When $\alpha \leq -2$ the divergences are such that the moving average is an insufficient tool to smooth the quantum field. Note that $\omega(r)$ is the same in the logarithmic case as the case where $\alpha > 0$. The logarithmic case can be thought of as the critical case where the smoothness implied by Dudley's criterion starts to lessen.

A.4.4.3 Log Correlated Scalar Field in 3D

It is useful to work out a case in higher dimensions as well. The massless scalar field in $n+1$ space-time dimensions has correlation

$$\langle \phi(x)\phi(y) \rangle \propto \frac{1}{|x-y|^{n-1}}\,. \tag{A.43}$$

Thus for $n > 1$ will we get power law, instead of logarithmic correlations. Yet a logarithmically correlated, nonrelativistic, scalar field in 3 space dimensions is still of interest in cosmology [27, 85]. As with the log-correlated scalar field in 1D we must average it over a test function

$$\langle \phi[h]\phi[h'] \rangle = -\int \log|x-y| h(x)h'(y)dx^3dy^3, \qquad \int h(x)dx^3 = 0. \tag{A.44}$$

Recall that

$$-\log|x| + \text{const} = c\int e^{ik\cdot x}\frac{1}{|k|^3}\frac{d^3k}{(2\pi)^3} \tag{A.45}$$

where the constant $c = 2\pi^2$. The integral is not absolutely convergent, so we define it through zeta regularization.

We perform our moving average over the interior of a sphere with radius t, centered at tu

$$\bar{\phi}_t(u) \equiv \int_{|w|\leq 1} \{\phi(t[u-w]) - \phi(t[0-w])\} dw \tag{A.46}$$

$$\langle \bar{\phi}_t(u)\bar{\phi}_t(v) \rangle = \int_{|w|\leq 1} \langle \phi(t[u-w_1])\, \phi(t[v-w_1]) \rangle dw_1dw_2 \tag{A.47}$$

$$= c^2\int \frac{1}{|k|^3}e^{ik\cdot(u-v)t}\frac{d^3k}{(2\pi)^3}\int_{|w|\leq 1} e^{-ik\cdot(w_1-w_2)t}dw_1dw_2. \tag{A.48}$$

Taking $t = 1$, this can be reduced to the form

$$\langle [\bar{\phi}_t(u) - \bar{\phi}_t(v)]^2 \rangle = 2^6\pi^4 G(r) \tag{A.49}$$

where

$$G(r) = \int_0^\infty dk \frac{1}{k^7} \left[\sin k - k\cos k\right]^2 \left[1 - \frac{\sin kr}{kr}\right] \tag{A.50}$$

and $r = |u - v|$. We are not able to evaluate the integral analytically, but its convergence is clear, justifying the independence on t. In the large r limit, the integral is dominated by small k contribution. We then have

$$G(r) \approx \int_0^\infty dk \frac{1}{3k} \left[1 - \frac{\sin kr}{kr}\right] \tag{A.51}$$

$$\sim \log(r) + O(1). \tag{A.52}$$

In the case of small r, the dominant contribution comes from the first peak of $\frac{1}{k^7}\left[\sin k - k\cos k\right]^2$, which must occur for $k < 2\pi$ (i.e., $k \approx 5.678$). This allows us to treat kr as small, yielding the behavior

$$G(r) \approx r^2 \int_0^\infty dk \frac{1}{3k^5} \left[\sin k - k\cos k\right]^2 \ \sim r^2. \tag{A.53}$$

The approximation can be verified numerically. This small r behavior dictates the continuity modulus discussed above. Namely, we have that for the 3D log-correlated scalar field,

$$\rho(r) \ \sim \ r, \tag{A.54}$$

$$\omega(r) \ \sim \ r\log(1/r). \tag{A.55}$$

A similar metric can be obtained for a log correlated field in other dimensions. Note that, once we have $\rho(r) \sim r$ for small r, the logarithm in the Dudley integral ensures that ω will not depend on the dimensionality (up to proportionality). This is not true if $\rho(r)$ has some other short distance behavior.

A.5 Outlook

Gaussian processes correspond to free fields. The most elegant way to introduce interactions into a scalar field theory is to let it take values in a curved Riemannian manifold. This is the nonlinear sigma model in physics language, or the wave map in the mathematical literature. In 1+1 dimensions, such a theory, with a target space of a sphere or a compact Lie group, is well studied in the physics literature. The short distance behavior is approximated by free fields with corrections computable in perturbation theory (asymptotic freedom). The only case for which mathematically rigorous results are known is that of the Wess-Zumino-Witten model, which has non-Gaussian behavior at short distances; i.e., a "nontrivial fixed point" for the renormalization group. The related measure for the ground state of the quantum field has been constructed by Pickrell. (For a review, see [86]). It is natural to ask for regularity results analogous to ours in this case.

Looking further out, it would be interesting to quantify the regularity of quantum fields of the nonlinear sigma model in two dimensional space time; and even further out, $\lambda\phi^4$ theory in four dimensions. It is possible that the "naturalness problem" of the standard model of particle physics has a resolution in terms of such a deeper understanding of the regularity of scalar quantum field theory. The "modern" theory [81] of regularity of non-Gaussian processes ought to help with this daunting task. Even harder is the case of Yang-Mills fields. An analogue of our moving average is the Wilson loop. The measure of integration over the space of gauge fields is only known rigorously for the two dimensional case [87]. Regularity of Yang-Mills fields satisfying classical evolution equations (let alone random processes) is already a formidable problem under active investigation (see for example [88]).

Appendix B

Highly nonlinear wave solutions in a dual to the chiral model

Reprinted from Ref. [32] with permission. Co-authored with S.G. Rajeev.

B.1 Introduction

We study the field theory [89–91] with equations of motion

$$\ddot{\phi} = \lambda \left[\dot{\phi}, \phi'\right] + \phi'', \tag{B.1}$$

where ϕ is valued in a Lie algebra, $\phi : \mathbb{R}^{1,1} \to \mathfrak{su}(2)$. This follows from the action

$$S_1 \equiv \int \mathcal{L}_1 dxdt = \int \mathrm{Tr}\left\{\frac{1}{2\lambda}\dot{\phi}^2 - \frac{1}{2\lambda}{\phi'}^2 + \frac{1}{3}\phi[\dot{\phi}, \phi']\right\} dxdt. \tag{B.2}$$

In the $\lambda \to 0$ limit, these equations admit linear wave solutions. But in the high-coupling regime, the theory is dominated by nonlinear effects.

S_1 is closely tied to other models and subjects, which we elaborate on in section B.2. These include the study of slow light, the Wess-Zumino-Witten (WZW) model, and the mathematical theory of hypoelliptic operators. Of particular interest in this paper, the model described by S_1 is also classically dual to the well-studied principal chiral model, described by the action

$$S_2 = \int \mathcal{L}_2 dxdt = \frac{1}{2f} \int \mathrm{Tr} \left\{ \left(g^{-1}\dot{g}\right)^2 - c^2 \left(g^{-1}g'\right)^2 \right\} dxdt \tag{B.3}$$

where $g : \mathbb{R}^{1,1} \to SU(2)$. This is a special case of the nonlinear sigma model, with target space $SU(2)$.

Despite their classical equivalence, S_1 and S_2 lead to entirely different quantum theories. S_2 gives an asymptotically free theory: at short distances $f \to 0$, giving us free massless excitations. But the true particles that survive to long distances are bound states of non-zero mass [92, 93]. For this reason, the principal chiral model is often used as a toy model for 4-dimensional Yang-Mills theory, notorious for its mathematical complexity. Not only do the two theories share similar short-distance behavior, but the existence of a mass gap in the principal chiral model has served as a proof of concept for the conjectured mass gap in Yang-Mills (though neither can yet be proven with full mathematical rigor).

S_1, on the other hand, leads to a renormalizable but not asymptotically free quantum theory. At short distances the coupling constant $\lambda \to \infty$, while at long distances we have weakly nonlinear massless excitations. It makes sense to use S_1 as a 2-D toy model for strongly coupled theories, in particular 4-dimensional $\lambda\phi^4$

theory[1]. The behavior of quantum field theories at high coupling is notoriously intractable, and the physical meaning of such theories is still up for debate. For this reason, it is still necessary to search for simple examples of such theories and try to gleam what meaning, if any, they have in the short distance limit.

In addition to sharing short-distance behavior, both the S_1 model and $\lambda\phi^4$ theory can be described by hypoelliptic hamiltonian operators with a step-3 nilpotent bracket algebra, suggesting some algebraic structure in common (see Section B.2.2 and the supplementary material in the original publication of Ref. [32] for more detail). The S_1 model's relative simplicity makes it a good candidate for attempting to probe the high coupling regime of field theories in general, but the connection to $\lambda\phi^4$ theory seems the closest. Additionaly, its classical duality to the principal chiral model motivates a juxtaposition of the two theories in the classical and quantum formulations.

To glimpse what becomes of our theory in the high coupling limit, we take the modest approach of finding nonlinear wave-type solutions to the classical model which survive the $\lambda \to \infty$ limit (section B.3). This set of solutions defines a mechanical system or "reduced system" in each of the dual models. While they physically appear very different, the resulting classical solutions can be mapped from one system to another. We quantize these collective variables to determine their dispersion relation (section B.4) in the short distance limit for each theory. We have in mind the sine-Gordon theory, whose solitons turns out to be the fundamental constituents which bind to form the scalar particles [94, 95].

These reduced quantum theories yield two different results. In particular, the reduced model of S_1 has an exotic dispersion relation in the short distance limit. We

[1] that is, pure $\lambda\phi^4$ theory, describing a Higgs-like particle with no coupling to fermions

postulate that its spectrum may hint at the fundamental constituents of the highly coupled theory, which need not behave like traditional particles at all. In section B.5 we offer concluding remarks and a side-by-side comparison of our work with S_1 and S_2.

On The Notation

We regard $\phi = \frac{1}{2i}\left[\phi_1\sigma_1 + \phi_2\sigma_2 + \phi_3\sigma_3\right] = \frac{1}{2i}\boldsymbol{\phi}\cdot\boldsymbol{\sigma}$ as a traceless anti-hermitian matrix. Recall then that the commutator and cross product are related by

$$[X, Y] = \frac{1}{2i}\left(\mathbf{X}\times\mathbf{Y}\right)\cdot\boldsymbol{\sigma}. \tag{B.4}$$

Also, we define $\mathrm{Tr}X \equiv -2\mathrm{tr}X$ so that

$$\mathrm{Tr}\phi^2 = \phi_1^2 + \phi_2^2 + \phi_3^2. \tag{B.5}$$

In relativistically invariant notation, (B.1) and (B.2) can be written as

$$\partial^\mu\partial_\mu\phi_a - \frac{\lambda}{2}\,\epsilon_{abc}\epsilon^{\mu\nu}\partial_\mu\phi^b\partial_\nu\phi^c = 0, \tag{B.6}$$

$$S_1 = \frac{1}{2\lambda}\int \partial_\mu\phi^a\partial_\nu\phi^a\eta^{\mu\nu}d^2x + \frac{1}{6}\int \epsilon_{abc}\phi^a\partial_\mu\phi^b\partial_\nu\phi^c\epsilon^{\mu\nu}d^2x \tag{B.7}$$

where $\mu, \nu = 0, 1$ and $a, b, c = 1, 2, 3$; also, $\epsilon^{\mu\nu}, \epsilon_{abc}$ are the Levi-Civita tensors. This is a particular case of the general sigma model studied in [96] as the background of string theory, with a flat metric on the target space and a constant 3-form field ϵ_{abc}.

B.2 relation to other models

B.2.1 The $c \to 0$ Limit and Slow Light

Consider the equations of motion (B.1) where the speed of linear propagation at low coupling is taken to be c rather than 1:

$$\ddot{\phi} = \lambda \left[\dot{\phi}, \phi'\right] + c^2 \phi''. \tag{B.8}$$

If we rescale $\phi \to \lambda^a \phi$, $t \to \lambda^b t$, this becomes

$$\lambda^{a-2b} \ddot{\phi} = \lambda^{1+2a-b} \dot{\phi} \times \phi' + c^2 \lambda^a \phi''. \tag{B.9}$$

Set $a = 2b$ and $1 + 2a = b$ to get

$$\ddot{\phi} = \dot{\phi} \times \phi' + c^2 \lambda^{-\frac{2}{3}} \phi''. \tag{B.10}$$

Thus the strong coupling limit $\lambda \to \infty$ at fixed c is equivalent to the limit $c \to 0$ with $\lambda = 1$:

$$\ddot{\phi} = \dot{\phi} \times \phi'. \tag{B.11}$$

The strongly coupled limit can be thought of as the limit in which the waves move very slowly. It has been noted in that literature [97] that when the speed of light in a medium is small, nonlinear effects are magnified. Although the specific equations appearing there are different, it is possible that the solutions of the sort we study are of interest in that context as well.

From a field theoretic context, the equivalence of these limits seems troubling. At short distances, the highly-coupled theory will not be relativistic. It is a sort of "post-relativistic" regime, where $c \to 0$. This is much the opposite of the case in the theory of S_2; there the short-distance excitations are massless, but form massive bounds states which survive to long distances and can be non-relativistic in the traditional $c \to \infty$ sense. Perhaps some exotic excitations at high coupling are in fact the fundamental constituents in the S_1-model, forming as bound states the ordinary massless particles which appear in the long distance limit. As we know from the quark model, the short distance excitations do not need to be particles in the usual sense; they could be confined. In any case, it is important to know what solutions might survive the high coupling limit, whether they be unphysical or simply unintuitive.

We will see an example of wave solutions which classically survive the $c \to 0$ limit, continuing to propagate through nonlinearity alone. Since the energy density is constant, these solutions do not violate causality: they are analogous to the Continuous Wave solutions in a medium where the phase velocity is greater than c.

B.2.2 Sub-Riemannian Geometry and the Strong Coupling Limit

Many physical problems (Yang-Mills, Fluid Mechanics) become intractable in the strong coupling limit where the non-linearities dominate. It would be nice to have a unified geometric approach to understanding these systems. We have such an approach in the weak coupling limit: small perturbations around a stable equilibrium are equivalent to a harmonic oscillator.

A larger picture emerges if we think in terms of Riemannian and sub-Riemannian geometry. The orbits of many mechanical systems of physical interest (again, Yang-Mills or incompressible Fluids) can be thought of as geodesics in some appropriate Riemannian manifold. In the simplest case, the harmonic oscillator describes geodesics in the Heisenberg group. The anharmonic oscillator (and many nonlinear field theories with quartic coupling) can also be thought of as geodesic motion on a nilpotent Lie group, by introducing an additional generators (see the supplementary material in the full publication of Ref. [32] for additional discussion).

In the limit of strong coupling, the metric degenerates and becomes sub-Riemannian [98]. That is, the contravariant metric tensor has some zero eigenvalues so that it can be written as $\sum_j X_j \otimes X_j$ for some vector fields X_j whose linear span may be smaller than the tangent space. Moreover, in the cases of interest, these vector fields satisfy the celebrated Hörmander condition: X_j along with their repeated commutators span the tangent spaces at every point. In such a case, there are still geodesics connecting every pair of sufficiently close points (Chow-Rashevskii theorem, [98]). Thus, we can define a distance between pairs of points as the shortest length of geodesics.

These ideas came to the notice of many physicists following a model for the self-propulsion of an amoeba [99], though they have roots in the Carnot-Caratheodory geometric formalism of thermodynamics and in control theory. Hörmander [100] discovered independently that the same criterion is sufficient for the sub-Riemannian Laplace operator $\Delta = \sum_j X_j^2$ to be *hypoelliptic*, meaning the solution f to the inhomogenous equation $\Delta f = u$ is smooth whenever the source u is smooth. This can be thought of the quantum version of the above condition on subgeodesic connectivity.

This kind of sub-Riemannian geometry may present a powerful geometric framework for strongly-coupled field theories. The example we work out in this paper is arguably the simplest interesting case of a strongly coupled field theory, and the solutions we study correspond to sub-Riemannian geodesics in the limit $\lambda \to \infty$. We hope to apply such geometric ideas to other cases in the future, using this as a prototype.

B.2.3 Relation to the WZW model

We can also regard our equations as a limiting case of the Wess-Zumino-Witten model[2] [101]

$$S_{WZW} = \frac{1}{4\lambda_1^2} \int \mathrm{tr}\, \partial^\mu g \partial_\mu g^{-1} d^2x + \frac{n}{24\pi} \int_{M_3} \mathrm{tr}\, (g^{-1}dg)^3 \tag{B.12}$$

as $n \to \infty$ and $\lambda_1 \to 0$, keeping $\lambda = \lambda_1^2 (n/2\pi)^{\frac{2}{3}}$ fixed[3]. To see this, let $g(x) = e^{bi\sigma_a \phi^a(x)}$ and expand in powers of b:

$$S_{WZW} = \frac{b^2}{2\lambda_1^2} \int \partial^\mu \phi^a \partial_\mu \phi^a + \frac{n}{24\pi} b^3 \int_{M_3} 2\epsilon_{abc}\, d\phi^a d\phi^b d\phi^c \;+ \cdots \tag{B.13}$$

To this order the WZW term is an exact differential, so we can write it as an integral over space-time

$$S_{WZW} = \frac{1}{2} \frac{b^2}{\lambda_1^2} \int \partial^\mu \phi^a \partial_\mu \phi^a + \frac{n}{12\pi} b^3 \int \epsilon_{abc}\, \phi^a d\phi^b d\phi^c \;+ \cdots \tag{B.14}$$

[2]Witten's Tr is our tr. His λ is our λ_1.

[3]Here, M_3 is a 3-manifold of which the two-dimensional space-time is the boundary. We do not require λ_1^2 to take the conformally invariant value $\frac{4\pi}{n}$.

S_{WZW} reduces to S_1 if we identify $b^3 = 2\pi/n$ and λ as above. By taking this limit, we can easily get the renormalization of our model. Recall that [101] the one loop renormalization group equation of the $O(N)$ WZW model is

$$\frac{d\lambda_1^2}{d\log\Lambda} = -\frac{\lambda_1^4(N-2)}{2\pi}\left[1-\left(\frac{\lambda_1^2 n}{4\pi}\right)^2\right] \tag{B.15}$$

We need the particular case of $N = 4$ corresponding to the target space being $S^3 \approx SU(2)$. Thus in our limit $n \to \infty$, $\lambda_1 \to 0$ keeping λ fixed,

$$\frac{d\lambda}{d\log\Lambda} = \frac{\lambda^4}{4\pi} \tag{B.16}$$

It is useful to take this limit rather than calculating loop corrections from scratch, as the renormalization group evolution of the WZW has been studied to high order [102,103]. Including these higher order terms does not alter the short-distance divergence of λ.

B.2.4 Duality with the principal chiral model

We have now seen that the S_1 model is strongly coupled in the short-distance limit. Yet, as a classical field theory, it can be viewed [89,90] as a dual to the asymptotically free principal chiral model with equation of motion

$$\partial^\mu[g^{-1}\partial_\mu g] = 0, \qquad g : \mathbb{R}^{1,1} \to SU(2). \tag{B.17}$$

To see this, we define the currents

$$I = \frac{1}{\lambda}\dot{\phi}, \quad J = \phi' \tag{B.18}$$

so the equations of motion become

$$\dot{J} = \lambda I', \quad \dot{I} = \lambda [I, J] + \frac{1}{\lambda} J'. \tag{B.19}$$

We can solve the second equation with the relations

$$I = \frac{1}{\lambda^2} g^{-1} g', \quad J = \frac{1}{\lambda} g^{-1} \dot{g}. \tag{B.20}$$

Then the first equation becomes

$$\partial_0 \left[g^{-1} \dot{g} \right] = \partial_1 \left[g^{-1} g' \right] \tag{B.21}$$

which is the non-linear sigma model. Thus, the same classical equations of motion follow from the action

$$S_2 = \frac{1}{2f} \int \mathrm{Tr} \left\{ \left(g^{-1} \dot{g} \right)^2 - c^2 \left(g^{-1} g' \right)^2 \right\} dx dt \tag{B.22}$$

if we identify $f = \lambda^2$. A summary of relevant correspondences in the dual models can be found in table B.1 at the end of section B.5.

We also briefly note that our theory is closely related to the sigma model on the Heisenberg group (see [104]).

B.3 Reduction to A Mechanical System

We will look at propagating waves of the form

$$[compact]\phi(t,x) = e^{Kx}R(t)e^{-Kx} + mKx, \quad K = \frac{i}{2}k \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{B.23}$$

for constants k, m. These solutions are equivariant under translations: the "potential" ϕ changes by an internal rotation and a constant shift under translation, while the currents only change only by the internal rotation. Thus, the energy density is constant. They are to be contrasted with soliton solutions, which have energy density concentrated at the location of the soliton. They are more analogous to the plane wave solutions of the wave equation, or a Continous Wave (CW) laser beam. Moreover, the currents

$$I = \frac{1}{\lambda}e^{Kx}\dot{R}e^{-Kx}, \quad J = e^{Kx}\left\{[K,R] + mK\right\}e^{-Kx} \tag{B.24}$$

are periodic in space with wavelength $\frac{2\pi}{k}$. Defining

$$L \equiv [K,R] + mK, \quad S \equiv \dot{R} + \frac{1}{\lambda}K, \tag{B.25}$$

We can write the equations of motion and identity (B.19) in a symmetric form

$$\dot{L} = [K,S], \quad \dot{S} = \lambda[S,L]. \tag{B.26}$$

This new choice of variables will allow us to connect to the dual theory, identify the conserved quantities and to pass to the quantum theory more easily.

B.3.1 The reduced system Lagrangian

Three conserved quantities follow immediately:

$$
\begin{aligned}
s^2k^2 &\equiv \mathrm{Tr}\, S^2 \\
C_1k^2 &\equiv \mathrm{Tr} SL \\
C_2k^2 &\equiv \mathrm{Tr}\left[\frac{1}{2}L^2 - \frac{1}{\lambda}KS\right].
\end{aligned} \tag{B.27}
$$

The quantity s will be of importance in the dual picture, while the other constants have less obvious roles there. Moreover, we have the identity

$$\mathrm{Tr} KL = mk^2. \tag{B.28}$$

Of the six independent variables in S and L, only two remain after taking into account these constants of motion. The dynamics are described by the effective lagrangian density (dropping a total time derivative and an overall factor of volume of space divided by λ)

$$
\begin{aligned}
\mathcal{L}_1 &= \mathrm{Tr}\left\{\frac{1}{2}\dot{R}^2 + \frac{\lambda}{3}R\left[\dot{R},\ [K,R]+mK\right] - \frac{1}{2}\left([K,R]+mK\right)^2\right\} \\
&= \mathrm{Tr}\left\{\frac{1}{2}(S-\frac{1}{\lambda}K)^2 + \frac{\lambda}{3}R\left[S-\frac{1}{\lambda}K,\ L\right] - \frac{1}{2}L^2\right\}
\end{aligned} \tag{B.29}
$$

and hamiltonian density

$$H_1 = \mathrm{Tr}\left[\frac{1}{2}\left(S-\frac{1}{\lambda}K\right)^2 + \frac{1}{2}L^2\right], \tag{B.30}$$

B.3.2 Reduction to One Degree of Freedom

It is useful to work with the first two components of R as a single complex variable. Defining $Z = R_1 + iR_2$, we can write explicitly

$$L = \frac{k}{2}\begin{pmatrix} im & \bar{Z} \\ -Z & -im \end{pmatrix}. \tag{B.31}$$

To describe the third component, we define

$$u \equiv \frac{1}{k}\dot{R}_3 - \frac{1}{\lambda}, \tag{B.32}$$

allowing us to write a similarly compact expression for S,

$$S = \frac{1}{2i}\begin{pmatrix} uk & \dot{\bar{Z}} \\ \dot{Z} & -uk \end{pmatrix}. \tag{B.33}$$

The three conserved quantities (B.27) can now be written in terms of Z and u as

$$\begin{aligned} s^2k^2 &= u^2k^2 + |\dot{Z}|^2 \\ C_1k^2 &= \frac{ik}{2}\left[\bar{Z}\dot{Z} - \dot{\bar{Z}}Z\right] - mk^2u \\ C_2k^2 &= \frac{k^2}{2}\left(m^2 + \frac{2u}{\lambda} + |Z|^2\right). \end{aligned} \tag{B.34}$$

Using the identity

$$\left(\frac{d}{dt}|Z|^2\right)^2 = 4|Z|^2|\dot{Z}|^2 + (\dot{\bar{Z}}Z - \bar{Z}\dot{Z})^2, \tag{B.35}$$

we can combine these three equations to eliminate Z and yield an ODE for $u(t)$,

$$\dot{u}^2 = k^2\lambda^2 \left\{ \left[2C_2 - m^2 - \frac{2}{\lambda}u\right](s^2 - u^2) - [mu + C_1]^2 \right\}. \tag{B.36}$$

B.3.3 Solution in terms of elliptic functions

The ODE for $u(t)$ describes an elliptic curve. Setting $u = av + b$, we can pick the constants

$$a = \frac{2}{k^2\lambda}, \quad b = \frac{C_2\lambda}{3} \tag{B.37}$$

to bring our ODE to Weierstrass normal form in terms of v:

$$\dot{v}^2 = 4v^3 - g_2 v - g_3. \tag{B.38}$$

The somewhat unsightly expressions for g_2 and g_3 can be obtained by symbolic computation:

$$g_2 = \frac{1}{3}k^4\lambda^2\left(3C_1\lambda m + C_2^2\lambda^2 + 3s^2\right) \tag{B.39}$$

$$g_3 = \frac{1}{108}k^6\lambda^4(27C_1^2 + 18C_1C_2\lambda m + 4C_2^3\lambda^2 - 36C_2 s + 27m^2s^2) \tag{B.40}$$

The solution to the Weierstrass differential equation (B.38) is then

$$v(t) = \wp(t+\alpha) \implies u(t) = \frac{2}{k^2\lambda}\wp(t+\alpha) + \frac{C_2\lambda}{3}, \tag{B.41}$$

where $\wp$ is the Weierstrass P-function and α is a complex constant determined by the initial conditions. We can most immediately solve for $R_3(t)$. Recalling (B.32), we have

$$\dot{R}_3 = \frac{2}{k\lambda}\wp(t+\alpha) + k\left(\frac{C_2\lambda}{3} + \frac{1}{\lambda}\right). \tag{B.42}$$

In order for to obtain a sensible solution, $\wp(t+\alpha)$ must be real and bounded. This requires $\mathrm{Im}(\alpha) = |\omega_2|$, where ω_2 is the imaginary half-period of the Weierstrass P-function (which depends on the elliptic invariants g_2, g_3). The real part of α merely shifts our solution in time, so we can take $\alpha = \omega_2$ for simplicity. Using the relationship

$$\int \wp(u)du = -\zeta(u), \tag{B.43}$$

where ζ is the Weierstrass ζ-function, and taking $R_3(0) = 0$ gives the solution

$$R_3(t) = \frac{2}{k\lambda}\left[\zeta(\omega_2) - \zeta(t+\omega_2)\right] + \left(\frac{C_2\lambda}{3} + \frac{1}{\lambda}\right)kt. \tag{B.44}$$

The solution for the other two components is found by making the substitution $Z = re^{i\theta}$ in (B.34). Writing $|Z^2| = r^2$ quickly yields

$$r^2(t) = \frac{4}{3}C_2 - m^2 - \frac{4}{k^2\lambda^2}\wp(t+\omega_2). \tag{B.45}$$

Note that the choice of $\mathrm{Re}(\alpha) = 0$ we made earlier implies that $t = 0$ is a turning point of the radial variable, as $\wp'(\omega_2)$ is necessarily 0. It is useful to write

$$r^2(t) = \frac{4}{k^2\lambda^2}[\wp(\Omega) - \wp(t+\omega_2)], \tag{B.46}$$

where

$$\wp(\Omega) = k^2\lambda^2 \left(\frac{C_2}{3} - \frac{m^2}{4} \right). \tag{B.47}$$

Then we can use the identity

$$\wp(z) - \wp(\Omega) = -\frac{\sigma(z+\Omega)\sigma(z-\Omega)}{\sigma^2(z)\sigma^2(\Omega)}, \tag{B.48}$$

where σ is the Weierstrass σ-function, in order to simplify a later result. We obtain the solution

$$r(t) = \frac{2}{\lambda k \sigma(\Omega)} \frac{\sqrt{\sigma(t+\omega_2+\Omega)\sigma(t+\omega_2-\Omega)}}{\sigma(t+\omega_2)} \tag{B.49}$$

To find $\theta(t)$ from (B.34), we substitute $(\bar{Z}Z - \bar{Z}\dot{Z}) = -2ir^2\dot{\theta}$, obtaining

$$\dot{\theta} = \frac{C_3}{r^2} + \frac{km\lambda}{2} = \frac{k^2\lambda^2 C_3}{4[\wp(\Omega) - \wp(t+\omega_2)]} + \frac{km\lambda}{2}, \tag{B.50}$$

where

$$C_3 \equiv k \left[\frac{m^3\lambda}{2} - C_1 - m\lambda C_2 \right] \tag{B.51}$$

Using the identity

$$\int \frac{dz}{\wp(z) - \wp(\Omega)} = \frac{1}{\wp'(\Omega)} \left[2z\zeta(\Omega) + \log \frac{\sigma(z-\Omega)}{\sigma(z+\Omega)} \right] \tag{B.52}$$

and taking $\theta(0) = 0$, we have

$$\theta(t) = \frac{k^2\lambda^2 C_3}{4\wp'(\Omega)}\left[2t\zeta(\Omega) + \log\frac{\sigma(t+\omega_2-\Omega)\sigma(\omega_2+\Omega)}{\sigma(t+\omega_2+\Omega)\sigma(\omega_2-\Omega)}\right] + \frac{km\lambda}{2}t \tag{B.53}$$

We can use the Weierstrass differential equation (B.38) directly to obtain $\wp'(\Omega) = (i/2)k^2\lambda^2 C_3$, leading to a seemingly remarkable cancellation. We then have

$$e^{i\theta(t)} = \sqrt{\frac{\sigma(t+\omega_2+\Omega)\sigma(\omega_2-\Omega)}{\sigma(t+\omega_2-\Omega)\sigma(\omega_2+\Omega)}} \cdot \exp\left[-\left(\zeta(\Omega) + \frac{ikm\lambda}{2}\right)t\right]. \tag{B.54}$$

Finally, a few terms cancel in the overall expression for Z, yielding

$$Z(t) = \left[\frac{2}{\lambda k\sigma(\Omega)}\sqrt{\frac{\sigma(\omega_2-\Omega)}{\sigma(\omega_2+\Omega)}}\right]\frac{\sigma(t+\omega_2+\Omega)}{\sigma(t+\omega_2)} \cdot \exp\left[-\left(\zeta(\Omega) + \frac{ikm\lambda}{2}\right)t\right]. \tag{B.55}$$

A sample solution is plotted in Fig. B.1. We can see that, in the R_1-R_2 plane, the solution traces an oscillating curve in between some inner and outer radius. Meanwhile, the solution propagates in the R_3 direction with non-uniform speed. This behavior is typical over all parameter values we tested.

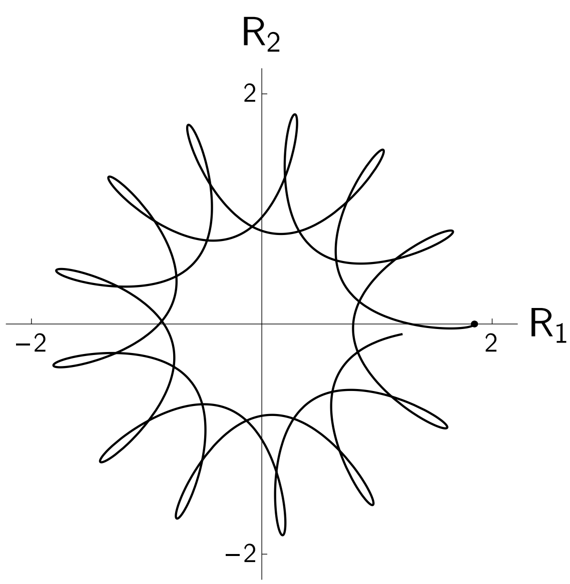

Figure B.1: The orbit in the R_1-R_2 plane (above) and the evolution of R_3 with time (below). The sample solution is plotted for $0 < t < 21$ and uses parameters $k = 1$, $\lambda = 2.4$, $C_1 = 0.5$, $C_2 = 1$, $s = 2$, $m = 0.5$.

B.4 Mechanical Interpretation and Quantization of the Reduced Systems

The equations of motion following from the ansatz (B.26), defining the reduced system for S_1, can be written as

$$\ddot{R} = \lambda[\dot{R}, [K, R] + mK] + [K, [K, R]]. \tag{B.56}$$

These are the equations of motion of a particle in a static electromagnetic field, given by (working in cylindrical polar coordinates where $R_1 = r\cos\theta,\ R_2 = r\sin\theta,\ z = R_3$)

$$\vec{B} = kr\hat{\theta} + mk\hat{z}, \quad \vec{E} = k^2 r\hat{r}, \tag{B.57}$$

which follow from the vector and scalar potentials

$$\vec{A} = \frac{\lambda k}{2}\left(mr\hat{\theta} + r^2\hat{z}\right), \quad V = \frac{k^2}{2}r^2. \tag{B.58}$$

The classical Hamiltonian is then

$$H_1 = \frac{1}{2}p_r^2 + \frac{1}{2}\frac{\left[p_\theta - A_\theta\right]^2}{r^2} + \frac{1}{2}\left[p_z - A_z\right]^2 + V(r). \tag{B.59}$$

It is clear that p_θ and p_z are conserved. This formulation lends some physical intuition to the solutions found in section B.3. We can pass to the quantum theory as usual by finding the covariant Laplacian in cylindrical co-ordinates,

$$\hat{H}_1\psi = -\frac{\hbar^2}{2}\frac{1}{r}\frac{\partial}{\partial r}\left[r\frac{\partial\psi}{\partial r}\right] + \frac{1}{2r^2}\left[-i\hbar\partial_\theta - A_\theta\right]^2\psi + \frac{1}{2}\left[-i\hbar\partial_z - A_z\right]^2\psi + V(r)\psi. \tag{B.60}$$

The conservation of $p_\theta,\ p_z$ leads us to seek a solution to the Schrodinger equation of the separable form

$$\psi(r,\theta,z) = \frac{1}{\sqrt{r}}\rho(r)e^{il\theta}e^{i\frac{p_z z}{\hbar}} \tag{B.61}$$

for integer $l = \frac{p_\theta}{\hbar}$. The system is then reduced to a one-dimensional Schrodinger equation

$$-\frac{\hbar^2\rho''(r)}{2} + U(r)\rho(r) = E\rho(r) \tag{B.62}$$

with effective potential

$$U(r) = -\frac{\hbar^2}{8r^2} + \frac{1}{2}\frac{[\hbar l - A_\theta]^2}{r^2} + \frac{1}{2}[p_z - A_z]^2 + V(r)$$

$$= \frac{1}{2}\left[\frac{\hbar^2\left[l^2 - \frac{1}{4}\right]}{r^2} - \frac{\hbar k\lambda ml}{r} + \left(\frac{k^2\lambda^2 m^2}{4} + p_z^2\right) + (k - \lambda p_z)kr^2 + \frac{\lambda^2 k^2 r^4}{4}\right]. \quad \text{(B.63)}$$

At low coupling ($\lambda \to 0$), we have

$$-\hbar^2\rho'' + \left\{\frac{\hbar^2\left[l^2 - \frac{1}{4}\right]}{r^2} + p_z^2 + \frac{k^2}{2}r^2\right\}\rho = E\rho, \quad \text{(B.64)}$$

and we see by dimensional analysis

$$[\hbar k] = 1/L^2 \quad \Rightarrow E \sim |\hbar k|. \quad \text{(B.65)}$$

These are weakly coupled massless excitations. But in the high coupling limit ($\lambda \to \infty$), we have

$$-\hbar^2\rho'' + \left\{\frac{k^2\lambda^2}{4}(m^2 + r^4)\right\}\rho = E\rho, \quad \text{(B.66)}$$

which yields a much more peculiar spectrum

$$[\hbar^2 k\lambda]^2 = 1/L^6 \quad \Rightarrow E \sim |\hbar^2\lambda k|^{2/3}. \quad \text{(B.67)}$$

If this dispersion relation describes some fundamental constituents of the theory, then they are certainly not particles in the traditional sense. We propose that this may be a glimpse of some post-relativistic constituents as mentioned in section B.2.1.

B.4.1 quantization of the dual reduced system

In the dual picture (nonlinear sigma model), our ansatz picks out a class of solutions that correspond to a different mechanical system. Though the equations of motion in each picture can be mapped to one another via the duality, the correspondence is not immediately obvious, and the systems will appear very different upon quantization.

After the ansatz, the duality relations (B.20) read

$$g^{-1}g' = \lambda e^{Kx}\left(S + \frac{1}{\lambda}K\right)e^{-Kx}, \quad g^{-1}\dot{g} = \lambda e^{Kx}Le^{-Kx}. \tag{B.68}$$

Writing $g = h(t,x)e^{-Kx}$ yields

$$h^{-1}h' = \lambda S \quad h^{-1}\dot{h} = \lambda L. \tag{B.69}$$

We further suppose that h is separable as $h(t,x) = F(x)Q(t)$. Then the equation for S can be separated as

$$F^{-1}(x)F'(x) = \lambda Q(t)S(t)Q^{-1}(t) \tag{B.70}$$

Both sides are equal to some constant traceless matrix C. Since $Q(t)$ is only unique up to multiplication on the left by a constant matrix in $\mathbb{SU}(2)$, we can use this to choose C to be diagonal and thus proportional to K. We then have

$$Q(t)S(t)Q^{-1}(t) = sK, \tag{B.71}$$

implying that $\mathrm{Tr}S^2 = s^2k^2$. (B.70) is satisfied if

$$F(x) = e^{\lambda sKx}. \tag{B.72}$$

Thus, the full corresponding ansatz for the field variable in the dual theory is

$$g(t,x) = e^{\lambda sKx}Q(t)e^{-Kx}, \tag{B.73}$$

where Q is related to the previous variables by

$$S = sQ^{-1}(t)KQ(t), \quad L = \frac{1}{\lambda}Q^{-1}\dot{Q}. \tag{B.74}$$

The dual Lagrangian can now be written as

$$\mathcal{L}_2 = \frac{1}{2f^2}\mathrm{Tr}\left[\left(Q^{-1}\dot{Q}\right)^2 - \left(\lambda sQ^{-1}KQ - K\right)^2\right]. \tag{B.75}$$

It is useful to parameterize Q in terms of the Euler angles:

$$Q = e^{\frac{i}{2}\sigma_3\gamma}e^{\frac{i}{2}\sigma_1\beta}e^{\frac{i}{2}\sigma_3\alpha}. \tag{B.76}$$

The traces in $\mathcal{L}_2$ can then be computed directly, yielding

$$\mathcal{L}_2 = \frac{1}{\lambda^2}\left\{\frac{\dot{\alpha}^2+\dot{\beta}^2+\dot{\gamma}^2}{2} + \cos\beta\dot{\alpha}\dot{\gamma} - V(\beta)\right\} \tag{B.77}$$

where (dropping a constant shift)

$$V(\beta) = -2k^2 \lambda s \cos\beta. \tag{B.78}$$

As a mechanical system, this is the well known spinning top (isotropic Lagrange top). It is instructive to write

$$\mathcal{L}_2 = \frac{1}{\lambda^2}\left[\frac{1}{2} g_{ij}\dot{\alpha}^i \dot{\alpha}^j - V\right] \tag{B.79}$$

where

$$g_{ij} = \begin{pmatrix} 1 & 0 & \cos\beta \\ 0 & 1 & 0 \\ \cos\beta & 0 & 1 \end{pmatrix} \tag{B.80}$$

is the metric of the rotation group and V is the gravitational potential of the top. The overall constant $\frac{1}{\lambda^2}$ in the action leads to a rescaling of $\hbar \mapsto \hbar\lambda^2$ upon quantization.

To pass to the quantum theory, we find the Laplacian operator with respect to the metric g of Eulerian coordinates,

$$\nabla^2 \psi = \frac{1}{\sqrt{g}} \partial_i \left[\sqrt{g} g^{ij} \partial_j \psi\right] \tag{B.81}$$

The Hamiltonian is then

$$\hat{H}_2 = -\frac{\hbar^2 \lambda^4}{2}\left[\frac{\partial_\alpha^2 + \partial_\gamma^2 - 2\cos\beta \partial_\alpha \partial_\gamma}{\sin^2\beta} + \partial_\beta^2 + \cot\beta \partial_\beta \psi\right] + V\psi \tag{B.82}$$

We can again reduce the Schrodinger equation $\hat{H}_2\psi = E\psi$ to a one-dimensional Schrodinger equation with the ansatz

$$\psi(\alpha, \gamma, \beta) = e^{im_\alpha \alpha} e^{im_\gamma \gamma} \frac{B(\beta)}{\sqrt{\sin\beta}}, \tag{B.83}$$

yielding

$$-\hbar^2\lambda^4 \frac{B''(\beta)}{2} + U(\beta)B(\beta) = EB(\beta) \tag{B.84}$$

where

$$U(\beta) = -\frac{\hbar^2\lambda^4}{8} + \frac{\hbar^2\lambda^4}{2\sin^2\beta}\left[{m_\alpha}^2 + m_\gamma^2 - 2m_\alpha m_\gamma \cos\beta - \frac{1}{4}\right] - 2k^2\lambda s\cos\beta. \tag{B.85}$$

This can be studied by standard techniques for periodic potentials (Floquet theory or Bloch waves etc.) We content ourselves with a quick look at low energy excitations: small oscillations around the classical equilibrium $q = 0$ and setting $m_\alpha = 0 = m_\gamma$. Changing variables $\beta = \hbar\lambda^2 q$ and expanding around the classical minimum $q = 0$ gives

$$-\frac{1}{2}\frac{d^2B}{dq^2} + \left\{q^2\left(\hbar^2k^2s\lambda^5\right) - \frac{1}{8q^2} - 2k^2s\right\}B \approx EB \tag{B.86}$$

The solutions involve Laguerre polynomials and the spectrum is, in this approximation $E_n \approx \sqrt{2}(2n+1)\hbar k\sqrt{s}\lambda^{\frac{5}{2}}$. If we remove the zero-point energy ($n = 0$), we have the energy of n free particles each of energy $e_1 = \hbar k\sqrt{8s}\lambda^{\frac{5}{2}}$. This is the dispersion relation of massless particles, except for a rescaling of the speed.

B.5 Conclusions and outlook

Because they only exist in the short distance limit, it is difficult to say whether objects like "preons" we discuss could correspond to directly observable objects in an experiment. Quarks were not considered at first to be directly observable things either, as they could not be created as isolated particles. In the S_1-model's strong coupling limit, the Minkowski geometry of space-time appears to be lost, and wave propagation is sustained entirely by the non-linearity. However, these waves do not appear to transmit information, and perhaps any "post-relativistic" effects are hidden by some sort of confinement when they form bound states.

It is at least intriguing to question whether highly coupled theories have fundamental constituents whose nature is so exotic that they have been overlooked. Drawing paralells with $\lambda\phi^4$ theory, it is tempting to speculate that the Higgs particle of the standard model is such a composite of some strongly bound preons existing only at short distances. Were this the case, one could sensibly describe a "pure Higgs" at short distances.

For a more complete understanding, we must quantize the whole theory rather than just its mechanical reduction. Since the equations have a Lax pair, it should be possible to perform a canonical transformation to angle-variables and then pass to the quantum theory. Such a quantization was achieved for sine-Gordon theory [95], proving that the solitons are fermions which bind to form the scalar waves. A similar analysis of our model is a lengthy endeavor, and we hope to return to this later after laying the groundwork and motivation here.

We present a side by side comparison of comparison of our work with the two models in Table B.1 below.

Nilpotent Field Theory (S_1)		Principal Chiral Model (S_2)
$\mathcal{L}_1 = \mathrm{Tr}\left\{\frac{1}{2\lambda}\dot{\phi}^2 - \frac{1}{2\lambda}\phi'^2 + \frac{1}{3}\phi[\dot{\phi},\phi']\right\}$	Lagrangian density	$\mathcal{L}_2 = \frac{1}{2\lambda^2}\mathrm{Tr}\left\{(g^{-1}\dot{g})^2 - (g^{-1}g')^2\right\}$
$I = \frac{1}{\lambda}\dot{\phi}, \quad J = \phi'$	Currents	$I = g^{-1}g', \quad J = g^{-1}\dot{g}$
$\lambda I' = \dot{J}$	Current Identity	$\dot{I} - \frac{1}{\lambda}J' + \lambda\,[J,I] = 0$
$\dot{I} - \frac{1}{\lambda}J' + \lambda\,[J,I] = 0$	Equation of Motion	$\lambda I' = \dot{J}$
Reduced System of S_1		Reduced System of S_2
$\phi(t,x) = e^{Kx}R(t)e^{-Kx} + mKx$	Wave Ansatz	$g(t,x) = e^{\lambda sKx}Q(t)e^{-Kx}$
$I = e^{Kx}\dot{R}e^{-Kx}$ $J = e^{Kx}\{\lambda[K,R] + mK\}\,e^{-Kx}$	Wave Currents	$I = e^{Kx}\left\{\lambda sQ^{-1}KQ - K\right\}e^{-Kx}$ $J = e^{Kx}\{Q^{-1}\dot{Q}\}e^{-Kx}$
$S = \dot{R} + \frac{1}{\lambda}K$ $L = [K,R] + mK$	Common Variables	$S = sQ^{-1}KQ$ $L = \frac{1}{\lambda}Q^{-1}\dot{Q}$
$\dot{L} = [K,S]$ $\dot{S} = \lambda[S,L]$	Current Identity Eqn. of motion	$\dot{S} = \lambda[S,L]$ $\dot{L} = [K,S]$
$H_1 = \mathrm{Tr}\left\{\frac{1}{2}\left(S - \frac{1}{\lambda}K\right)^2 + \frac{1}{2}L^2\right\}$	Hamiltonian $H_1 = H_2$	$H_2 = \mathrm{Tr}\left\{\frac{1}{2}\left(S - \frac{1}{\lambda}K\right)^2 + \frac{1}{2}L^2\right\}$
$\mathcal{L}_1 = \mathrm{Tr}\Big\{\frac{1}{2}\left(S - \frac{1}{\lambda}K\right)^2$ $+\frac{1}{3}R\,[S-K,L] - \frac{1}{2}L^2\Big\}$	Lagrangian $\mathcal{L}_1 \neq \mathcal{L}_2$	$\mathcal{L}_2 = \frac{1}{2}\mathrm{Tr}\left\{L^2 - (S-K)^2\right\}$
$E \sim \lvert k\rvert^{2/3} \quad (\lambda \to \infty)$	short-range dispersion	$E \sim \lvert k\rvert \quad (\lambda \to 0)$

Table B.1: A comparison of results in the dual models

Appendix C

Measurement of the top quark Yukawa coupling from $\mathrm{t\bar{t}}$ kinematic distributions in the dilepton final state at $\sqrt{s} = 13$ TeV

Partially reprinted from publicly available preliminary result in Ref. [79] with permission. Co-authored with the CMS collaboration

Since the discovery of the Higgs boson in 2012 [10, 11], one of the main goals of the CERN LHC program has been to study in detail the properties of this new particle. In the standard model (SM), all fermions acquire their mass through interaction with the Higgs field. More specifically, the mass of fermions, m_{f}, arises from a Yukawa interaction with coupling strength $g_{\mathrm{f}} = \sqrt{2}m_{\mathrm{f}}/v$, where v is the vacuum expectation value of the Higgs field. For the purpose of this measurement, we define for the top quark the parameter $Y_t = g_t/g_t^{SM}$, which is equivalent to the modifier κ_t defined in the κ–framework [105].

Among all such couplings, the top quark Yukawa coupling is of particular interest. It is not only the largest, but also remarkably close to unity. Given the measured top quark mass $m_t = 172.4 \pm 0.5\,\text{GeV}$ [106], the mass-Yukawa relation implies a value of the Yukawa coupling $g_t^{SM} \approx 0.99$. However, physics beyond the SM, such as two-Higgs-doublet and composite Higgs models, introduce modified couplings that alter the interaction bestween the top quark and the Higgs field [47,48]. This makes the top–Higgs interaction one of the most interesting properties of the Higgs field to study today at the LHC, especially because it is experimentally accessible through multiple avenues, both direct and indirect.

Recent efforts have seen the first success in directly probing g_t via $t\bar{t}H$ production [49,50], while the most precise determination comes from the $\kappa-$framework fit in Ref. [51], which yields $Y_t = 0.98 \pm 0.14$ by combining information from several Higgs boson decay channels. These measurements, however, fold in assumptions of SM branching fractions via Higgs couplings to other particles. Another way to constrain g_t, which does not depend on the Higgs coupling to other particles, was presented in the search for four top quark production in Ref. [78], yielding a 95% confidence limit of $Y_t < 1.7$. However, it is also possible to constrain g_t indirectly using the kinematic distributions of reconstructed $t\bar{t}$ pair events, a technique which recently yielded a similar 95% confidence limit of $Y_t < 1.67$ [76] in the lepton+jets decay channel from a data set with an integrated luminosity $36\,\text{fb}^{-1}$. The measurement presented in this note follows closely the latter approach.

Current simulations of $t\bar{t}$ production include next-to-leading-order (NLO) precision in perturbative quantum chromodynamics (QCD). Subleading-order correc-

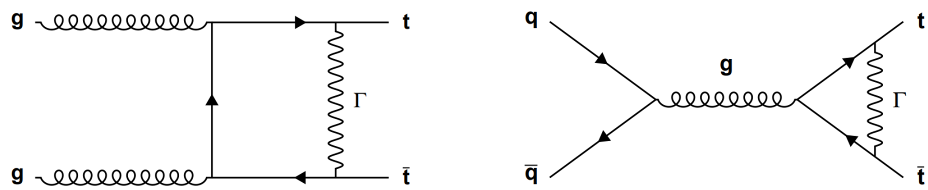

Figure C.1: Sample diagrams for EW contributions to gluon-induced and quark-induced top quark pair production, where Γ stands for neutral gauge bosons or the Higgs boson. (repeated from page 87)

tions arise from including electroweak (EW) terms in a perturbative expansion of the strong coupling α_S and the EW coupling constant α. Such terms begin to noticeably affect the cross section only at loop-induced order $\alpha_s^2\alpha$, and are typically ignored. While these terms have a very small effect on the total cross section, they can alter the shape of kinematic distributions to a measurable extent. Such changes will be more noticeable if the Yukawa coupling affecting the loop correction (Fig. C.1) is anomalously large. Therefore, these corrections are of particular interest in deriving upper limits on g_{t}. For example, the distribution of the invariant mass of the $\mathrm{t\bar{t}}$ system, $M_{\mathrm{t\bar{t}}}$, will be affected significantly by varying Y_{t}. In particular, doubling the value of Y_{t} can alter the $M_{\mathrm{t\bar{t}}}$ distribution by about 9% near the $\mathrm{t\bar{t}}$ production threshold, as described in Ref. [52] and shown in Fig. C.2 (left). Another variable sensitive to the value of Y_{t} is the difference in rapidity between top quarks and antiquarks, $\Delta y_{\mathrm{t\bar{t}}} = y(\mathrm{t}) - y(\mathrm{\bar{t}})$, shown in Fig. C.2 (right). Together, $M_{\mathrm{t\bar{t}}}$ and $\Delta y_{\mathrm{t\bar{t}}}$ are proxies for the Mandelstam kinematic variables s and t respectively, and can be used to include the EW corrections on arbitrary samples via weight corrections.

The effect of EW corrections can thus be evaluated without generating entirely new Monte Carlo (MC) simulations. After deriving the dependence of these corrections on the parameter of interest (POI) Y_{t}, a measurement is performed. In order for this analysis to be sensitive to deviations in Y_{t} from unity arising from new

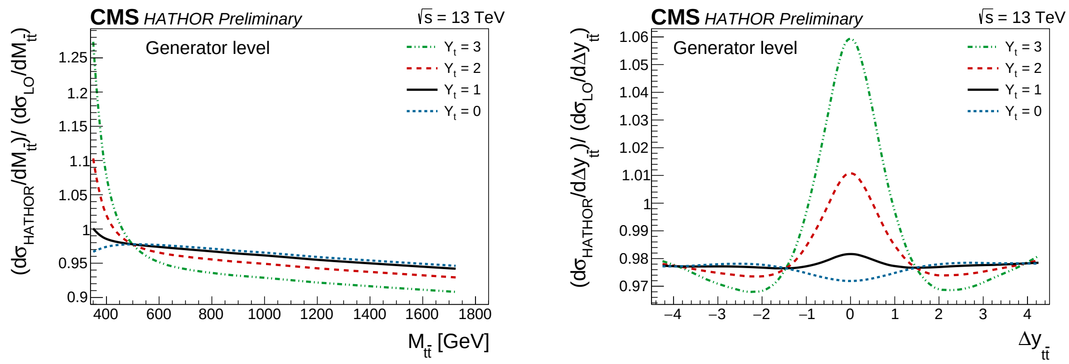

Figure C.2: Effect of the EW corrections on $t\bar{t}$ differential kinematic distributions for different values of Y_t, after reweighting of simulated events. (repeated from page 88)

physics beyond the SM, we avoid correcting our simulated data for any potential mismodeling with unclear physical origin, such as the discrepancy in top quark p_T distributions observed in Ref. [74]). We target events in the dilepton final state, for which this type of measurement has not yet been performed. While this decay channel has a smaller branching fraction than the lepton+jets channel studied in Ref. [76], it has much lower backgrounds due to the distinguishing presence of two final-state high-p_T leptons. However, two neutrinos are also expected in this final state, which escape detection and pose challenges in kinematic reconstruction. For this reason, we do not perform a full kinematic reconstruction as was done in the previous measurement in the lepton+jets channel. This measurement also utilizes the full data set of 137 fb^{-1} collected during Run 2 at the LHC from 2016 to 2018, allowing us to achieve comparable precision from a decay channel with a much lower branching rate.

In this note, we will first discuss the data and MC samples (Section C.1), followed by the methods for event selection (Section C.2) and reconstruction (Section C.3). We then present an outline of the measurement technique (Section C.4) and the contributing uncertainty sources (Section C.5), and conclude with the results of the measurement (Section C.6).

C.1 Simulation of top quark pair production and backgrounds

The production of $t\bar{t}$ events is simulated at the matrix-element (ME) level with NLO QCD precision, using the POWHEG generator [54–57]. This calculation is performed with the renormalization and factorization scales, μ_r and μ_f, set to the transverse top quark mass, $m_T = \sqrt{m_t^2 + {p_T}^2}$, where p_T is the transverse momentum of the top quark. The default value of m_t is set to 172.5 GeV in all simulations. The matrix-element calculations obtained from POWHEG are combined with parton shower (PS) simulation from PYTHIA8 [58–60], using the underlying-event tune M2T4 [61] to simulate data taken in 2016, and tune CP5 [62] to simulate data taken in 2017 and 2018 . The parton distribution functions (PDFs) set NNPDF3.0 at NLO [63] is used in 2016 and updated to NNPDF3.1 [64] at next-to-NLO (NNLO) in 2017 and 2018. These samples are normalized to a $t\bar{t}$ cross section calculated at NNLO in QCD including resummation of next-to-next-to-leading logarithmic (NNLL) soft gluon terms with TOP++2.0 [39].

A high purity of $t\bar{t}$ events can be obtained in the dilepton channel. For the estimation of the remaining small fraction of backgrounds, we rely on simulations. In particular, we account for dilepton production due to Drell–Yan type processes and single top quark production. Other SM processes, in particular W boson and diboson productions, were investigated and found to have negligible contributions.

About 1% of the events identified as $\mathrm{t\bar{t}}$ dilepton decays are misidentified $\mathrm{t\bar{t}}$ lepton+jets decays. EW corrections are applied to all $\mathrm{t\bar{t}}$ events, even misidentified ones, so their kinematic distributions remain dependent on Y_{t}. Thus, these events are still considered as signal even though the quality of reconstruction may generally be poor.

Single top quark events are simulated at NLO with POWHEG in combination with PYTHIA8, while diboson events are simulated with PYTHIA8 with leading-order (LO) QCD precision. Drell–Yan production is simulated at LO using the MADGRAPH generator [67] interfaced to PYTHIA8 using the MLM matching algorithm [68].

The detector response to all simulated events is modeled with the GEANT4 software toolkit [46]. In addition, the effects of multiple pp interactions per event are included in simulations and the number of these pileup interactions agrees with what is observed in data.

C.1.1 Simulation of electroweak corrections

Contributions to the top quark pair production arising from QCD+EW diagrams are evaluated using the HATHOR (v2.1) package [53]. A double-differential cross section is computed as a function of $M_{\mathrm{t\bar{t}}}$ and $\Delta y_{\mathrm{t\bar{t}}}$ at order $\alpha_S^2\alpha$. These diagrams involve massive gauge boson exchange and are shown in Fig. C.1. The contributions from photon-mediated interactions, which are comparatively small [107], are ignored. Formally, these will be referred to as EW corrections, though certain electromagnetic interactions are not included.

The ratio of this double-differential cross section is evaluated with respect to the LO QCD computation, in order to obtain a multiplicative weight correction $w(M_{\mathrm{t\bar{t}}}, \Delta y_{\mathrm{t\bar{t}}})$. Applying this weight at parton level to MC samples produced at NLO QCD approximates the inclusion of EW corrections in the simulation. This *multiplicative approach* to including EW corrections was used previously in Ref. [76], and has the benefit of approximating the inclusion of diagrams at order $\mathcal{O}(\alpha_S^3\alpha)$. Because EW corrections factorize in some kinematic regimes, this is a better-motivated approach than the alternative *additive approach*, in which one adds the fixed-order result at order $\mathcal{O}(\alpha_S^2\alpha)$ while ignoring all potential contributions of order $\mathcal{O}(\alpha_S^3\alpha)$ (see Ref. [65] for a detailed discussion). In other words, the additive approach applies a EW correction factor only to the proportion of POWHEG events present at LO QCD, while any interplay between EW corrections and higher-order QCD simulation is ignored. Though the multiplicative approach is clearly favored, neither approach can account for the effects of two-loop contributions near the $\mathrm{t\bar{t}}$ production threshold. Accordingly, we take the difference between the two predictions as a modeling uncertainty in this regime. The implementation of this uncertainty is discussed further in Section C.4.

The EW correction weights are calculated for discrete integer values of $Y_{\mathrm{t}} = 0, 1, \ldots, 5$. Since the dependence of the production rate on Y_{t} is exactly quadratic, these discrete values are sufficient to parametrize event yields as a continuous function of Y_{t} (see Section C.4). This allows us to measure which value of Y_{t} best describes the data.

C.2 Event and object selection

Events are selected using single-electron or muon-triggers. Data taking at CMS is punctuated by technical stops at the end of each year, leading to minor changes in configuration and modeling between 2016, 2017, and 2018. For events selected by the single-electron trigger, we require a trigger p_{T} threshold of 27 GeV with the exception of 2018, where a threshold of 32 GeV is used. In the case of the single-muon trigger, we select events with a trigger p_{T} threshold of 24 GeV, which is raised to 27 GeV in 2017 due to high event rates.

We ensure that all electrons and muons are within the silicon tracker coverage by requiring a pseudorapidity $\mid \eta \mid< 2.4$. To operate well above the trigger threshold we then require at least one isolated electron or muon reconstructed with $p_{\mathrm{T}} >$ 30 GeV, except in 2018, where we require leading p_{T} electrons to have $p_{\mathrm{T}} > 34$ GeV in accordance with the trigger availability. After selecting the leading-p_{T} lepton, a second isolated electron or muon with $p_{\mathrm{T}} > 20$ GeV is required. Events with three or more isolated leptons with $p_{\mathrm{T}} > 15$ GeV are discarded.

Jets are clustered from PF objects through the anti-$\mathrm{k_T}$algorithm [69, 70] using a distance parameter of 0.4. The jet momentum is calculated as the vectorial sum of the momenta of its constituents. Corrections to the jet energy are derived as a function of jet p_{T} and η in simulation and improved by measurements of energy balance in data [71]. We select jets with $\mid \eta \mid< 2.4$ and $p_{\mathrm{T}} > 30$ GeV. Jets originating from quarks are identified using the DeepCSV algorithm [41]. We require events to have at least two jet candidates passing the algorithm's loose working point, using a tagging criterion with an efficiency of 84%. This is the most efficient and least stringent working point of the DeepCSV algorithm. Events with more than two jets

Source	2016 (X fb^{-1})	2017 (Y fb^{-1})	2018 (Z fb^{-1})	All (X+Y+Z fb^{-1})	% total MC
$t\bar{t}$ MC	140 831±134	170 554±95	259 620±146	571 005±220	96.1%
V + jets MC	1944±54	3178±75	4958±133	10 080±162	1.7%
Single t MC	3022±25	3521±24	5827±32	12 370±47	2.1%
Diboson MC	142±9	151±10	250±15	543±20	0.1%
Total MC	145 940±148	177 404±124	270 655±201	593 999±279	100%
Data	144 817	178 088	264 791	587 696	98.9%

Table C.1: MC and data event yields for all three years and combined. Statistical uncertainty on simulated event counts is given. For an estimate of systematic uncertainty, see Figs. 8.5-8.7. (repeated from page 96)

meeting this criterion are considered only if exactly two of the jets pass one of the more stringent medium or tight working points supplied by the algorithm, which have efficiencies of 68% and 50%, respectively. In this case, the two jets passing the most stringent identifier are assumed to be the two b jets originating from top quark decays. We find that in $t\bar{t}$ events, the two jets are among the candidates passing the looser -tagging requirement in roughly 93% of the events.

In order to remove Drell–Yan background events in the ee and $\mu\mu$ channels, we reject events in which the two leptons together have an invariant mass below 50 GeV or within 10 GeV around the Z boson mass of 91.2 GeV.

The negative vector sum of all transverse momenta, $\vec{\not{E}}_T$, is generally of large magnitude in dilepton decays because of the two undetected final-state neutrinos. To further aid in removing Drell–Yan type decays, we impose an additional selection requirement on the magnitude of the missing transverse momentum vector of $p_{\mathrm{T}}^{\mathrm{miss}} > 30\,\mathrm{GeV}$ in all events with ee or $\mu\mu$ in the final state.

The breakdown of expected signal and background yields is shown by year in Table C.1. The Drell–Yan background is estimated to be about 2%. Single top quark production accounts for roughly another 2% of the estimated sample composition.

C.3 Event reconstruction

The EW corrections are calculated based on $M_{\mathrm{t\bar{t}}}$ and $\Delta y_{\mathrm{t\bar{t}}}$. However, to evaluate these quantities it is necessary to reconstruct the full kinematic properties of the $\mathrm{t\bar{t}}$ system, including the two undetected neutrinos. While it is possible to completely reconstruct the neutrino momenta in the on-shell approximation, such a reconstruction is highly sensitive to $p_{\mathrm{T}}^{\mathrm{miss}}$, which introduces large resolution effects and additional systematic uncertainties. We found that using the proxy variables $M_{\mathrm{b}\ell} = M(+ \bar{} + \ell + \bar{\ell})$ and $|\Delta y|_{\mathrm{b}\ell} = |y(+ \bar{\ell}) - y(\bar{} + \ell)|$, where ℓ represents a final state electron or muon, results in a more sensitive measurement.

Unlike $M_{\mathrm{b}\ell}$, the accurate reconstruction of $|\Delta y|_{\mathrm{b}\ell}$ requires that each of the two jets is matched to the correct lepton, i.e., both originating from the same top quark decay. In order to make this pairing, we utilize the information from the kinematic constraints governing the neutrino momenta.

If one assumes the top quarks and W bosons to be on-shell, the neutrino momenta are constrained by a set of quadratic equations arising from the conservation of four-momentum at each vertex. We refer to these kinematic equations, collectively, as the mass constraint. The mass constraint for each top quark decay results in a continuum of possible solutions for neutrino momenta, which geometrically can be presented as an intersection of ellipsoids in three-dimensional momentum-space [108]. For certain values of input momenta of jets and leptons these ellipsoids do not intersect at all, such that the quadratic equations have no real solution. In these scenarios the mass constraint cannot be satisfied.

In cases where the mass constraint can be satisfied, one may also constrain $p_{\mathrm{T}}^{\mathrm{miss}}$ in the event to equal the sum of the p_{T} of the two undetected neutrinos. We call this the $p_{\mathrm{T}}^{\mathrm{miss}}$ constraint. This constraint reduces the remaining solutions to a discrete set, containing either 2 or 4 possibilities that fully specify the momenta of both neutrinos. Similar to the case of the mass constraint, there are some values of the input parameters, for which the $p_{\mathrm{T}}^{\mathrm{miss}}$ constraint cannot be satisfied exactly.

Assuming the jets are correctly reconstructed and paired, the mass constraint can be satisfied in 96% of all cases, while the mass and $p_{\mathrm{T}}^{\mathrm{miss}}$ constraints can both be satisfied in 55% of cases. In contrast, if the jets are correctly reconstructed but incorrectly paired to leptons, the mass constraint can be satisfied in only 23% of cases, while both mass and $p_{\mathrm{T}}^{\mathrm{miss}}$ constraints can be satisfied in only 18% of cases.

Pairings with no solution to the mass constraint are thus frequently incorrect. When the mass constraint can be satisfied, pairings with a solution to the $p_{\mathrm{T}}^{\mathrm{miss}}$ constraint are more likely to be correct. This information is used as part of the pairing procedure, which proceeds in three steps.

1. The mass constraint is checked for both possible pairings. If only one pairing is found to satisfy the mass constraint, that pairing is used. If both pairings fail to satisfy the mass constraint, the event is discarded. If both pairings satisfy the mass constraint, we check the $p_{\mathrm{T}}^{\mathrm{miss}}$ constraint.

2. If only one pairing allows for the $p_{\mathrm{T}}^{\mathrm{miss}}$ constraint while the other does not, the pairing yielding an exact solution to the $p_{\mathrm{T}}^{\mathrm{miss}}$ constraint is used.

3. If the neutrino kinematics do not suggest a clear pairing, the -jets, $_1$ and $_2$, are paired with the leptons (ℓ, $\bar{\ell}$) by minimizing the quantity

$$\Sigma_{1(2)} = \Delta R(_{1(2)}, \ell) + \Delta R(_{2(1)}, \bar{\ell})$$

among the two possible pairings, where $\Delta R(,\ell)^2 = (\eta_\mathrm{b} - \eta_\ell)^2 + (\varphi_\mathrm{b} - \varphi_\ell)^2$.

After these steps, we obtain the correct jet pairing in simulation in 82% of the events for which both jets originating from top quark decays are correctly identified.

The sensitivity of our chosen kinematic variables to Y_t, before and after reconstruction, is shown in Fig. C.3. We see that, in the chosen proxy variables, not much sensitivity is lost in the reconstruction process. This is especially true for the proxy mass observable, $M_{\mathrm{b}\ell}$, providing an advantage over $M_{\mathrm{t\bar{t}}}$, which cannot be reconstructed as accurately.

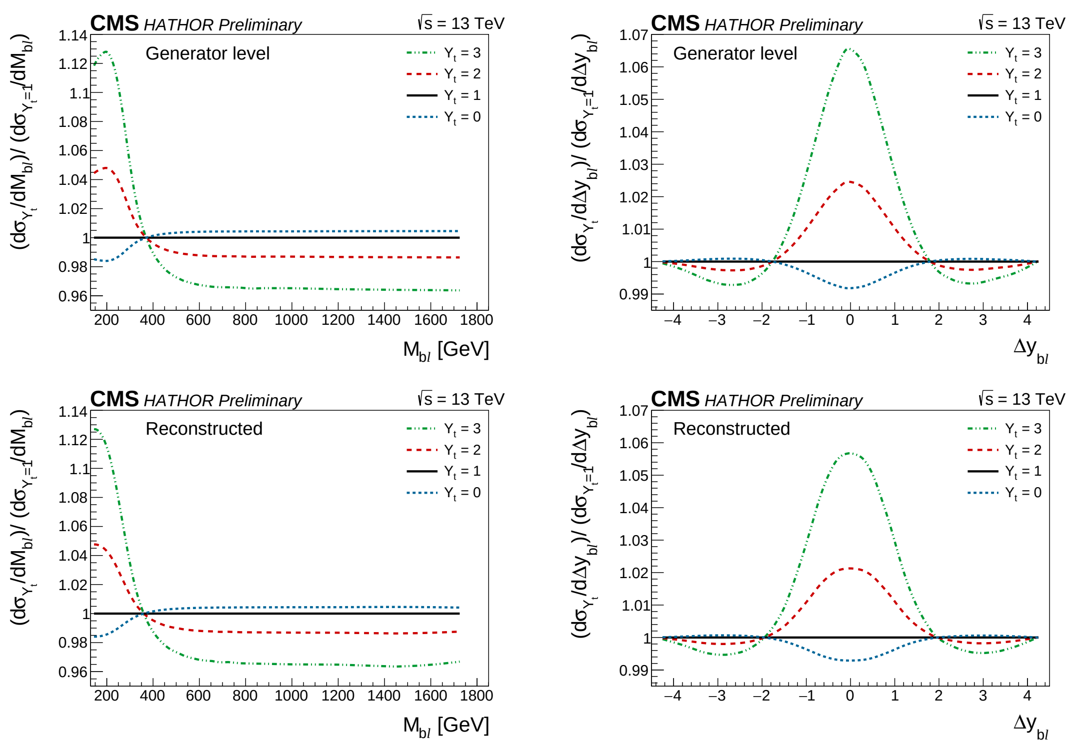

Figure C.3: The ratio of reconstructed kinematic distributions with EW corrections (evaluated for various values of Y_t) to the SM kinematic distribution is shown, demonstrating the sensitivity of these distributions to the Yukawa coupling. The upper figures show the information at generator level, while the lower figures are obtained from reconstructed events. (repeated from page 99)

C.3.1 Comparison between data and simulation

Comparisons between data and simulation is shown in Fig. C.4, where the agreement is generally seen to be within the systematic uncertainty. Some slope is

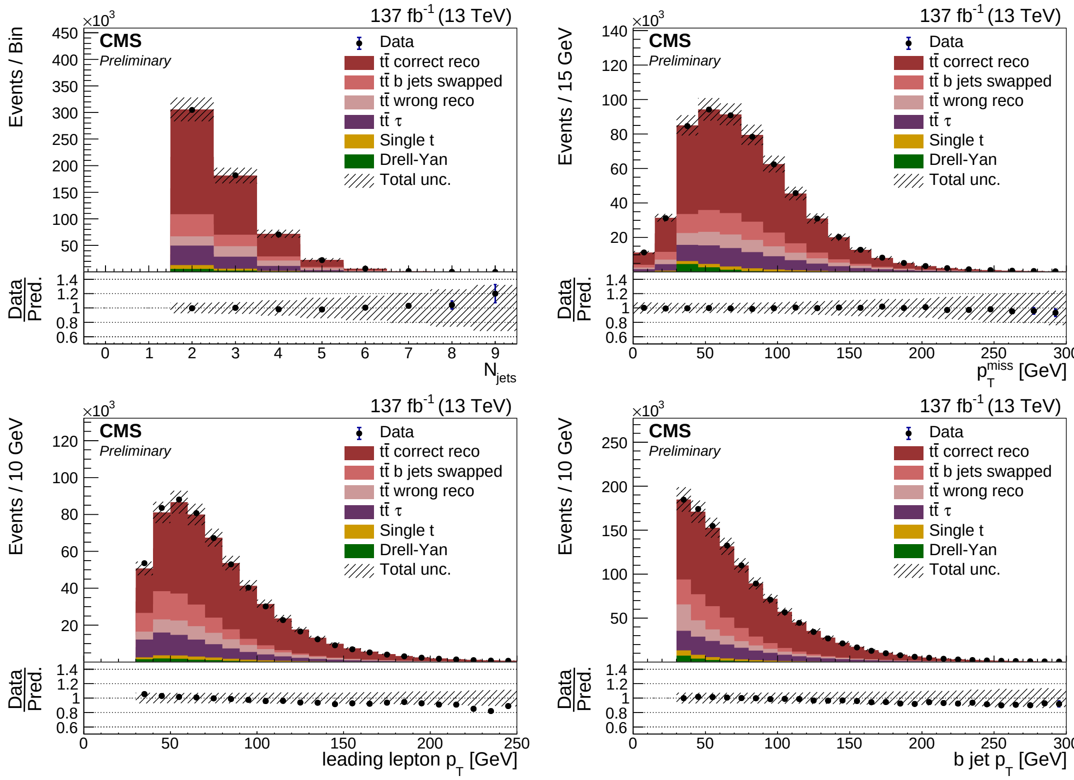

Figure C.4: Data to MC simulation comparison is shown for jet multiplicity, $p_{\mathrm{T}}^{\mathrm{miss}}$, lepton p_{T}, and b jet p_{T}. Uncertainty bands are derived by varying each known uncertainty source up and down across the full data-taking run, and summing the effects in quadrature. Here the signal simulation is divided into the following categories: correctly reconstructed events ($t\bar{t}$ correct reco), events with correctly reconstructed leptons and b jets but incorrect pairing of b jets and leptons ($t\bar{t}$ b jets swapped), events with incorrectly reconstructed b jets or leptons including non-dilepton $t\bar{t}$ decays ($t\bar{t}$ wrong reco), and a separate category for incorrectly reconstructed dilepton events which have a τ lepton in the final state ($t\bar{t}$ τ).

apparent in the ratio of data to MC prediction in the p_{T} distributions of lepton and jet. These trends may be related to a previously observed feature of the nominal POWHEG +PYTHIA8, in which a harder top quark p_{T} distribution is observed than in data (see, for example, Ref. [74]). Fixed-order NNLO calculations are available that generally show a softer top quark p_{T} spectrum than in the POWHEG +PYTHIA8 simulation, but these are typically evaluated with a different dynamical scale choice, as described in Ref. [109]. The dynamical scale m_{T} at which ME calculations are evaluated in POWHEG is generally preferred for modeling $M_{\mathrm{t\bar{t}}}$, which is more closely related to our measurement sensitivity. The potential effects of NNLO calculations and the top quark p_{T} spectrum on this measurement are discussed further in Section C.5.

C.4 Measurement strategy and statistical methods

After reconstruction, events are binned coarsely in $|\Delta y|_{\mathrm{b}\ell}$ and more finely in M_{bl}. The binning is chosen to ensure no fewer than $\approx$10 000 events in each bin and in each data-taking year, as seen in Fig. C.5, to ensure a low statistical uncertainty and improved uncertainty estimation. In each bin, the expected yield is parametrized as a function of Y_{t}. The effect is exactly quadratic, as a consequence of the order at which EW corrections are evaluated. Therefore we perform a quadratic fit to extrapolate the effect of the EW corrections on a given bin as a continuous function of Y_{t} (Fig. C.6). This correction for each bin can be applied as a rate parameter $R^{\mathrm{bin}}_{\mathrm{EW}}(Y_{\mathrm{t}}) = 1 + \delta_{\mathrm{EW}}(Y_{\mathrm{t}})$ affecting the expected bin content.

We construct a likelihood function $\mathcal{L}$, which is defined as a product of a Poisson probability describing statistical fluctuations in each bin and penalty terms encoding systematic uncertainty sources as nuisance parameters:

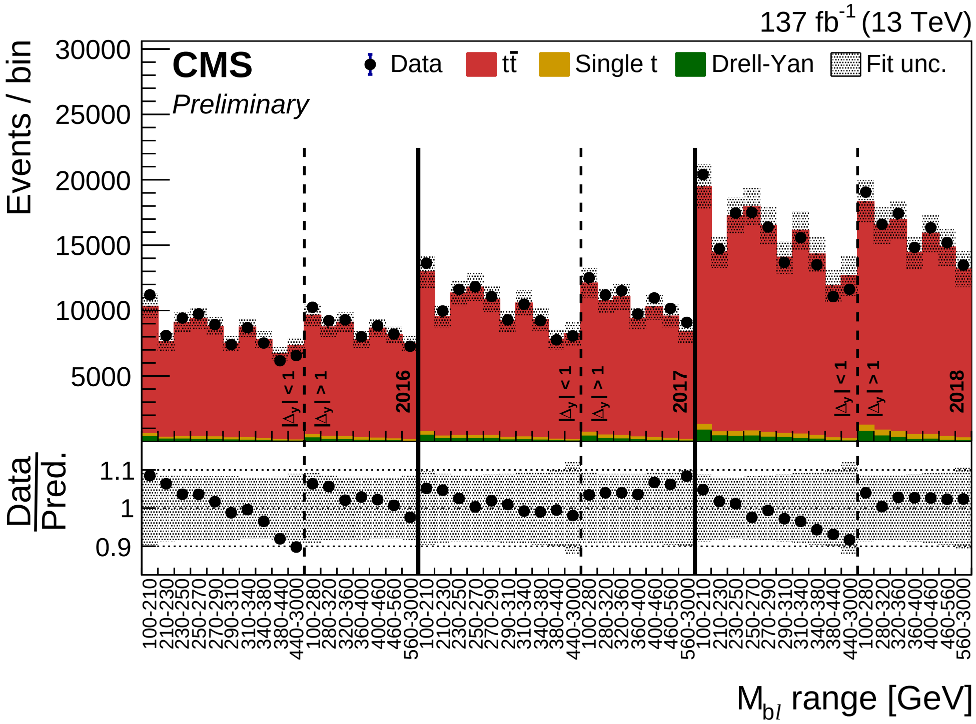

Figure C.5: Binned data and simulated events prior to performing the fit. The solid lines divide the three datataking periods, while the dashed lines divide the two $|\Delta y|_{b\ell}$ bins in each datataking period, with $M_{b\ell}$ bin ranges displayed on the x axis. (repeated from page 103)

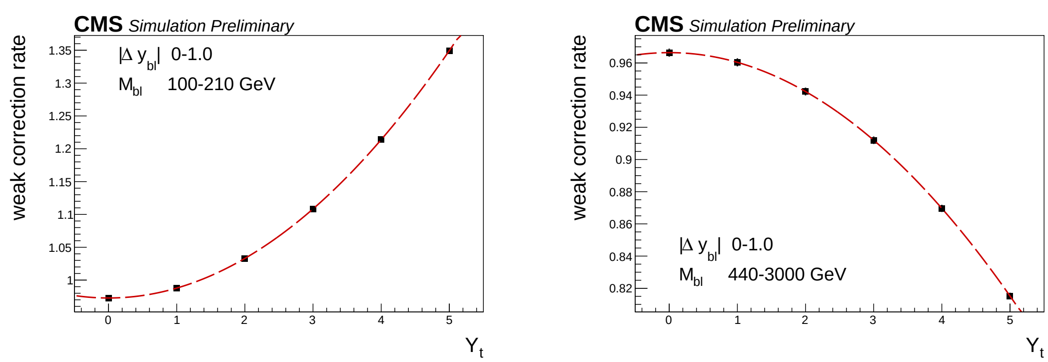

Figure C.6: The EW correction rate modifier $R^{\text{bin}}_{\text{EW}}$ in two separate $[M_{b\ell}, \Delta y_{b\ell}]$ bins from 2017 data, demonstrating the quadratic dependence on Y_t. All bins will have an increasing or decreasing quadratic yield function, with the steepest dependence on Y_t found at lower values of $M_{b\ell}$. (repeated from page 106)

$$\mathcal{L} = \prod_{\text{bin}\in(M_{\text{b}\ell},|\Delta y|_{bl})} \mathcal{L}_{\text{bin}} \times p(\phi) \times \prod_i p(\theta_i), \tag{C.1}$$

$$\mathcal{L}_{\text{bin}} = \text{Poisson}\left[n_{\text{obs}}^{\text{bin}} \middle| s^{\text{bin}}(\{\theta_i\}) \times R_{\text{EW}}^{\text{bin}}(Y_{\text{t}}, \phi) + b^{\text{bin}}(\{\theta_i\})\}\right]. \tag{C.2}$$

Here $n_{\text{obs}}^{\text{bin}}$ is the total observed bin count, with the expected bin count being the sum of the predicted signal yield s^{bin} and background yield b^{bin}. The number of expected signal events is modified by the additional rate parameter R_{EW}, which depends on the Yukawa coupling ratio Y_{t} and the special nuisance parameter ϕ, discussed further below. Each of the other nuisance parameters, $\{\theta_i\}$, that affect the expected counts in each bin, is described by a probability distribution function of $p(\theta_i)$ that describes the likelihood of the deviation of θ_i from its expected value. We include an estimate of the uncertainty in the multiplicative application of EW corrections derived at order $\mathcal{O}(\alpha_S^2\alpha)$. The full expression for the rate $R_{\text{EW}}^{\text{bin}}$, including this uncertainty term in the bins near the $\text{t}\bar{\text{t}}$ production threshold, is given by

$$R_{\text{EW}}^{\text{bin}}(Y_{\text{t}}, \phi) = (1 + \delta_{\text{EW}}^{\text{bin}}(Y_{\text{t}})) \times (1 + \delta_{\text{QCD}}^{\text{bin}}\delta_{\text{EW}}^{\text{bin}}(Y_{\text{t}}))^{\phi}, \tag{C.3}$$

where we have defined

$$\delta_{\text{EW}}^{\text{bin}} = \frac{n_{\text{HATHOR}}^{\text{bin}} - n_{\text{LO QCD}}^{\text{bin}}}{n_{\text{LO QCD}}^{\text{bin}}}, \qquad \delta_{\text{QCD}}^{\text{bin}} = \frac{n_{\text{POWHEG}}^{\text{bin}} - n_{\text{LO QCD}}^{\text{bin}}}{n_{\text{POWHEG}}^{\text{bin}}}. \tag{C.4}$$

Intuitively, δ_{EW} represents the marginal effect of EW corrections included in HATHOR relative to the LO QCD calculation, while δ_{QCD} represents the marginal effect of higher-order terms included in the POWHEG sample relative to the LO QCD calculation. The multiplicative approach to including EW corrections assumes that these two corrections factorize. The quantity $\delta_{\mathrm{QCD}}^{\mathrm{bin}}\delta_{\mathrm{EW}}^{\mathrm{bin}}$ represents the cross term arising from the difference in multiplicative and additive approaches. The normally distributed nuisance parameter ϕ modulates the uncertainty generated by this cross term. We note that the uncertainty in the EW corrections is unique because it depends on the value of Y_{t} at which the EW corrections are evaluated. Thus, it is described by its own term and nuisance parameter ϕ separate from other systematic uncertainties. For bins away from the threshold where EW corrections decrease as a function of Y_{t}, we do not include this uncertainty. These bins do not contribute much sensitivity to the measurement, and also enter a kinematic regime in which this method of uncertainty estimation is no longer meaningful. At the large values of the Mandelstam variable s that correspond to these bins, the dominant terms contributing to δ_{EW} are Sudakov logarithms resulting from W and Z exchange. These terms factorize well and do not contribute to the uncertainty we wish to model [65].

Nuisance parameters corresponding to uncertainties that have an overall normalization effect, that have an overall normalization, such as uncertainty on luminosity or cross section values, follow a log-normal distribution. Uncertainties with shape effects associated to nuisance parameters $\{\theta_i\}$ are handled by generating up and down variations of the bin content s^{bin} for each θ_i. These variations result from changing the underlying theoretical/experimental sources, which are outlined in Section C.5, usually by an amount of one standard deviation (σ) based on the uncertainty in our best estimates. Since θ_i is normally distributed, these shapes are then enforced as the bin modifiers associated with $\theta_i = \pm 1$, while $\theta_i = 0$ corresponds to the nominal estimate where the simulated data is reproduced. The collection of bin modifiers for these up and down variations are referred to as shape templates, with examples shown in Section C.6. A vertical template morphing is applied to alter the shape as a function of the underlying nuisance parameter θ_i, where in each bin the modifier is interpolated as a cubic spline for values of $\theta_i \in [-1, 1]$ and linearly outside of that region, assuring that $s^{\text{bin}}(\theta_i)$ remains continuous and differentiable.

The measurement of Y_t is then performed by a maximum likelihood fit. By repeating this maximization over a fine array of fixed values of Y_t and comparing to the absolute maximum, we evaluate confidence intervals (CI) at the 68% and 95% level around the best-fit value.

C.5 Sources of experimental and theoretical uncertainty

The list of uncertainties considered is very similar to that of the previous measurement presented in Ref. [76]. The main differences are the lack of QCD multijet background and the use of data from the full Run 2 data-taking period. Full or partial correlations are imposed on the underlying uncertainty sources between data-taking periods where appropriate, as discussed further in Section C.5.1. Uncertainties which strictly do not alter the shape of final distribution are treated as normalization uncertainties, while all others are treated as shape uncertainties on the binned data. Shape effects are considered for the distributions of $\text{t}\bar{\text{t}}$ events only, as the contribution of background events is small. Correlations of uncertainties between different data-taking periods are treated on a case-by-case basis. Because the measurement is more sensitive to shape effects than normalization effects, the dominant uncertainties are not necessarily those with the largest magnitude of effect.

The dominant experimental uncertainty in this analysis comes from the calibration of the detector's jet energy response. Corrections to the reconstructed jet energies are applied as a function of p_{T} and η. We follow the standard approach outlined in Ref. [71] to consider 26 separate uncertainties which are involved in determining these calibrations. In this approach, uncertainty in the resolution of the jet reconstruction is also considered in addition to the energy response. The effect of these uncertainties is propagated to the reconstruction of $p_{\mathrm{T}}^{\mathrm{miss}}$.

Other experimental sources of uncertainty are comparatively minor. The overall uncertainty in the integrated luminosity of 2.5%, 2.3%, and 2.5% is included as a normalization uncertainty applied to all signal and background events in 2016, 2017, and 2018 respectively [110–112]. The uncertainty in the number of pileup events included in simulation is assessed by varying the inelastic cross section, 69.2 mb, by 4.6% [75].

Efficiencies in jet identification and misidentification are corrected to match data [41]. The uncertainty in this correction manifests approximately as an overall normalization effect of around 3%.

Similarly, scale factors are applied in bins of p_{T} and η to correct simulated efficiencies of lepton reconstruction, identification, isolation, and triggers to match data. These are derived from a fit using the tag-and-probe method on Z boson decays [72, 73]. This fit accounts for the uncertainty from the limited number of events in the data sample as well as differences in performance based on the jet multiplicity. Overall, the effect is assessed to be below 2%.

As a standard technique to estimate the contributions of higher-order QCD terms at the ME level, the QCD renormalization scale μ_r and factorization scale μ_f are each varied separately up and down in the POWHEG simulation by a factor of 2. The same-sign variation of the two scales together is included as an additional uncertainty. Since an NNLO $t\bar{t}$ cross section is already used to improve the normalization of the MC simulation, the normalization effect induced by the scale variations is overestimated. As our measurement is predominantly sensitive to shape effects, this overall normalization effect is therefore removed in generating the corresponding shape templates.

Because a higher-order cross section calculation is used, the uncertainty on sample normalization coming from scale variations is fictitiously large. In generating the corresponding shape templates, this overall normalization effect is removed, as the NNLO $t\bar{t}$ cross section is already used to improve MC normalization, and we are predominantly sensitive to shape effects. The shape effects remain significant and these are among the dominant uncertainties in the fit.

A 5% normalization uncertainty is assumed in the $t\bar{t}$ normalization, which covers expected contributions from higher-order terms not included in the NNLO+NNLL cross section calculation [39]. The backgrounds in this analysis are small enough ($\approx$2% sample composition each) that we do not generate shape templates for their response to individual systematic uncertainties. A 15% normalization uncertainty is included on single top quark MC samples, which was seen to cover expected ME scale variation and jet energy correction uncertainties associated with these samples. Drell–Yan and diboson MC samples are assigned a 30% normalization uncertainty, to cover the larger ME scale variation uncertainties associated with these LO simulations. The background normalizations can alter slightly the expected shape of the data, but are not among the dominant uncertainties.

We include an uncertainty in the EW corrections, based on our methods for generating and applying these additional terms, as outlined in Sections C.1 and C.4. Like the scale variations, this uncertainty is designed to cover higher-order effects at the ME level, specifically those arising from diagrams of order $\alpha_s^3\alpha$. This is a dominant uncertainty in the fit, and in particular it reduces sensitivity to values of $Y_\mathrm{t} < 1$.

As mentioned in Section C.3.1, fixed-order calculations are available for top quark pair production at NNLO, and are generally in better agreement with the top quark p_T spectrum observed in data. Some measurements have included additional uncertainties or reweighting in attempting to account for this effect. To assess the potential effects of NNLO calculations on this measurement, differential cross sections in top quark p_T as well as $M_{\mathrm{t\bar{t}}}$ were studied using publicly available FASTNLO tables [113, 114]. Multidifferential cross sections at NNLO [115] in $M_{\mathrm{t\bar{t}}}$ and $\Delta y_{\mathrm{t\bar{t}}}$ were also studied, as these are the same variables through which EW corrections are applied in this measurement. The NNLO calculation in Ref. [113] does not show a significantly softer top quark p_T spectrum than the corresponding NLO calculation evaluated with the same choice of dynamical scale. This suggests that the differences in the top quark p_T spectrum between the POWHEG +PYTHIA8 and the calculation in Ref. [113] arise partly from the choice of dynamical scale and not only from higher-order effects. Meanwhile, cross section calculations as a function of $M_{\mathrm{t\bar{t}}}$ appear consistent with existing uncertainties. For this reason, no additional uncertainties are included with the explicit attempt to reflect knowledge of higher-order simulations, as this would begin to double-count the existing theory uncertainties.

The uncertainty in modeling the initial- and final-state radiation in the parton shower algorithm is assessed by varying the value of the renormalization scales in the initial- and final-state radiation by a factor of two. These are among the dominant modeling uncertainties in the measurement. Uncertainty for other parameters in the parton shower description are considered separately. The h_{damp} parameter, which controls the matrix-element parton shower matching in POWHEG + PYTHIA8, is set to the nominal value of $h_{\mathrm{damp}} = 1.58\,m_{\mathrm{t}}$ (1.39 m_{t}) in 2016 (2017–2018). Dedicated MC samples are generated with this parameter varied down to $1\,m_{\mathrm{t}}$ ($0.874\,m_{\mathrm{t}}$) and up to $2.24\,m_{\mathrm{t}}$ ($2.305\,m_{\mathrm{t}}$) in 2016 (2017-18), in order to estimate the effect of this uncertainty. Dedicated MC samples are also generated with variations of the PYTHIA8 underlying-event tune, a set of parameters which impact the hadronization. The uncertainty in h_{damp} and the underlying event tune are minor compared to the parton shower scale variations.

Dedicated MC samples are generated with the top quark mass varied up and down by 1 GeV from the nominal value $m = 172.5$ GeV to estimate the effect of the top quark mass uncertainty. This is a significant uncertainty, but not among the most dominant. It should be noted that, although the mass and Yukawa coupling are generally treated as independent in this measurement, varying the mass will slightly modify the definition of $Y_{\mathrm{t}} = 1$. However, this effect, which is below 1%, is much smaller than the sensitivity of the measurement and can therefore be ignored.

The NNPDF sets [63] contain 100 individual variations as uncertainties. Following the approach in Ref. [76], similar variations are combined to reduce the number of variations to a more manageable set of 10 shape templates. The variation of the strong coupling α_s used by NNPDF is treated separately from the other PDF variations. The effect of uncertainties in the PDF set is typically smaller than 1%, and has a minor impact on the fit.

The branching fraction of semileptonic B decays affects the jet response. The effect of varying this quantity within its measured uncertainty [36] is included as an uncertainty, which has a small effect relative to other modeling uncertainties.

The momentum transfer from quarks to B hadrons is modeled with a transfer function dependent on $x_b = \frac{p_{\mathrm{T}}(B)}{p_{\mathrm{T}}(-\mathrm{jet})}$. To estimate the uncertainty, the transfer function is varied up and down within uncertainty of the Bowler–Lund parameter [77] in PYTHIA8. The resulting effect is included by modifying event weights to reproduce the appropriate transfer function. This has a noticeable shape effect of the order 4%, but was not found to be a dominant uncertainty in the fit.

C.5.1 Processing of systematics

In this analysis, the effect of the POI Y_{t} manifests itself as a smooth shape distortion of the kinematic distributions as shown in Fig. C.7. Though the nuisance parameters describing the sources of uncertainty should induce smooth shape effects as well, their effects are sometimes obscured by statistical noise or imprecise methods of estimation. This is noticeable for the uncertainties associated with jet energy scale, jet energy resolution, parton shower modeling, pileup reweighting, and top quark mass. For these shape templates only, we apply a one-iteration LOWESS algorithm [116] to smooth the templates and remove fluctuations that may disturb the fit. The underlying-event tune and h_{damp} uncertainties in the parton showering

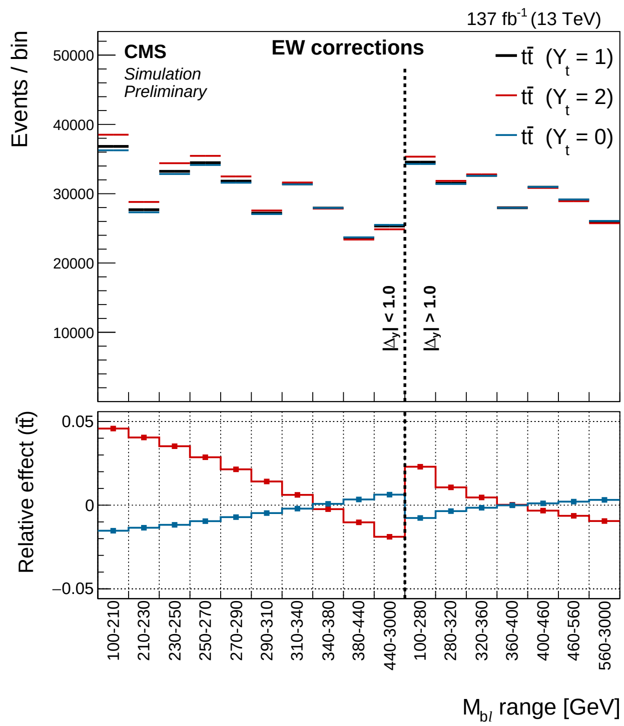

Figure C.7: The effect of the Yukawa parameter Y_t on reconstructed events in the final binning. The POI induces a shape distortion on the kinematic distributions. (repeated from page 105)

are small enough for their shapes disappear into statistical noise, and are therefore treated only as normalization uncertainties.

Most shape templates are also symmetrized, by taking the larger magnitude of the up and down variations in each bin and negating this value to replace the effect of the opposite variation. This step helps ensure a stable minimum in the likelihood fit, but is skipped for templates whose natural shape effect is notably asymmetric. In the few cases where this may be an overly conservative approach, it nonetheless helps to ensure the performance and reliability of the minimization procedure, and has little effect on the final result.

Full or partial correlations between the 2016, 2017, and 2018 data samples are assumed for many uncertainties. In general, all theoretically motivated uncertainties are considered fully correlated between years. Exceptions are made in cases where modeling methods were altered between years, in particular for initial-state radiation, final-state radiation, and PDF uncertainties. The modeling of these uncertainties differs in the 2016 simulation, so the associated nuisance parameter in this year is either partially or fully decorrelated from the other years. Additionally, uncertainties whose effects disappear into statistical noise due to limited MC sample size (underlying-event tune, h_{damp}) are converted to uncorrelated normalization uncertainties.

Some experimental uncertainties can be broken into components which are either fully correlated or uncorrelated between years (large jet energy scale contributions, luminosity). The uncertainty in the number of pileup events is considered fully correlated as it is evaluated by varying the total inelastic cross section. For minor uncertainties from jet and lepton scale factors, which have both correlated and statistical components, a 50% correlation is assumed between years. Lastly, the jet energy resolution uncertainties are treated as uncorrelated between years.

C.6 Results

We obtain a best fit value of $Y_t = 1.16$, with a 68% CI of $[0.81, 1.40]$, and a 95% CI of $[0, 1.54]$. The scan of the negative log likelihood as a function of Y_t, used to build these intervals, is shown in Fig. C.8. We also show the agreement of data and simulation after performing the fit in Fig. C.9. The maximum of the likelihood occurs at a configuration with good agreement between data and simulation. The shape templates for the four uncertainties with the greatest effect on the fit are shown in Fig. C.10.

This result is in agreement with the previously obtained measurement in the lepton+jets final state in Ref. [76], while obtaining a slight increase in sensitivity. Using a different decay channel and a larger data set provides a measurement complementary to the previous result.

C.7 Summary

A measurement of the Higgs Yukawa coupling to the top quark is presented, based on data from proton-proton collisions at the CMS experiment. Data at a center-of-mass energy 13 TeV is analyzed from the full LHC Run 2, collected from 2016–2018, corresponding to an integrated luminosity of 137 fb^{-1}. The result is a best fit value of $Y_t = 1.16^{+0.24}_{-0.35}$ and a 95% confidence interval (CI) of $Y_t \in [0, 1.54]$. This

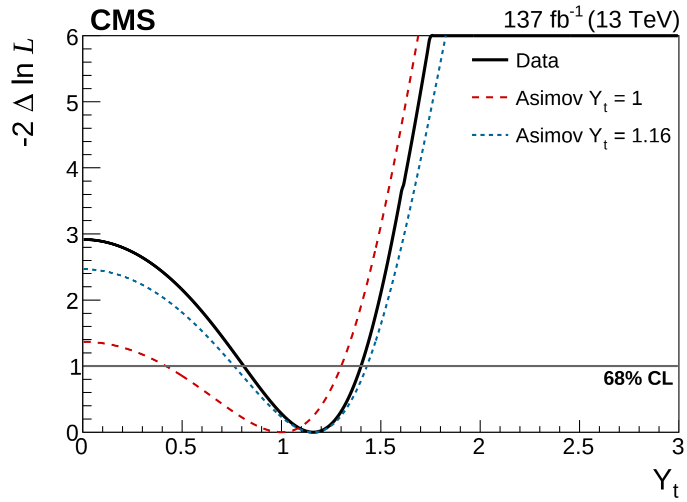

Figure C.8: The result of a likelihood scan, performed by fixing the value of Y_t at values over the interval [0, 3]. Expected curves from fits on simulated data are shown produced at the SM value $Y_t = 1.0$ (dashed) and at the final best-fit value of $Y_t = 1.16$ (dotted). (repeated from page 115)

measurement uses the effects of virtual Higgs boson exchange on $t\bar{t}$ kinematic properties to extract information about the coupling from kinematic distributions. Although the sensitivity is lower compared to constraints obtained from studying processes involving Higgs boson production in Refs. [49] and [51], this measurement avoids dependence on other Yukawa coupling values through additional branching assumptions, making it a compelling orthogonal measurement. This measurement also achieves a slightly higher precision than the only other Y_t measurement that does not make additional branching fraction assumptions, performed in the search for production of four top quarks. The four top quark search places a 95% CI of $Y_t < 1.7$ [78] while this measurement achieves a 95% CI of $Y_t < 1.54$.

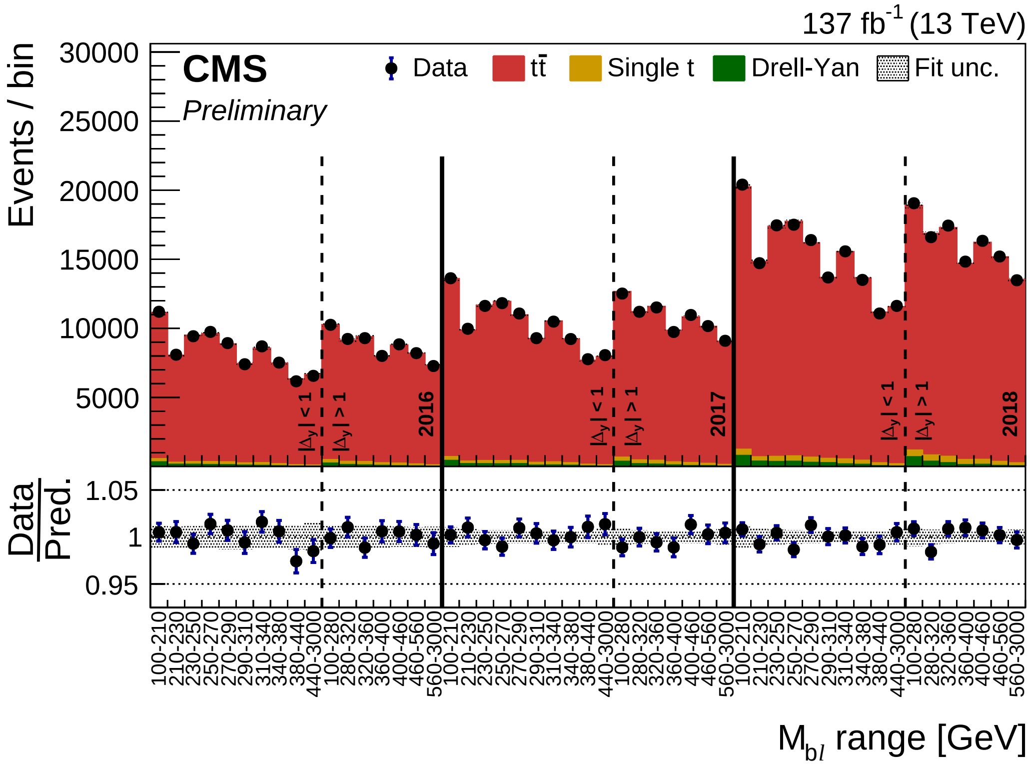

Figure C.9: The agreement between data and MC simulation at the best fit value of $Y_t = 1.16$ after performing the likelihood maximization, with shaded bands displaying the post-fit uncertainty. The solid lines divide the three datataking periods, while the dashed lines divide the two $|\Delta y|_{b\ell}$ bins in each datataking period, with $M_{b\ell}$ bin ranges displayed on the x axis. (repeated from page 115)

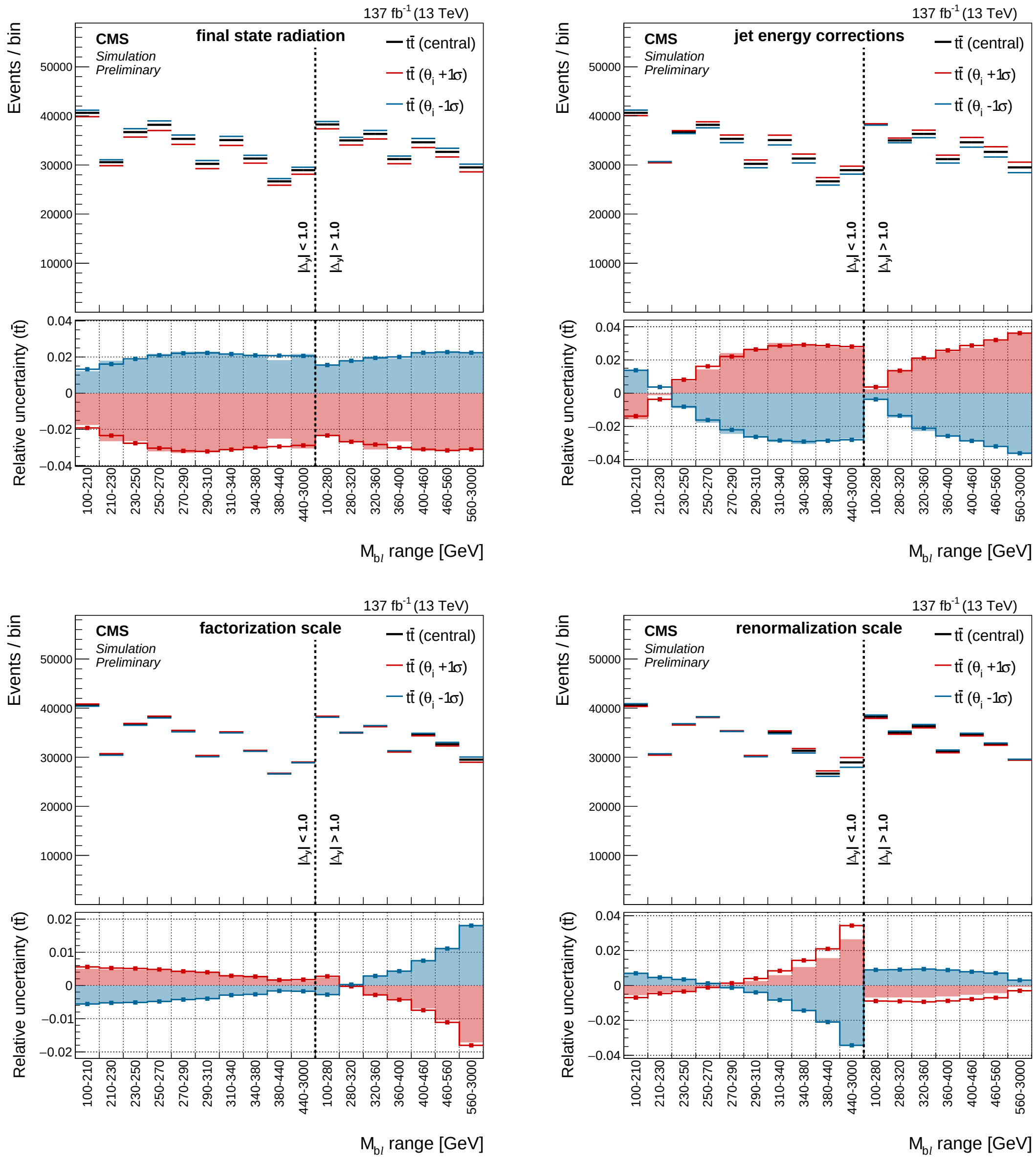

Figure C.10: Shape templates are shown for the uncertainties associated with the final-state radiation in PYTHIA8, the jet energy corrections, the factorization scale, and the renormalization scale. Along with the intrinsic uncertainty included on the EW corrections, these are seen to be the dominant uncertainties in the fit. The shaded bars represent the raw template information, while the lines show the shapes after smoothing and symmetrization procedures have been applied. In the fit, the jet energy corrections are split into 26 different components, but for brevity only the total uncertainty is shown here. Variation between years is minimal for each of these uncertainties, though they are treated separately in the fit. (repeated from page 112)

Bibliography

[1] A. Einstein, "Über einem die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt", *Annalen der physik* **4** (1905).

[2] C. Jönsson, "Elektroneninterferenzen an mehreren künstlich hergestellten Feinspalten", *Zeitschrift für Physik* **161** (1961), no. 4, 454–474.

[3] P. G. Merli, G. Missiroli, and G. Pozzi, "On the statistical aspect of electron interference phenomena", *Am. J. Phys* **44** (1976), no. 3, 306–307.

[4] D. Hanneke, S. F. Hoogerheide, and G. Gabrielse, "Cavity control of a single-electron quantum cyclotron: Measuring the electron magnetic moment", *Physical Review A* **83** (2011), no. 5, 052122.

[5] T. Aoyama, M. Hayakawa, T. Kinoshita et al., "Tenth-order electron anomalous magnetic moment: Contribution of diagrams without closed lepton loops", *Physical Review D* **91** (2015), no. 3, 033006.

[6] H. Georgi, "Lie algebras in particle physics: from isospin to unified theories", volume 54. Westview press, 1999.

[7] A. Zee, "Group theory in a nutshell for physicists", volume 17. Princeton University Press, 2016.

[8] H. Yukawa, "On the interaction of elementary particles. I", *Proceedings of the Physico-Mathematical Society of Japan. 3rd Series* **17** (1935) 48–57.

[9] C.-S. Wu, E. Ambler, R. Hayward et al., "Experimental test of parity conservation in beta decay", *Physical review* **105** (1957), no. 4, 1413.

[10] CMS Collaboration, "Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC", *Phys. Lett. B* **716** (2012) 30, `doi:10.1016/j.physletb.2012.08.021`, `arXiv:1207.7235`.

[11] 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. B* **716** (2012) 1, `doi:10.1016/j.physletb.2012.08.020`, `arXiv:1207.7214`.

[12] Y. Sofue and V. Rubin, "Rotation curves of spiral galaxies", *Annual Review of Astronomy and Astrophysics* **39** (2001), no. 1, 137–174.

[13] Supernova Cosmology Project Collaboration, "Measurements of Ω and Λ from 42 high-redshift supernovae", *The Astrophysical Journal* **517** (1999), no. 2, 565.

[14] High-Z SN Search Team Collaboration, "Observational evidence from supernovae for an accelerating universe and a cosmological constant", *The Astronomical Journal* **116** (1998), no. 3, 1009.

[15] K. G. Wilson, "The renormalization group and critical phenomena", *Reviews of Modern Physics* **55** (1983), no. 3, 583.

[16] A. Kar, "Renormalization from classical to quantum physics", *University of Rochester PhD Thesis* (2014).

[17] S. Aoki, Y. Aoki, D. Bečirević et al., "Review of lattice results concerning low-energy particle physics", *The European Physical Journal C* **77** (2017), no. 2, 112.

[18] "Millenium Problems". `https://www.claymath.org/millennium-problems`.

[19] R. D. Peccei, "The strong CP problem and axions", in *Axions*, pp. 3–17. Springer, 2008.

[20] M. Dine and A. Kusenko, "Origin of the matter-antimatter asymmetry", *Reviews of Modern Physics* **76** (2003), no. 1, 1.

[21] A. R. Liddle, "An introduction to cosmological inflation", *High energy physics and cosmology* (1998) 260.

[22] D. G., "The pre-LHC Higgs hunt", *Phil. Trans. R. Soc. A.* **373** (2015) 20140039, `doi:10.1098/rsta.2014.0039`.

[23] A. Blondel, "Precision electroweak physics at LEP", `doi:10.1016/0920-5632(94)90655-6`.

[24] T. A. Collaboration, the DELPHI Collaboration, the L3 Collaboration et al. *Physics Reports* **427** (May, 2006) 257–454, `doi:10.1016/j.physrep.2005.12.006`, `arXiv:hep-ex/0509008`.

[25] M. Hairer, "Regularity structures and the dynamical Φ^4_3 model", *arXiv preprint arXiv:1508.05261* (2015).

[26] R. J. Adler, "An Introduction to Continuity, Extrema, and Related Topics for General Gaussian Processes", *Lecture Notes-Monograph Series* **12** (1990) i–155.

[27] B. Duplantier, R. Rhodes, S. Sheffield et al., "Log-correlated Gaussian fields:

an overview", *arXiv preprint arXiv:1407.5605* (2014).

[28] S. Rajeev and E. Ranken, "Continuity of scalar fields with logarithmic correlations", *Physical Review D* **92** (2015), no. 4, 045021.

[29] A. Kar and S. Rajeev, "Non-Riemannian metric emergent from scalar quantum field theory", *Physical Review D* **86** (2012), no. 6, 065022.

[30] E. Ogievetsky, N. Reshetikhin, and P. Wiegmann, "The principal chiral field in two dimensions on classical lie algebras: The Bethe-ansatz solution and factorized theory of scattering", *Nuclear Physics B* **280** (1987) 45 – 96, `doi:https://doi.org/10.1016/0550-3213(87)90138-6`.

[31] A. Polyakov and P. Wiegmann, "Goldstone fields in two dimensions with multivalued actions", *Physics Letters B* **141** (1984), no. 3, 223 – 228, `doi:https://doi.org/10.1016/0370-2693(84)90206-5`.

[32] S. Rajeev and E. Ranken, "Highly nonlinear wave solutions in a dual to the chiral model", *Physical Review D* **93** (2016), no. 10, 105016.

[33] G. S. Krishnaswami and T. R. Vishnu, "On the Hamiltonian formulation, integrability and algebraic structures of the Rajeev-Ranken model", *Journal of Physics Communications* **3** (feb, 2019) 025005, `doi:10.1088/2399-6528/ab02a9`.

[34] G. S. Krishnaswami and T. Vishnu, "Invariant tori, action-angle variables, and phase space structure of the Rajeev-Ranken model", *Journal of Mathematical Physics* **60** (2019), no. 8, 082902.

[35] "Constraint on the matter–antimatter symmetry-violating phase in neutrino oscillations", *Nature* **580** (Apr, 2020) 339–344,

doi:10.1038/s41586-020-2177-0.

[36] Particle Data Group Collaboration, “Review of Particle Physics”, *Phys. Rev. D* **98** (2018) 030001, doi:10.1103/PhysRevD.98.030001.

[37] D0 Collaboration Collaboration, “Observation of the Top Quark”, *Phys. Rev. Lett.* **74** (Apr, 1995) 2632–2637, doi:10.1103/PhysRevLett.74.2632.

[38] CDF Collaboration Collaboration, “Observation of Top Quark Production in $\bar{p}p$ Collisions with the Collider Detector at Fermilab”, *Phys. Rev. Lett.* **74** (Apr, 1995) 2626–2631, doi:10.1103/PhysRevLett.74.2626.

[39] M. Czakon and A. Mitov, “Top++: A program for the calculation of the top-pair cross-section at hadron colliders”, *Comput. Phys. Commun.* **185** (2014) 2930, doi:10.1016/j.cpc.2014.06.021, arXiv:1112.5675.

[40] M. Czakon, D. Heymes, and A. Mitov, “High-Precision Differential Predictions for Top-Quark Pairs at the LHC”, *Physical Review Letters* **116** (Feb, 2016) doi:10.1103/physrevlett.116.082003.

[41] CMS Collaboration, “Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV”, *JINST* **13** (2018) P05011, doi:10.1088/1748-0221/13/05/P05011, arXiv:1712.07158.

[42] L. Evans and P. Bryant, “LHC Machine”, *Journal of Instrumentation* **3** (aug, 2008) S08001–S08001, doi:10.1088/1748-0221/3/08/s08001.

[43] G. Roland, K. Šafařík, and P. Steinberg, “Heavy-ion collisions at the LHC”, *Progress in Particle and Nuclear Physics* **77** (2014) 70–127.

[44] CMS Collaboration, “Particle-flow reconstruction and global event

description with the CMS detector", *JINST* **12** (2017) P10003, doi:10.1088/1748-0221/12/10/P10003, arXiv:1706.04965.

[45] CMS Collaboration, "The CMS experiment at the CERN LHC", *JINST* **3** (2008) S08004, doi:10.1088/1748-0221/3/08/S08004.

[46] GEANT4 Collaboration, "GEANT4—a simulation toolkit", *Nucl. Instrum. Meth. A* **506** (2003) 250, doi:10.1016/S0168-9002(03)01368-8.

[47] G. C. Branco, P. M. Ferreira, L. Lavoura et al., "Theory and phenomenology of two-Higgs-doublet models", *Phys. Rept.* **516** (2012) 1, doi:10.1016/j.physrep.2012.02.002, arXiv:1106.0034.

[48] K. Agashe, R. Contino, and A. Pomarol, "The Minimal composite Higgs model", *Nucl. Phys. B* **719** (2005) 165, doi:10.1016/j.nuclphysb.2005.04.035, arXiv:hep-ph/0412089.

[49] CMS Collaboration, "Observation of $\mathrm{t\bar{t}H}$ production", *Phys. Rev. Lett.* **120** (2018) 231801, doi:10.1103/PhysRevLett.120.231801, arXiv:1804.02610.

[50] ATLAS Collaboration, "Observation of Higgs boson production in association with a top quark pair at the LHC with the ATLAS detector", *Phys. Lett. B* **784** (2018) 173, doi:10.1016/j.physletb.2018.07.035, arXiv:1806.00425.

[51] CMS Collaboration, "Combined measurements of Higgs boson couplings in proton-proton collisions at $\sqrt{s} = 13$ TeV", *Eur. Phys. J. C* **79** (2019) 421, doi:10.1140/epjc/s10052-019-6909-y, arXiv:1809.10733.

[52] J. H. Kühn, A. Scharf, and P. Uwer, "Weak interactions in top-quark pair production at hadron colliders: an update", *Phys. Rev. D* **91** (2015) 014020,

doi:10.1103/PhysRevD.91.014020, arXiv:1305.5773.

[53] M. Aliev, H. Lacker, U. Langenfeld et al., "HATHOR: HAdronic Top and Heavy quarks crOss section calculatoR", *Comput. Phys. Commun.* **182** (2011) 1034, doi:10.1016/j.cpc.2010.12.040, arXiv:1007.1327.

[54] P. Nason, "A new method for combining NLO QCD with shower Monte Carlo algorithms", *JHEP* **11** (2004) 040, doi:10.1088/1126-6708/2004/11/040, arXiv:hep-ph/0409146.

[55] S. Frixione, P. Nason, and C. Oleari, "Matching NLO QCD computations with parton shower simulations: the POWHEG method", *JHEP* **11** (2007) 070, doi:10.1088/1126-6708/2007/11/070, arXiv:0709.2092.

[56] S. Alioli, P. Nason, C. Oleari et al., "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX", *JHEP* **06** (2010) 043, doi:10.1007/JHEP06(2010)043, arXiv:1002.2581.

[57] J. M. Campbell, R. K. Ellis, P. Nason et al., "Top-pair production and decay at NLO matched with parton showers", *JHEP* **04** (2015) 114, doi:10.1007/JHEP04(2015)114, arXiv:1412.1828.

[58] T. Sjöstrand, S. Mrenna, and P. Skands, "PYTHIA 6.4 physics and manual", *JHEP* **05** (2006) 026, doi:10.1088/1126-6708/2006/05/026, arXiv:hep-ph/0603175.

[59] T. Sjöstrand, S. Mrenna, and P. Z. Skands, "A brief introduction to PYTHIA 8.1", *Comput. Phys. Commun.* **178** (2008) 852, doi:10.1016/j.cpc.2008.01.036, arXiv:0710.3820.

[60] R. Frederix and S. Frixione, "Merging meets matching in MC@NLO", *JHEP*

12 (2012) 061, doi:10.1007/JHEP12(2012)061, arXiv:1209.6215.

[61] P. Skands, S. Carrazza, and J. Rojo, "Tuning PYTHIA 8.1: the Monash 2013 tune", *Eur. Phys. J. C* **74** (2014) 3024, doi:10.1140/epjc/s10052-014-3024-y, arXiv:1404.5630.

[62] CMS Collaboration, "Extraction and validation of a new set of CMS PYTHIA8 tunes from underlying-event measurements", *Eur. Phys. J. C* **80** (2020) 4, doi:10.1140/epjc/s10052-019-7499-4, arXiv:1903.12179.

[63] NNPDF Collaboration, "Parton distributions for the LHC Run II", *JHEP* **04** (2015) 040, doi:10.1007/JHEP04(2015)040, arXiv:1410.8849.

[64] NNPDF Collaboration, "Parton distributions from high-precision collider data", *Eur. Phys. J. C* **77** (2017) 663, doi:10.1140/epjc/s10052-017-5199-5, arXiv:1706.00428.

[65] M. Czakon, D. Heymes, A. Mitov et al., "Top-pair production at the LHC through NNLO QCD and NLO EW", *JHEP* **10** (2017) 186, doi:10.1007/JHEP10(2017)186, arXiv:1705.04105.

[66] D. Pagani, I. Tsinikos, and M. Zaro, "The impact of the photon PDF and electroweak corrections on $t\bar{t}$ distributions", *Eur. Phys. J. C* **76** (2016) 479, doi:10.1140/epjc/s10052-016-4318-z, arXiv:1606.01915.

[67] J. Alwall, R. Frederix, S. Frixione 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, doi:10.1007/JHEP07(2014)079, arXiv:1405.0301.

[68] J. Alwall, S. Höche, F. Krauss et al., "Comparative study of various

algorithms for the merging of parton showers and matrix elements in hadronic collisions", *Eur. Phys. J. C* **53** (2008) 473, doi:10.1140/epjc/s10052-007-0490-5, arXiv:0706.2569.

[69] M. Cacciari, G. P. Salam, and G. Soyez, "The anti-k_t jet clustering algorithm", *JHEP* **04** (2008) 063, doi:10.1088/1126-6708/2008/04/063, arXiv:0802.1189.

[70] M. Cacciari, G. P. Salam, and G. Soyez, "FastJet user manual", *Eur. Phys. J. C* **72** (2012) 1896, doi:10.1140/epjc/s10052-012-1896-2, arXiv:1111.6097.

[71] CMS Collaboration, "Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV", *JINST* **12** (2017) P02014, doi:10.1088/1748-0221/12/02/P02014, arXiv:1607.03663.

[72] CMS Collaboration, "Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s} = 13$ TeV", *JINST* **13** (2018) P06015, doi:10.1088/1748-0221/13/06/P06015, arXiv:1804.04528.

[73] CMS Collaboration, "Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at $\sqrt{s} = 8$ TeV", *JINST* **10** (2015) P06005, doi:10.1088/1748-0221/10/06/P06005, arXiv:1502.02701.

[74] CMS Collaboration, "Measurement of differential cross sections for the production of top quark pairs and of additional jets in lepton+jets events from pp collisions at $\sqrt{s} = 13$ TeV", *Phys. Rev. D* **97** (2018) 112003, doi:10.1103/PhysRevD.97.112003, arXiv:1803.08856.

[75] ATLAS Collaboration, "Measurement of the inelastic proton-proton cross section at $\sqrt{s} = 13$ TeV with the ATLAS detector at the LHC", *Phys. Rev. Lett.*

117 (2016) 182002, `doi:10.1103/PhysRevLett.117.182002`, `arXiv:1606.02625`.

[76] CMS Collaboration, "Measurement of the top quark Yukawa coupling from $t\bar{t}$ kinematic distributions in the lepton+jets final state in proton-proton collisions at $\sqrt{s} = 13$ TeV", *Phys. Rev. D* **100** (2019) 072007, `doi:10.1103/PhysRevD.100.072007`, `arXiv:1907.01590`.

[77] M. G. Bowler, "e^+e^- Production of heavy quarks in the string model", *Z. Phys. C* **11** (1981) 169, `doi:10.1007/BF01574001`.

[78] CMS Collaboration, "Search for standard model production of four top quarks with same-sign and multilepton final states in proton-proton collisions at $\sqrt{s} = 13\,\mathrm{TeV}$", *Eur. Phys. J. C* **78** (2018) 140, `doi:10.1140/epjc/s10052-018-5607-5`, `arXiv:1710.10614`.

[79] CMS Collaboration Collaboration, "Measurement of the top quark Yukawa coupling from $t\bar{t}$ kinematic distributions in the dilepton final state at $\sqrt{s} = 13$ TeV", Technical Report CMS-PAS-TOP-19-008, CERN, Geneva, (2020).

[80] V. I. Bogachev, "Gaussian measures". Number 62. American Mathematical Soc., 1998.

[81] M. Ledoux and M. Talagrand, "Probability in Banach Spaces: isoperimetry and processes", volume 23. Springer Science & Business Media, 2013.

[82] R. J. Adler, "An introduction to continuity, extrema, and related topics for general Gaussian processes", *Lecture Notes-Monograph Series* (1990) i–155.

[83] R. M. Dudley, "Sample functions of the Gaussian process", in *Selected Works*

of RM Dudley, pp. 187–224. Springer, 2010.

[84] F. Biagini, Y. Hu, B. Øksendal et al., "Stochastic calculus for fractional Brownian motion and applications". Springer Science & Business Media, 2008.

[85] S. Dodelson, "Modern cosmology". Academic press, 2003.

[86] D. Pickrell, "Invariant measures for unitary groups associated to Kac-Moody Lie algebras", volume 693. American Mathematical Soc., 2000.

[87] A. Sengupta, "The Yang-Mills measure for S 2", *Journal of functional analysis* **108** (1992), no. 2, 231–273.

[88] N. Charalambous and L. Gross, "The Yang-Mills heat semigroup on three-manifolds with boundary", *Communications in Mathematical Physics* **317** (2013), no. 3, 727–785.

[89] C. R. Nappi, "Some properties of an analog of the chiral model", *Phys. Rev. D* **21** (1980), no. 2, 418.

[90] C. R. Nappi and E. Witten, "Wess-Zumino-Witten model based on a nonsemisimple group", *Phys. Rev. Lett.* **71** (1993), no. 23, 3751.

[91] T. Curtright and C. Zachos, "Currents, charges, and canonical structure of pseudodual chiral models", *Phys. Rev. D* **49** (1994), no. 10, 5408.

[92] E. Ogievetsky, N. Reshetikhin, and P. Wiegmann, "The principal chiral field in two dimensions on classical lie algebras: the Bethe-ansatz solution and factorized theory of scattering", *Nucl. Phys. B* **280** (1987) 45–96.

[93] A. Polyakov and P. Wiegmann, "Goldstone fields in two dimensions with

multivalued actions", *Phys. Lett. B* **141** (1984), no. 3, 223–228.

[94] S. Coleman, "Quantum sine-Gordon equation as the massive Thirring model", *Phys. Rev. D.* **11** (1975), no. 8, 2088.

[95] L. D. Faddeev and V. E. Korepin, "Quantum theory of solitons", *Phys. Rep.* **42** (1978), no. 1, 1–87.

[96] C. G. Callan, D. Friedan, E. Martinec et al., "Strings in background fields", *Nucl. Phys. B* **262** (1985), no. 4, 593–609.

[97] J. E. Heebner, R. W. Boyd, and Q. H. Park, "Slow light, induced dispersion, enhanced nonlinearity, and optical solitons in a resonator-array waveguide", *Phys. Rev. E* **65** (2002), no. 3, 036619.

[98] R. Montgomery, "A tour of subriemannian geometries, their geodesics and applications". Number 91. American Mathematical Soc., 2006.

[99] A. Shapere and F. Wilczek, "Self-propulsion at low Reynolds number", *Phys. Rev. Lett.* **58** (1987), no. 20, 2051.

[100] L. Hormander, "Hypoelliptic second order differential equations", *Acta Math.* **119** (1967), no. 1, 147–171.

[101] E. Witten, "Non-abelian bosonization in two dimensions", *Comm. Math. Phys.* **92** (1984), no. 4, 455–472.

[102] M. Bos, "Dimensional regularization in the Wess-Zumino-Witten model", *Phys. Lett. B* **189** (1987), no. 4, 435–441.

[103] X. Zheng-Min, "Dimensional regularization and three-loop beta function of the Wess-Zumino-Witten model", *Nucl. Phys. B* **314** (1989), no. 1, 112–128.

[104] B. E. Baaquie and K. K. Yim, "Sigma model Lagrangian for the Heisenberg group", *Phys. Lett. B* **615** (2005), no. 1, 134–140.

[105] LHC Higgs Cross Section Working Group, "Handbook of LHC Higgs Cross Sections: 3. Higgs Properties", doi:10.5170/CERN-2013-004, arXiv:1307.1347.

[106] CMS Collaboration, "Measurement of the top quark mass using proton-proton data at $\sqrt{s}$ = 7 and 8 TeV", *Phys. Rev. D* **93** (2016) 072004, doi:10.1103/PhysRevD.93.072004, arXiv:1509.04044.

[107] W. Hollik and M. Kollar, "NLO QED contributions to top-pair production at hadron collider", *Phys. Rev. D* **77** (2008) 014008, doi:10.1103/PhysRevD.77.014008, arXiv:0708.1697.

[108] B. A. Betchart, R. Demina, and A. Harel, "Analytic solutions for neutrino momenta in decay of top quarks", *Nucl. Instrum. Meth. A* **736** (2014) 169, doi:10.1016/j.nima.2013.10.039, arXiv:1305.1878.

[109] M. Czakon, D. Heymes, and A. Mitov, "Dynamical scales for multi-TeV top-pair production at the LHC", *JHEP* **04** (2017) 071, doi:10.1007/JHEP04(2017)071, arXiv:1606.03350.

[110] CMS Collaboration, "CMS Luminosity Measurements for the 2016 Data Taking Period", CMS Physics Analysis Summary CMS-PAS-LUM-17-001, (2017).

[111] CMS Collaboration, "CMS luminosity measurement for the 2017 data-taking period at $\sqrt{s} = 13$ TeV", CMS Physics Analysis Summary CMS-PAS-LUM-17-004, (2018).

[112] CMS Collaboration, “CMS luminosity measurement for the 2018 data-taking period at $\sqrt{s} = 13$ TeV”, CMS Physics Analysis Summary CMS-PAS-LUM-18-002, (2019).

[113] M. Czakon, D. Heymes, and A. Mitov, “fastNLO tables for NNLO top-quark pair differential distributions”, arXiv:1704.08551.

[114] M. Czakon, D. Heymes, and A. Mitov, “High-precision differential predictions for top-quark pairs at the LHC”, *Phys. Rev. Lett.* **116** (2016) 082003, doi:10.1103/PhysRevLett.116.082003, arXiv:1511.00549.

[115] M. Czakon, S. Dulat, T.-J. Hou et al., “A study of the impact of double-differential top distributions from CMS on parton distribution functions”, arXiv:1912.08801.

[116] W. S. Cleveland, “Robust Locally Weighted Regression and Smoothing Scatterplots”, *Journal of the American Statistical Association* **74** (1979) 829, doi:10.1080/01621459.1979.10481038.